



## Draft Online Safety Bill

### Corrected oral evidence: Consideration of government's draft Online Safety Bill

Thursday 21 October 2021

9.50 am

Watch the meeting: <https://parliamentlive.tv/event/index/4df31e2e-50c7-4b92-a745-a5fddc227498>

Members present: Damian Collins MP (The Chair); Debbie Abrahams MP; Lord Black of Brentwood; Lord Clement-Jones; Lord Gilbert of Panteg; Baroness Kidron; Darren Jones MP; Lord Knight of Weymouth; John Nicolson MP; Lord Stevenson of Balmacara; Suzanne Webb MP.

Evidence Session No. 8

Heard in Public

Questions 135 - 142

### Witnesses

**I:** Professor Richard Wilson, Associate Dean for Faculty Development and Intellectual Life, Gladstein Chair and Professor of Anthropology and Law, University of Connecticut (appearing virtually by Zoom); Barbora Bukovská, Senior Director, Law and Policy, Article 19; Matthew d'Ancona, journalist, formerly of Index on Censorship, currently Editor at Tortoise Media (appearing virtually via Zoom); Silkie Carlo, Director, Big Brother Watch.

### USE OF THE TRANSCRIPT

1. This is an uncorrected transcript of evidence taken in public and webcast on [www.parliamentlive.tv](http://www.parliamentlive.tv).
2. Any public use of, or reference to, the contents should make clear that neither Members nor witnesses have had the opportunity to correct the record. If in doubt as to the propriety of using the transcript, please contact the Clerk of the Committee.
3. Members and witnesses are asked to send corrections to the Clerk of the Committee within 14 days of receipt.

## Examination of witnesses

Professor Richard Wilson, Barbora Bukovská, Matthew d'Ancona and Silkie Carlo.

Q135 **The Chair:** Good morning, and welcome to this further evidence session of the Joint Committee on the draft Online Safety Bill. We are grateful to our witnesses in the room for attending. We also have Matthew d'Ancona and Professor Richard Wilson joining remotely. Professor Wilson, we are very grateful to you, given that you are on the east coast and got up very early in the morning to give evidence to us today.

I will start with a question that is, in principle, one of the most important questions we have to consider. How do we draw the balance between freedom of speech and the harm that speech can cause other people, particularly in the context of social media? The harm might not be based on an individual posting but, rather, the systems effect that some people experience on social media online, where they are bombarded over a sustained period with individual pieces of content that on their own may not be unlawful, but the effect itself could be harmful. Could I ask each member of the panel for their opening thoughts on that question? Professor Wilson, perhaps you could start us off.

**Professor Richard Wilson:** Sure. Thanks very much for the invitation. The short answer is that I am not exactly sure if the Bill strikes a balance between freedom of speech and protection from harm. There are positive attributes, in so far as the Bill seeks to mitigate the harms of the internet that you describe, systemic harms. The central focus is on terrorist content and CSEA content, which is already illegal and fairly non-controversial. The protection of children online is laudable.

The concerns I have are that the protection from harm element in the Bill is extensively elaborated, whereas freedom of expression and privacy appear only as abstractions. It is unclear how those are reconciled. There is little guidance in the Bill, and in practice my concern is that the freedom of expression elements are likely to be ephemeral or easily swept aside in the actual implementation.

The devil is in the detail of the implementation. There is some vagueness in the Bill that may contribute to the misapplication of the regulatory powers of the Bill. By vagueness, I am referring to Section 11 and the "legal but harmful" content. That adds to my concerns. I would prefer a more precise definition of the harms in question, the content and the context in which their regulation is warranted. Protecting adults from psychological harm is the piece of this that I have the most concerns about.

Finally, there is a reliance on the service providers' risk assessments and their machine-learning algorithms. These algorithms do not yet integrate the social and political context in order to accurately assess what is legal but harmful content. This could lead to overzealous removal of legitimate speech and the limiting of freedom of expression.

In the literature, there are many examples of such algorithmic mistakes, including the removal of videos and images documenting war crimes in Syria. I also have concerns about the impact on minority communities—religious, ethnic and racial minorities. I can share all of this with you in a follow-up. One study found that AI models for processing hate speech were one and a half times more likely to flag tweets as offensive or hateful when written by African Americans.

I understand the need for the risk assessment model in the Bill. I think it is important, but it relies on algorithms that have been shown to be flawed and inaccurate at times. Those are my concerns.

**The Chair:** Perhaps I could ask you a follow-up question on that, which was all very helpful. I think you are right to identify the definition of legal but harmful as one of the hardest and, at the moment, probably least well-defined aspects of the legislation.

We have lots of laws that cover things like harassment, incitement and malicious communications. One of the things that would seem to be new about this is, as I said in my opening remarks, the sort of systems effect. The danger is that, if we try to come up with a tight legal definition of the systems effect, it could be a bit like the 1990s, when the British Government tried to define what a rave was for the purpose of public safety legislation. How do we crack that, do you think?

**Professor Richard Wilson:** You raise very valid points. My recommendation would be to go after the low-hanging fruit first of all. Plenty of incitement, harassment and co-ordinated campaigns of online abuse are already not currently regulated. Rather than going for the more difficult and grey area of disparaging and potentially racist or misogynist comments, there is plenty of material already on social media platforms and through search engines—you can access other content—that already breaks existing laws on incitement to terrorism and so on. My recommendation would be to go the more cautious route and suppress that which is already illegal.

**The Chair:** Thank you. We will go remote first, and then into the room. Matthew d’Ancona, thank you for joining us. I am interested in where you see we should draw the balance. I was also interested to read in your most recent book that when we used to talk about freedom of expression we used to think about people like Lech Walesa standing up to oppressive regimes. At the moment, the debate on freedom of expression seems to be about how abusive people can be towards each other.

**Matthew d’Ancona:** Thank you, Chair. I am very conscious of the context in which this Bill is being launched. As important as it is to frame a statute correctly, it is also important for lawmakers to think about the jurisprudence that has both given rise to it and to which it will give rise. I guess that is where my primary concerns lie.

First of all, it cannot be said often enough that the Bill is necessary. The scale and speed of the technological change that it seeks to address is

probably unprecedented in the history of human communications. Some form of measure or set of measures such as this is essential.

The question in your original point is how we do it without sacrificing key liberties, which, as your question implied, are not perhaps as cherished as they were in the late 1980s or early 1990s. I think that is partly because of the huge impact that the internet, and social media specifically, has had on the way in which we live. Worrying about harms has become more of a day-to-day activity than worrying about liberties, which is all the more reason for lawmakers to be careful that freedoms are not lost in concern about the immediate problem. One wants statute to be robust and resilient.

To echo the previous speaker, I think there is a genuine issue of both ethics and practical law in the idea of legal but harmful. I am not sure I really understand it. To the extent that I do understand it, it seems to me that it involves handing over the definition of harm, to a huge extent, to a group of very large tech companies. The definition in the Bill as it stands allows huge latitude around what constitutes harm.

I think we could all agree that speech or content that might have a foreseeable risk of causing a significant adverse physical impact is fair, but psychological impact seems to me to be dangerously elastic. It might be taken to include anything, from something that constituted harassment under the law already to an extract from fiction that an individual found unsettling or traumatising. I think that with words like "harm" and "safety" there is a slippage or a kind of semantic mission creep going on in their use. We used to talk about safety, and what we really meant was physical safety. Now, when people talk about safety, they often mean convenience or comfort. It is not the task of democratic legislators to make people feel comfortable. I think that is stretching the job description.

My main concern lies in that field. What exactly is content that is legal and harmful? To name another couple of concerns in that area, the Bill exempts journalism and comments below the line. There are difficulties there. Of course, some of the most brutal things that appear are comments below the line. Also, what is journalism? To take an example, the Twitter thread has become almost a modern form of column. Is that journalism? I do not know. Is a series of tweets by a prominent journalist categorised as journalism or not? What about Substack? Is everything on Substack journalism? One longs to hear.

There is a series of opacities that could usefully be cleared up, but, like the previous speaker, I would err on the side of caution. There is plenty to do. This Bill does not have to be the final statement on the matter. Indeed, how could it be? To conclude this answer, we should be very careful before thinking that in order to solve all problems of online harms, which no Bill, however brilliantly drafted, could do, we put the ancestral liberties that actually underpin all communication, and indeed all democracy, at real risk.

**The Chair:** Thank you. Perhaps I could ask Silkie Carlo and Barbora this as well. Matthew d'Ancona made the point that he recognised the need for the Bill. Do you recognise the need for this Bill? Do you think there is a problem that needs to be addressed, and do you think the Bill goes about addressing it in the right way?

**Silkie Carlo:** I think there is a problem with online abuse and abusive communications that seem not to be being policed, whether that is criminal communications that are not dealt with quickly enough by the tech companies or that are not, for some reason, being dealt with by the police. What I do not see a need for, given the scale of that problem in particular, is new legislation that specifically targets free speech and lawful content.

From my perspective, for that reason, given the sheer scale of the impact of this Bill, which will affect millions of people and billions of communications, it is, with Clause 11, one of the most dangerous pieces of legislation for freedom of expression of recent years. We are now putting squarely in the crosshairs of censorship lawful communications - by foreign companies with state backing. We are detaching from the rule of law with regards to one of the most fundamental rights we have.

I absolutely see the need for more robust processes around dealing with abusive and criminal content online. I wish the Bill was more explicit about how that would happen, with perhaps more about a portal for people to report abusive and criminal content, or how tech companies and law enforcement can be more joined up in the way they work. All of that is missing, but there is an awful lot of focus on free speech and lawful content. For that reason, I think it is a danger to democracy.

**The Chair:** Let me ask you about the point we discussed earlier. We talk a lot about it in this context. There is a lot of communication out there that is unpleasant and hurtful although not necessarily unlawful. Do you recognise that social media creates a different sort of problem, which is often the frequency of exposure to that content, or even whether people are directed towards content that could cause harm but is not necessarily unlawful? That is a different sort of problem. It is a problem that was not envisaged in the past. It is something that technology has created and we might need a new solution to address it.

**Silkie Carlo:** I am not the biggest fan of the social media companies, as you will not be surprised to hear. I think there are problems with the kinds of experiences that people have online. That is a very different thing from creating state-backed censorship of lawful communications. The spillover is so plainly obvious. For example, public figures need to be protected from criminal abuse. However, if they receive a large volume of criticism online, are we going to say that this is something that technically should not happen and that there should be legal powers to prevent it? Then we are preventing democracy taking its course, potentially.

Reinventing the wheel with freedom of expression is a very dangerous thing. Detaching from the rule of law is a very dangerous thing. I agree with the other witnesses you have heard that, surely, the first priority has to be getting a grip on the sheer amount of illegal content online, and criminal communications and criminal abuse, not creating the kind of state-backed Silicon Valley speech police that we will see under this Bill.

**The Chair:** Where do you put the barriers for that? A lot of the abuse that might not necessarily cross the threshold of being unlawful under the Equality Act could be highly abusive. The sustained nature of it could be highly upsetting and distressing to somebody; a young person could be targeted with content that glamourises self-harm, extreme dieting or anorexia. Any one of those posts might not necessarily cross the threshold for being unlawful, but nevertheless the effect could be harmful.

For me, the question we have to consider when looking at this Bill is to what extent companies are responsible for the environment they create and the harm that it could do. How should they monitor it?

**Silkie Carlo:** They are responsible for it. The immediate way they measure it is by how people use their services. One thing that I think is important, and it is a great shame that this Bill misses it, is that one of the most effective ways you could reduce the actual harms that people experience is by limiting the extraordinary data gathering that these companies are doing. It is granular data gathering and microtargeting, which enables them to create addictive and sometimes unpleasant experiences for people, and means that people are being hit with a barrage of the same, sometimes unpleasant, types of material.

We want to make sure, as it is the Government's obligation to do, that freedom of expression is protected, even in these difficult circumstances. For example, the internet has had an incredible capability in helping people who experience self-harm or eating disorders to build communities, find sources of support, express themselves and go on a survival journey. The risk is that, by saying to the companies that they need to suppress those broad categories of information, we will see some of those vulnerable and marginalised groups being erased.

We have collected evidence of that, which is in our report, *The State of Free Speech Online*. It shows time and again that, when even young people who have recovered from self-harm but have scars on their bodies post photos of themselves on Instagram, they are being taken down. They are not even posting about self-harm, but the algorithms are searching for scars and then deleting their photos, so they become excluded from an online social life. When you are a young person and all your friends are online, that is a very difficult thing. I think there are real risks in seeking to suppress whole categories of information in that way.

**The Chair:** Barbora Bukovská, what is your view on the balance that the Bill tries to strike in this area?

**Barbora Bukovská:** Thank you very much. I represent Article 19 international freedom of expression organisations. Our focus is on how this Bill meets international freedom of expression standards, and whether, as you said, it strikes the balance between protecting freedom of expression as well as protecting people from harms.

When you look at it from the international human rights standards point of view, it does not. It does not meet the basic criteria that international human rights standards require for restrictions on freedom of expression. The restrictions need to be stated with a sufficient level of clarity and lack of vagueness that people and entities can adapt their conduct accordingly. This is so vague and so broad a concept that it does not meet even the first aspect.

The second point is that the restrictions are proportionate and will lead to the desired effect. When we look at this measure, it will not. Some of the preceding speakers have already touched on these issues and the concept of legal but harmful. The problem with the legal but harmful concept is that it is legally protected speech. If the Government think that people should be protected from certain speech, online and offline, this should be met with corresponding criminal or civil offences that will be put in place, but which are not there. The concept is extremely problematic and my colleagues have already mentioned why.

I also want to look at the concept of illegal, because even that concept is not sufficiently defined in this legislation. The Government do not exhaustively list the illegal content that the companies should consider, which also indicates the complexity of the task ahead. They expect the companies to make those determinations. Content that is illegal should always be determined by a court or an independent adjudicating body. That is not happening, and the companies are hardly to be expected to make such assessments.

Going back to what you said on how we crack the balance—

**The Chair:** I am sorry to interrupt you. There is nothing wrong with what you are saying, but unfortunately there is a problem with the broadcasting system, which has just gone down, and I believe the remote witnesses cannot currently hear us. Would you forgive us if we suspend for five minutes or so? We will resume when the problem has been fixed.

*The meeting was suspended due to a broadcasting problem.*

**The Chair:** My apologies for the technical break. Unfortunately, we had a problem with the system here in Parliament, but we appear to be back up and running.

Barbora, you were in mid flow, so apologies for interrupting you. Would you like to resume your evidence?

**Barbora Bukovská:** Yes. Before I was silenced by technology, I was going back to whether we see the need for this legislation or regulation and to your question about how we strike a proper balance.

At Article 19, we fully recognise that there is a problem. There is a problem in how the companies moderate the content on their platforms, a problem with transparency and a problem with the power the companies have over the users on their platforms. The need for some regulation and for improvement on what we have so far is definitely there, but the balance is not properly struck in this piece of legislation.

Something we are really missing, and has already been alluded to by Silkie, is that we think all these problems—which I have just listed, and there are many more—will be fixed if we put small patches on each of the processes, if we fix content moderation and fix transparency. Our view is that they will not be fixed, because that will not address the underlying problem we have, which is the exploitative business-driven model of these companies. If we want to fix the problem, we need to fix the business model. That is not addressed at all in the Bill. It focuses only on content moderation at the expense of solutions that can curb the excessive market power of the dominant companies and abuse of that dominant position. We are proposing some measures that could address the problem through addressing content moderation. This is through opening the content moderation markets and allowing new players to come into the content moderation the companies do, and empowering users.

To your question about how we crack it, we need to think beyond content moderation. The European Union is doing that quite well now by addressing it through the Digital Services Act and the Digital Markets Act together. They see them as part of the same coin. For content moderation, we think that the companies need to unbundle the services they provide. They need to separate their hosting obligations and their hosting services from content moderation services. That will also partially deal with the problem you raised, that people are targeted by abusive content and seeing a lot of stuff they do not want to see. If they can choose the content moderation service they experience on their platform, and where the regulator or the state could provide some basis of common ground, it could lead to a better experience for the users.

That is on the one hand. On the other hand, there is interoperability. None of the problems you are describing would happen if we did not have the concentration of so many users and so much content on very few platforms. The users cannot leave those platforms because they would have to take all their friends and all their content elsewhere, which is difficult. That is what we need to solve, and not obsess solely on content moderation. That is not say that it is in vain. It needs to improve, but we need to be more ambitious and have more infrastructure-based solutions to the problem than just content moderation.

**Q136 The Chair:** I want again to ask one question of all the members of the panel, perhaps going round in the same order. Silkie and Barбора, you both mentioned the content moderation aspect, which is really a data engagement model directing some people to more harmful content because of previous interests. It has been suggested that one of the

solutions to that problem is to say that it might be based not around content removal but around not amplifying harmful content. People have freedom of speech, but they do not have a right to be broadcast by Facebook to millions of people. Can I ask each of the panellists what their views are on that question, starting again with Professor Wilson?

**Professor Richard Wilson:** Yes, that was something that I noticed could be in the Bill. There is the issue of content, but there is also the issue of the degree to which the algorithms elevate harmful content in the user feeds. Why do they do that? It is for the reasons that my esteemed colleague speaking before me just identified. It is part of their business model. This is an attention economy; they want eyeballs on screens, and nothing does that like provocative and often harmful content.

This has been documented extensively. There is an article by Müller and Schwarz, who analysed the Alternative für Deutschland Facebook page. They analysed posts in 2017 that were anti-immigrant. What they found was that the algorithm elevated those anti-immigrant posts on the feeds of users. It is clear that the social media companies are amplifying and elevating harmful speech on user feeds, and that is what the algorithm does. It is mentioned in a few places in the Bill that the algorithm may be regulated, but it could be developed quite a bit.

I will end with a legal case that we had here in the US of Force v Facebook. A young American student, Taylor Force, was in Israel. He was killed, stabbed to death, by a Hamas supporter in 2016. The Hamas supporter had been radicalised online. In discovery, they found that Facebook had distributed Hamas television, and actually elevated it. It had brought people in, contacted through the Facebook networks, folks who were not even part of the Hamas networks. It drew them in and then elevated radical content on their feeds. They found in fact, through the discovery process, that the algorithm had played a part in creating the intentionality of the Hamas supporters to find an American to kill.

In the end, the case was thrown out by the court, but Judge Katzmann, dissenting, said that Section 230 of the Communications Decency Act protected content but not the algorithm. I think there is some room for government regulatory powers to either insist that the algorithm depress harmful content or, at least at a minimum, not elevate it in order to build up the business model of, essentially, selling ads.

**The Chair:** Thank you. Matthew d'Ancona, what is your view?

**Matthew d'Ancona:** As I understand your thought experiment, Chair, it is that we would have retained the right to say things that were potentially harmful, but we would so regulate the social media and digital world that proliferation could not take place. There would be a non-proliferation pact, to coin a phrase.

It is an interesting image. What it argues, which has truth in it, is that it is the proliferation that is the problem. I am not sure if that is the case,

but I think it would set up a precedent in common law, which is that you can say it, but you cannot have it said again and again. That is an interesting proposition. Of course, you do not need me to tell you that there are lots of get-arounds for that, with screenshots and so on.

What I think gets to the heart of any discussion of this—I cannot really improve on Richard Wilson’s explanation—is the all-importance of the algorithm. It remains, to a mysterious extent, the revered and sacred black box of the digital economy. It is, after all, just maths. While some of it deserves commercial confidentiality, the extent to which it is shrouded in a revered atmosphere of secrecy is quite absurd. I think all attempts, national and global, to take on this problem in its many manifestations will require some sort of presumption of access to algorithms and the right to form judgments about their appropriateness, and the extent to which, taking the horrible example Professor Wilson showed us, what is, essentially, an advertising economy can proliferate a call to death.

To soar slightly above the specific, until we get into that black box we are going to be whistling in the wind. I hope that the law, as finally passed, gives the various authorities the power to demand access to algorithms and really does call big tech’s bluff. For 20 years, they have been saying, “We couldn’t possibly allow you into that bit of the business”. This is horribly reminiscent of the way the arms trade used to claim that its transgressions were protected by “commercial confidentiality”, which was complete nonsense. We must not be gulled by that. The stronger the demands in the final statute on this front, the better.

**The Chair:** Thank you. We are reminded, of course, that algorithms were built by human engineers rather than being self-creating and autonomous beings.

**Matthew d’Ancona:** Exactly.

**The Chair:** Silkie, what is your view on the question of freedom of reach versus freedom of speech?

**Silkie Carlo:** There are different ways that policy in this area could play out. What we would be uncomfortable with is the state defining certain categories of lawful speech that foreign companies would then have to suppress. That is what we are talking about. They already amplify content because, essentially, they editorialise with newsfeeds, recommendations, advertisements and so on. We would be talking about setting out a specific number of categories and saying, “Don’t advertise this. This is semi-banned content”. I think that is a very slippery slope.

At a systems level, the kind of transparency that Matthew spoke about is vital. I think there is unanimous agreement on that. There is no reason that we should not be able to see, given the importance to democracy, social cohesion and everything else, how the algorithms work and what types of content are being pushed through the editorialisation process.

On the individual level though, I urge caution about designating lawful content that is, essentially, semi-banned. The most important thing that policymakers could do to prevent overamplification of harmful content to individuals and radicalisation spirals, et cetera, is to ban microtargeted advertising by the social media companies. It would be a revolutionary change. It would mean that the extremely harmful journeys that individuals go on would be prevented. It would also prevent a lot of the harms to democracy that we have seen perpetrated by some of the companies.

Microtargeted advertising currently exists within a total regulatory and legal gap. I was quite surprised that, after the Cambridge Analytica scandal and other things that we have seen in recent years, that is completely unaddressed in the Bill, despite its being one of the most significant online harms objectively that we have seen in recent years. The ICO has called for an ethical pause on microtargeted advertising as it relates to political campaigning. Even the Institute of Practitioners in Advertising has called for a pause to online microtargeted advertising. Certainly, lots of us in the rights sector would call for a pause to microtargeted online advertising. That would cut the head off a lot of the harmful journeys that you are concerned about in this committee.

**The Chair:** Advertising is an important area that we have discussed as well, but it does not address issues like News Feed or UpNext on YouTube, which are around the recommendation of previously posted content, not advertising. I think it is a slight myth that Facebook, in particular, puts out that the only way you can push content through the network is through advertising. There are other ways that can be done.

**Silkie Carlo:** If I could clarify, it is because they are selling the data that they collect it. Regulating the collection of microtargeted data, granular data collection from individuals, would be a very important thing, too. The reason they are collecting it is for advertising. I quite agree that the harm is not simply in paid-for adverts. It is the fact that users are designated according to very detailed amounts of information that then determine what they see and what they experience. That is something that should be prevented.

**The Chair:** Although I think they still gather the data to hold people on the platform, because the longer they are there, the more they see. I do not think stopping one would necessarily stop the other.

**Silkie Carlo:** Which is for advertising. That is how they make money.

**The Chair:** The point I was making is that it is an engagement model; they gather the data for targeting but also for holding attention. They do it for two reasons, both of which are problematic. I completely agree with you on that. Barbora?

**Barbora Bukovská:** I have one additional comment. I fully agree with what has been said by the previous speakers. In terms of transparency of the algorithm, it is difficult to provide viral solutions because, as Matthew

said, we do not really know what is happening. There have been attempts from some regulatory agencies to penetrate the problem. For example, in France, they had a working group embedded in Facebook for several months. After that period, they still had no idea what was happening.

As regards auditing, it is very difficult even to suggest specific recommendations. There should absolutely be transparency, not just by independent regulators but by independent researchers or academics. That needs to be enabled. We call them multi-stakeholder transparency audits or assessments and human rights audits or assessments, which cannot just be in the realm of regulators.

Another point from the free speech perspective is that the pluralism and diversity in what people see and engage with online is also a legitimate objective of freedom of expression and positive obligations. That absolutely needs to be taken into account. What you were saying about News Feed and how the companies run the content on users' individual pages and profiles is important from the pluralism and diversity point of view.

Q137 **John Nicolson:** Professor Wilson, there has been an explosion in homophobic hate crimes in these islands over the last couple of years. Crimes against trans people have quadrupled, real-life crimes; they are being beaten up in the street and elsewhere.

There has been an explosion of transphobia in our press and in our broadcast media. It is quite an extraordinary thing to witness. We have also seen it online. A lot of trans people have told me that, as they are often squeezed out of discussions on the "Today" programme and elsewhere, despite the fact that they are being attacked, online is one of the few places where they feel they can defend themselves. They can talk to other trans people. They can talk to trans teenagers and try to give them a bit of moral support. Do you think it is possible that this Bill will lead to minority communities, such as trans people, being silenced online?

**Professor Richard Wilson:** Thank you very much for the question. The short answer is that I do not know. It may be shocking to hear an academic say, "I do not know", but I do not. It is hard to tell. Could you tell me more about your concerns that the trans community or the LGBT community would be silenced online as a result of the Bill? Are any provisions giving you cause for concern?

**John Nicolson:** The LGBT Foundation, for example, in its witness statement to this committee said that LGBT content is taken down more often than other content. I suppose the argument would be that the social media companies might be overly concerned about the provisions of the Bill, and might decide to be proactive in taking down content because they are worried about the punishments if they do not, and they would not look too carefully at the kind of content that they were taking down. You already illustrated that when you mentioned African American communities being disproportionately taken down in the States.

**Professor Richard Wilson:** That would have to be under constant review. The question, I suppose, for all of us is the balance between protecting the trans or LGBT community from significant harms, hate speech, threats, harassment and incitement, and whether that is greater than the possibility that the content put up by trans users or LGBT users could be taken down. You are absolutely right that there is a possibility that the algorithm could be more aggressively policing that speech.

**John Nicolson:** For example, the social media companies might be concerned about the fines they could get. This is a very different world from the world we live in at the moment. At the moment, you can pretty much say anything, and the social media companies will not take it down if it is homophobic. They do not abide by their own community rules. We are looking ahead. If the Bill goes ahead as planned, and if there is overzealous takedown of content, what provisions do you think there could be or should be for folk to appeal?

**Professor Richard Wilson:** I defer to my colleagues on this, but I think that there are reasonable appeals provisions in the Bill, both at the level of the companies themselves and at the level of Ofcom. That is a real question. Are the appeals provisions significant, meaningful and robust?

Perhaps I could point out an event that took place recently in the United States. One of the more popular shock jocks who had a programme on YouTube—his name is Steven Crowder—was removed by YouTube over the summer for making trans and homophobic remarks. He simply lost his YouTube channel. That was quite a high-profile and significant move that YouTube made, I think largely on its own initiative, to send a message that trans and homophobic comments would not be tolerated on YouTube. It has been one of the better social media companies in regulating this type of speech.

If a high-profile message like that is sent, the Bill could be more protective of the trans and LGBT community. You are absolutely right that there is always the possibility that social media companies would be overzealous and suppressive in the silencing of that speech. That would have to be under constant review.

**John Nicolson:** It is interesting that you mention Ofcom. I know that Ofcom LGBT staff are very concerned with where Ofcom itself is going on the whole issue of LGBT rights internally. Many of them have contacted me and said that they are deeply worried about their own organisation and its treatment internally of its LGBT staff, but that is probably for another occasion. Matthew, you have your hand up.

**Matthew d'Ancona:** Thank you for the question. It is really important. Clearly, it is absolutely vital that the Bill protects trans and gay people from any form of speech that will lead to violence. On the other hand, we get back to the legal but harmful question, and the words "psychological impact", and, again, it is a question of the current climate.

The current climate is one where legally and legitimate gender-critical feminism is often categorised by some people as transphobic. One only has to look offline at the recent experience of Professor Kathleen Stock at the University of Sussex to realise this is not a hypothesis; it is a real prospect. While I take your point, one must be wary that the consequence of this Bill is not to limit the completely legitimate right of gender-critical feminists to raise their lawful concerns about the way in which the trans debate is being negotiated.

**John Nicolson:** I think “whipped up” is the term that I would use, with imported culture wars from the United States, because we are of course not in any way reducing the rights of women by enhancing the rights of trans people. It is important to remember that.

**Matthew d’Ancona:** I am sorry to contradict you, but I think that a lot of women feel that there is a conflict of rights and feel that social media is a place where they have a right to express that dissent.

**John Nicolson:** They may feel that, but I think that is wrong. I do not want to stray down—

**Matthew d’Ancona:** You may say they are wrong, again—[*Inaudible.*]

**John Nicolson:** This is a very specific discussion, and I do not think we should go down that particular rabbit hole at the moment.

**Matthew d’Ancona:** But you did. I am answering your question.

**John Nicolson:** I just raised the issue of LGBT rights being protected; that is all.

**Matthew d’Ancona:** Yes, quite, but hold on. The problem of a Bill of this scale and scope is that when you raise a question like that you cannot object if it then leads to another. That is in the nature of a piece of legislation of this scale and scope. Simply saying, “That wasn’t what I was asking about”, I am afraid, will not wash. Each question leads to another. That is etched into the fibre of this Bill.

**John Nicolson:** I will leave you with the final word on that because we could discuss it for the next hour, and it is not what this is about.

**The Chair:** I am sorry to interrupt, but if you do not mind, both Silkie and Barbora would like to come in.

**John Nicolson:** I was about to say that perhaps we could hear from both of the other witnesses, please.

**Barbora Bukovská:** To your question as to whether this could lead to suppression of content that transgender or LGBT people put online, I think that if you allow so much subjectivity in interpreting the terms the Bill puts forward, it will lead to such suppression. That is our experience at Article 19. We work around the world, and we see how companies moderate content, and it actually does lead to suppression of minority

voices when their content has been taken down. This goes to the vagueness of the definitions and how companies will determine it.

Even for some of the terms we already touched on, such as what kind of content causes a significant impact on ordinary people's sensitivities, which is probably the provision you are referring to, I cannot imagine the companies will make an assessment of each aspect of the definition. I do not know how they can confidently make such determinations in the absence of any context.

We need to understand that when discussions happen online they are very robust. For example, when JK Rowling put some comments about the transgender issue on Twitter, there were vitriolic comments on both sides. Many people felt harmed on the side of transgender people, but when transgender people were putting up comments, people from majority communities felt harmed. It is entirely possible to imagine that the companies will be making determinations that any content can reach that level of severity, and that will lead to suppression of legitimate speech.

Looking at the logic of the formulation, this "adult of ordinary sensibilities" does not really make sense, because people can be harmed by something even if they are not sensible. In that sense, the Government want to protect people from a certain harm, as they might be sensible to it, but they ask the companies to determine the level of sensibility. Legally, it does not make much sense to me, especially if you are trying to protect someone from the impact of speech. In order to protect minorities, the formulation should be extremely narrow. It needs to be extremely narrow to comply with international freedom of expression standards. The determination that certain speech is causing psychological harm to someone is really a recipe for trimming speech that some people find offensive and critical, and we are absolutely against that.

**Silkie Carlo:** I am really glad that this issue has come up because I think it is a very clear example of why we should be strengthening the rule of law in this area rather than relying on subjective interpretation. We have already seen here some cracks appearing in how much we can agree on the types of things that might fall in scope.

Certainly from our research, and we have been researching online censorship for the past three years at least, we can see a disproportionate amount of LGBTQ censorship online and over-removal. Coming back to algorithmic removal, I referenced earlier that Instagram algorithms, for instance, search for scars. That has impacted people who have gone through gender reassignment surgery. People have had their photos taken down when they have surgery scars. It is even jokes; for example, there was a gay man who posted, "Gay men are the worst", or, "Men are the worst", something like that, and it was taken down because it was deemed homophobic.

On the other hand, there are hundreds if not thousands of cases of women, perhaps what you would call gender-critical feminists, as Matthew said, who have been censored, suppressed or simply banned from popular social media sites. One of the most high-profile cases was a Canadian journalist who said that men are not women. She was permanently banned from Twitter for that. They have an extraordinary gender policy in their hate speech policy. The problem is—

**John Nicolson:** Are they doing that? I could go on Twitter now and find hundreds and hundreds of people saying exactly what you have said, or variations of it, and some really offensive things from that particular perspective, and they are all still up. Lots of them are pseudonyms and eggs. There is the idea that Twitter is taking down masses of this stuff, but if it is, a lot of it is still up.

**Silkie Carlo:** It has been, and it has been quite erratic in the way it has done it. It has gone through different phases. It responds to press criticism and the wider public conversation, so there is a fluidity as to how it actually enforces in this area, but the fact of the matter is that for a lot of people the damage is done. A lot of people have built up social media profiles or journalistic careers, et cetera, and in those two years, a lot of those women were being silenced for saying things like men are not women.

What we are doing is somehow aligning British law and a state regulator to Silicon Valley values about acceptable speech. That means that on these tricky issues of sex and gender, for instance, the companies take a self-ID approach, which we do not really have in the UK in the same way. In another high-profile case, there was an individual who, to all intents and purposes, was basically a predator who tried to get female-only services, while in his professional life having a male name. It went to court in Canada. Women who were talking about that online, talking about the specifics of that case, were banned from Twitter for talking about a court case that was happening in another country, which clearly was relevant to some of the gender debates that were going on here.

Another area where we see it is in gender and discussions of crime. We have detailed these examples in our state of free speech online report; you can see screenshots of the censored material. For example, numerous women have had their posts taken down or been banned from Twitter for saying things like only men rape, which is a fact in UK law, but it falls foul of the extraordinary sex and gender policies of those foreign companies.

I am glad that we are having this discussion because it is exactly where the discussion needs to be. Here we can see the fissures between a British approach and domestic law and the rule of law on our communications offences, compared with the values of foreign big tech companies.

**John Nicolson:** On a point of detail, you used the word British, but in fact the Gender Recognition Act reform is part of the programme of

government of the Scottish Government and will be passed, because it is a manifesto commitment of the Scottish Government and the Greens, so we will see it tested in practice in these islands.

**Q138 Lord Black of Brentwood:** Going back to the question of censorship and regulation, we heard some very powerful evidence from the Carnegie Trust, and we see in the submissions from the publishers we will hear from in the next session that, actually, censorship is already happening; Facebook and Google are taking content, and either altering it, sometimes altering headlines, or refusing to carry it. That is already there, as you have indicated. The question is whether this Bill will make the issue worse, because content from legitimate publishers who are exempted from the Bill will in fact be subjected to a form of double regulation. The platforms will look at it and say, "We don't think this adheres to the codes that Ofcom have put out", and therefore refuse to carry it, so a lot of democratic content, which should be out of the scope of the Bill, will be censored by the platforms themselves. That sort of carries on from what you were saying just now, Silkie.

**Silkie Carlo:** Could you clarify the question for me, please?

**Lord Black of Brentwood:** Is there a danger of double regulation, so that content that is already regulated will then be double regulated by the platforms? They will seek to apply the codes that Ofcom lays out, take their own judgments, and say, "Sorry, we're not going to carry this".

**Silkie Carlo:** That is right. There is a large degree of deference in the Bill as to what the policy should be to the companies, and Ofcom is in the position of ensuring that the companies enforce those policies. For example, where a Twitter policy is far removed from what would fall foul of the law in our law—there is even divergence there between the four nations—a state regulator is making sure that that content is censored. That is an extraordinary contradiction, and an undermining of freedom of expression by our Government. I think it is a real concern. The other side of the coin is of course that the Secretary of State has enormous capacity to set what some of those policies should be as well, so you really have the worst of both worlds.

**Barbora Bukovská:** I am looking at the clauses that provide protection from this kind of content, such as Clauses 12 and 13. There are some concerns at Article 19 that indeed the situation will be made worse. The first problem is the fact that the platforms will have an obligation to give regard to freedom of expression. That is insufficient; we think this will simply be a box-ticking exercise where services will have to show that they have considered freedom of expression or, as regards privacy, that they have taken it into account. We worry that, when weighted with all the other obligations they have to protect children and adults from such content, the freedom of expression and privacy aspects will be secondary and will not be given due consideration in the balance. They would also need to have a very good understanding of what freedom of expression encompasses. There is a magnitude of law that regulates such content.

Going to journalistic content, we need to recognise that in this day and age journalistic content is much broader than the news or the established media. We have a lot of citizen journalism or posts by individuals. They are probably not professional journalists, but are making information available to the public on current events, and on events such as terrorist attacks or earthquakes. Users report about those events in real time. There is a real question of how that content will be protected and how the platforms will balance those different rights, especially in the light of the other definitions, such as legal but harmful content. I can imagine that the situation will be made far worse with these provisions.

**Professor Richard Wilson:** I dissent slightly from what has been said, in so far as it strikes me that the Bill only lightly regulates and largely exempts mainstream journalism, so I do not see the same concerns about double regulation that my colleagues have expressed. If they have those concerns, I would appreciate them pointing to a particular section of the Bill that they find problematic. They may well be right, but I did not see it.

My concern about journalistic content is the grey area where you may have terrorist organisations closely allied to journalistic outlets. I do not have the section in front of me—I was just looking for it but could not find it—but I think I am right in saying that the Bill exempts press agencies run by foreign Governments. Let us say for the purposes of argument that there is a news outlet that is very pro-Hezbollah. Hezbollah is largely supported by the Iranian Government, so the Iranian Government have a press agency set up in a country like, for the purposes of argument, Lebanon, which is pro-Hezbollah—a known terrorist organisation. That would not come under the regulation of the Bill, so I would suggest that the Bill actually be a little more aggressive than it is with respect to those areas of news media that are essentially outlets for terrorist organisations. However, it is a grey area, and I do not claim to have the ability to define it. I can see that it is problematic.

**Lord Black of Brentwood:** Matthew, do you want to come in?

**Matthew d’Ancona:** Only to answer slightly obliquely and say, like Richard, that I did not see a problem of double regulation in the Bill. I saw a risk of half-regulation, which is that so much authority and power is given to the companies under the duty of care rubric that I think the problem is otherwise—that we will effectively end up finding ourselves having delegated all this to those who have a huge, vested interest in the status quo or, to put it slightly differently, a huge, vested interest in defining words such as liberty, harm, hate, and so on, in ways that suit their commercial intentions.

Of course, it is always wise to keep an eye open for duplication of roles. My concern, however, is that we are handing almost the entire role in practice to people who are actually protagonists in the debate and protagonists in the activity we are seeking to regulate. That would be my greater concern.

**Lord Black of Brentwood:** Silkie's point is that we are also giving considerable powers to the Secretary of State in this area. I would be interested in the views of the other panellists on the Secretary of State's powers.

**Barbora Bukovská:** These powers are some of the most concerning for us because they indeed give very broad powers to the Secretary of State to define the scope of the services and power over Ofcom. We think that many provisions of the Bill show that the Government will have a very close grip on the implementation of the Bill by Ofcom. It is absolutely necessary that the state takes action so as not to undermine Ofcom's independence in its day-to-day work, and these provisions are redundant.

**Lord Black of Brentwood:** Do you think we could get round that if the Secretary of State's authority was subject to parliamentary approval, for instance?

**Barbora Bukovská:** There is a parliamentary role in the powers that are given, but we believe the provisions regarding Ofcom and the powers of the Secretary of State should be removed, unequivocally.

**Professor Richard Wilson:** I also had some concerns about the role of the Secretary of State. Under Section 41(5) the Secretary of State can designate content as priority illegal content, bypassing Parliament. The question we should always ask of legislation is, "Would I like this in the hands of my political opponents?" because one day they will come to power.

I would be very concerned about the temptation to engage in political spin and engineering, where each party in power, through the Secretary of State, tilts the scales in its own direction. This could have very damaging effects on freedom of expression. I feel that Ofcom is not sufficiently independent. In Sections 32 and 33, it must submit the code of practice to the Secretary of State, who has the discretion to modify. In Section 110, the Secretary of State can provide guidance to Ofcom on all aspects of the exercise of its power. I think it is better in the long run to grant Ofcom more independence and authority than the Bill does, because that will give the exercise of its regulatory powers more legitimacy. I do not think I am alone in trusting career civil servants a little more than politicians.

**Barbora Bukovská:** For the avoidance of doubt, the provisions that we suggest should be fully removed are Clauses 33, 34(6), 57 and 113. Clause 41(5) should be reworded to give Ofcom an obligation to consider human rights in its deliberations, but those four other provisions need to be removed.

**Silkie Carlo:** I have concern about Clause 46, which gives the Secretary of State the ability to define harmful content, and is detached from the broader definition of harmful content in the Bill about psychological impact. It is just a standalone. Basically, 46(2)(b) gives a blank cheque to the Secretary of State to define what harmful content is.

It is important that we are talking about this in the context of the pandemic as well, and some of the social media policies we have seen in the wake of the pandemic. Thinking back to your point about journalistic content and heightened protections, we are in an odd situation where, for example, Sir Richard Dearlove would not be able say on his Twitter account, when this policy was in place, that he thought it was plausible that there might have been a lab leak in Wuhan, because that was not allowed on social media companies. However, he was able to do an interview with the *Telegraph*, and for that to be published. Under this Bill, the *Telegraph* would be able to post that to Twitter, but were he to have his own social media account, for example, he would not be able to post it because it would have fallen in breach of the rules. That is a good example of why this whole construction around harm is problematic, and a good demonstration of why some, frankly, knee-jerk speech policies, that are detached from the rule of law, made by Silicon Valley, and that are endorsed by the Government, because of course the Government—

**The Chair:** I am really sorry but I need to move this on. I think you have made your point about the areas where you want to see amendment. Could we move on to the next question?

**Silkie Carlo:** May I finish?

**The Chair:** Briefly, yes.

**Silkie Carlo:** The Government set up a counter disinformation unit to make exactly these kinds of extrajudicial requests. We should not be fooled into thinking the Secretary of State will not use that power to censor lawful content that actually, as we can see months later, is of great public importance.

**The Chair:** Thank you.

Q139 **Lord Clement-Jones:** We have quite a spectrum of opinion here. At one end, Silkie, I think you really want to scrap the whole Bill and go into a greater data protection area.

**Silkie Carlo:** Scrap Clause 11.

**Lord Clement-Jones:** It sounds as though there are a lot of other aspects that you would like to scrap. I think Barбора is probably next in line. Further down the spectrum there are greater degrees of acceptance of aspects of the Bill. The fundamentals of trying to protect freedom of speech and so on, on the face of it, lie in the form of risk assessment the Bill is going to introduce, overseen by Ofcom, which will require the social media platforms to carry out a particular form of risk assessment, and that is why phrases such as protecting freedom of expression and journalistic content appear.

To meet the criticisms we have heard today, do we try to improve that process in some way by prescribing more protections for freedom of expression, or is the risk assessment process idea contained in the Bill, and what Ofcom will do, itself flawed, and we need to find another way of

dealing with it, for instance, simply treating social media platforms as conventional publishers? Richard, is there some other way of approaching this?

**Professor Richard Wilson:** I refer to my colleague at Article 19, who I believe talked about interoperability. I do not know that the British Government want to get into anti-trust litigation with Facebook quite yet. I expect that will be coming in the United States at a certain point. On the interoperability point, which basically allows users to export all their contacts, photographs and networks on to another platform, I think the British Government could insist on that. There are a variety of different approaches.

I would end with a comment. It is interesting that what we are not arguing about, and what we are not subjecting to great scrutiny, are the provisions relating to terroristic content and child protection. There I would recommend a much leaner, slimmed-down Bill, which focuses on those aspects and runs, essentially, a kind of pilot programme to see how those components work, and to see how the complex mechanics in the relationship between the media companies, Ofcom, the Secretary of State and Parliament function. There has been a lot of speculation about how possible harmful or unintended consequences of the Bill play out, but I recommend a slimmed-down, pared-down Bill focusing on the essential content of the Bill related to terrorist and child protection activities, which then builds out later on. You can have lots of bites of the apple over time. It is interesting to think about the elements we have been focusing on in today's discussion and the elements that have been relatively non-controversial.

**Lord Clement-Jones:** Matthew?

**Matthew d'Ancona:** On which specific question?

**Lord Clement-Jones:** Do you think the risk assessment process is fundamentally flawed and that we need to add more protections to it?

**Matthew d'Ancona:** No. For the purposes of concision, I would put a big fat tick by what Richard just said, which I think is an excellent idea. Furthermore, I am nervous of adding to the pot of things that are prohibited things that are not yet illegal. I think what we might now call the Wilson doctrine is a very, very good way of going ahead.

Again, I keep being struck by the sheer scale of this. There is a danger that, in trying to do everything, one ends up doing too much or nothing. A focus to the Bill would be very wise, as would an approach to the Bill as a process rather than an event. Legislatively, it was going to be part of Parliament's calendrical activities, and given what it strives to do, which is to bring some sort of civic order to the most transformative revolution in technology of our lifetimes, that sort of makes sense. To begin with a narrow spectrum of risk and then watch, see what happens and build up from there, as Professor Wilson said, is a very, very good approach. I echo that wholeheartedly.

**Lord Clement-Jones:** Barbora, is it all about scope for you?

**Barbora Bukovská:** Yes, I think you are touching on a very important topic; the Government want to regulate here the whole scope of human interactions, which should not be the role of government policy. We are proposing an approach similar to Professor Wilson's. As he said, we need a targeted approach. The Government should only improve the content moderation practices of those companies that goes to illegal content. Their obligations should be narrowly and very specifically defined in that way.

The category you call legal but harmful should be resolved not through legislation but through improvement of what companies do on their platform. As we heard earlier, their approach is erratic and not clear. Some groups do not know why their content is removed. We need the platforms to improve content moderation. The question is how they do that. The Article 19 recommendation is to improve what we call the multi-stakeholder processes. We call for the creation of what we call social media councils. Facebook has the Facebook oversight board, which we think is problematic, but there is scope for the creation of stakeholder entities that would advise companies on their policies independently, and on how they design content moderation practices.

Legal but harmful content is content that dominant platforms can remove from their platforms under the community guidelines and for this they must improve consistency, transparency and accountability for the content. It is what I call tiered approach. On the one hand, the state should address illegal content with a level of clarity. Then there is the multi-stakeholder process, where there are independent oversight bodies for the companies. That would be combined with interoperability, which we have already discussed, and opening up the content moderation market. Facebook, Twitter or Google would not be the only ones moderating content on their platform. There would be competition. They would open their protocols for competitors to come. As a user, you would have a choice about who was doing content moderation for you, and it would not just be in the hands of Facebook.

That would also address the problem I described at the very beginning, which is abuse by the dominant power, the platform, and would give users more power and control over the choices they make on social media, and a better experience, because they could choose better content moderation services. If I am a minority group or a woman, I could choose what I am going to see on my platform. There would be some common base that a regulator could put together and impose on those who are competing in doing content moderation services on the platforms. It is what we call ex ante regulation, where competition and consumer protection come in. That is not just some blue-sky thinking by Article 19. Even Twitter has started to talk about it. A couple of years ago, Jim Dorsey came up with the idea that it should have a decentralised model for content moderation. I think it is called Bluesky. Twitter is apparently working on that.

We need to move beyond the obsession with harmful aspects and look at a more innovative solution to the problem at hand, a more technical solution, but *ex ante* legislation could be much more helpful. Also, it will have the desired effect you are trying to achieve: to protect users, with a better experience, and to protect freedom of expression on those platforms. You will not achieve it with this Bill, unfortunately. The Bill will probably more benefit lawyers than companies and users and the protection of human rights.

**Silkie Carlo:** I agree. You said we were on a spectrum, Lord Clement-Jones, but I think there is broad agreement that the focus has to be on the illegal content. Things are more visible than ever. For a lot of the harms we are talking about, the reason such a high volume of them is being referred to the authorities—for example, the awful child sexual abuse material on Facebook—is that they are more visible than ever. The authorities need to get a grip on that. I might even expand it to the abuse that is directed at, in particular, public figures—Members of Parliament, for example. We need to see law enforcement get a grip on that. I would have liked a portal, something that makes the reporting process easier, not just to the companies but to the police as well. That has to be the focus. I back the Wilson doctrine.

Q140 **Debbie Abrahams:** I am conscious of time and, to be honest, in some of the earlier remarks around algorithms you have probably already addressed this. On Monday, we took evidence from Sophie Zhang, who you may know is a Facebook whistleblower, and she made the very pertinent point that, obviously, we support freedom of speech and freedom of expression but there is no freedom of distribution. Do you agree or not? Perhaps we can go round as quickly as possible, so a sentence from each of you would be very helpful, starting with Silkie.

**Silkie Carlo:** I think we are in dangerous territory if we start to designate lawful content that should be suppressed. That is a different thing, but I think that is where the policy leads to.

**Debbie Abrahams:** This is about companies deciding what should be distributed, so it is companies and the algorithms, and I think that is what everybody was saying earlier. Have I misunderstood that—that companies decide the information they want to push?

**Silkie Carlo:** I support opening up algorithms at the systems level to transparency. As I said earlier, a lot of the harms could be reduced by not allowing granular data targeting to create individual experiences that are finely attuned to that person's demographics and views, for example. Putting your point in the context of this Bill, we would be talking about designating specific types of content that should be suppressed, and that is a dangerous route that I think should be avoided.

**Barbora Bukovská:** On whether we have freedom of distribution, I would not put it in those terms. I would put it in terms of what I said before about diversity and pluralism, because you as an individual have a right to access diverse and pluralistic information from a variety of

sources. The companies should also have an obligation and have instruction on how they achieve it.

**Debbie Abrahams:** That is an interesting distinction between what I was saying about the companies deciding what you should have and you saying it is about access—

**Barbora Bukovská:** You have access that the company shows you and gives you access to, because it is moderated by the companies.

**Debbie Abrahams:** It is an interesting point, is it not, and a subtle distinction? May we move on, please, to Richard?

**Professor Richard Wilson:** I simply echo the points my colleague just made about transparency. Governments have a legitimate interest in reviewing the algorithm. I believe that it would be a great contribution if the British Government could have access to the algorithms of the main social media companies for review. The platforms already decide what becomes viral and what does not. They are already curating our feed. They are already deciding, on the basis of their business interests alone, the material and content that we ought to see, and how it ought to be distributed. There is a legitimate interest of Governments to have a say on that, and a review of it.

Again, it should be temporary, targeted and transparent. I will give you an example. Public order and protection of religious minorities may be an overriding interest of a state in a particular context. In 2017 to 2018, Germany took in 1 million Syrian refugees. It was a remarkable thing. There were a number of attacks on those Syrian refugees and incitement of violence against those minorities by members of the AfD and far-right parties. The German NetzDG law came about as a result of the social media companies not engaging in a dialogue with the German Government, who sought to suppress anti-immigrant speech at the time. They may not want to do that for ever, but for a temporary period I believe the German Government, and Governments generally, may have a public order interest in negotiating with the social media companies about the virality of content.

**Debbie Abrahams:** Those are very useful examples, Richard, thank you. Matthew?

**Matthew d'Ancona:** My starting point in all this is liberty under the law; therefore, I am wary of anything that might add slightly hazy rubrics of content categories that were subject to specific restrictions outside the law. If Parliament decides that the discussion of horticulture is to be analysed in a particular way, that is fine by me, but up to and until that point I am nervous of—again I use the words—mission creep, where suddenly you find that lots and lots of categories of speech and expression are being aggregated into things that we insist the companies look at. As Richard said, they look at lots already. There is plenty of curation going on already. Super-curation is not desirable. What is desirable is a robust and effective law. Having listened to the excellent

conversation today, I think more and more that the more narrowly defined and focused that is, the better.

**Debbie Abrahams:** Thank you very much.

**The Chair:** As a point of clarification, Professor Wilson, of what you said in response to Debbie Abrahams, it would seem to me that what could happen in that context is that, if there was a public safety issue, a Government would seek to investigate the algorithms of a social media company to look at how they promoted speech that incited dangerous criminal behaviour against individuals. If such an inspection regime existed, the Government may determine, "We think you were irresponsible, you were party to an incitement that could be considered to be an offence, and you behaved irresponsibly".

In a situation like that, I can see that the regulator would start to build up, in effect, a form of case law based on determinations of particular situations, where only after the fact and an investigation could it be determined that a company had behaved irresponsibly and allowed incitement to take place that had led to physical violence and would, therefore, try to set rules and guidance to say, "In situations like that in the future, this is how you should behave, otherwise you will probably leave yourself liable to some form of prosecution under incitement legislation". I see you nodding. Is that what you think might happen in practical terms in that case, and do you think it would be desirable?

**Professor Richard Wilson:** I think there are two processes. One is the criminal law, and there are laws of course about incitement, harassment and threatening speech. The other is the administrative law through Ofcom and the regulatory powers it has with respect to the platforms. Those may reinforce one another. In the Bill it says—I cannot remember the exact clause but I could find it—that the platforms would not be inherently liable if a finding were made against them and they were found to be in non-compliance of the Bill, but that the evidence produced may be used in a court of law. These two elements may run alongside one another, they may reinforce one another, or they may be in parallel and independent of one another. The situation you have just described sounds plausible to me, however.

**The Chair:** On that basis, do you think the Bill should be more specific that that is a route that would be open to people? If the regulator determined following investigation that a company had been negligent and allowed a situation to occur where incitement took place, it could open up a route through the civil courts where the case would be taken on the basis of that judgment being made.

**Professor Richard Wilson:** From what I know about the process, it is likely.

**Barbora Bukovská:** It would probably also depend on the context, because if there was a causality between what the platform did and the impact happening, violence or whatever, it would probably need to be

very clearly determined. I can imagine that in countries like Myanmar, as regards Facebook and genocide, it would be quite likely because Facebook is the internet in Myanmar, but it is probably not the case in the UK that people only access information through that platform and not through other forms of media. A clear balance would need to be established, narrowly tailored to that effect.

**The Chair:** We have the Wilson doctrine based on illegal content, but algorithmic investigation could highlight negligence and, ultimately, it could be determined in the civil courts whether a company was liable because it was operating below the standards in the guidance set by the regulator. It could be the Collins corollary to the Wilson doctrine.

**Lord Clement-Jones:** A protocol.

Q141 **Baroness Kidron:** I do not mean to attack the Commons, but I want to dig down on this idea of a narrow Bill that only deals with illegal content. First, it rather leaves the black footballer who misses the goal with the potential of having to take people individually to court to get some satisfaction. It also leaves us with the 71% of young girls who currently feel silenced about giving their opinion. I am interested in how you feel that addresses the status quo when, actually, it is private companies setting out the rules of engagement. It is not that there is some absolute freedom. It is a freedom curtailed by the values of Silicon Valley, as someone said earlier. That is one part of the question.

Another part is about how we enforce it. One of the things I am very concerned about is the financial crimes that do not get dealt with. We have talked about CSAM, where there is a 10-year backlog on existing cases, even if one more case never came to us. We say that we have to resource policing, and then I am concerned about the criminalisation of this space, police oversight and a whole other set of problems. Richard, I would like you to give us a bit of hope, from your perspective, on some of those issues. Does this mean the status quo and does it mean that the police have to be all over it?

**Professor Richard Wilson:** Thank you for your impossible question. I will give you a merely inadequate answer. I start with the comment that there is no magic bullet. There are many different levers and they must be used simultaneously. We should not expect any Bill to end racism or misogyny in Britain, or the United States, or anywhere else for that matter. This is a long-term cultural and educational project. I applaud the element of the Bill that encourages the teaching of media literacy. That has to be a core piece. Much of the hateful speech on line is reprehensible and offensive but not illegal, and must be met with counter speech—civil but confronting speech that challenges it.

I do not have a close enough grasp of the British context to comment very accurately on this, but it seems from our conversation today that there is a certain amount of frustration with the public prosecution service in this regard, and that the Bill is, in a sense, compensating for the lack of prosecutorial robustness on this topic, and there is a lot of

inciting and threatening speech going unchecked and unremarked. Of course, the prosecution service cannot deal with all of it, but the most egregious examples are happening with impunity.

There is also the civil process raised by Mr Collins, I think correctly so. In fact, I am very hesitant to use the criminal law, which is a very blunt instrument, on speech matters. The civil law and opening up for civil liability may be a better way to go, because the civil courts have a history and an experience of evaluating emotional and psychological harm and awarding damages on that basis. In a sense, they are trained to do that.

I guess I would end with where I started, which is that there are many different levers and they all have to work in the same direction. No one element will do it all. In a sense, I think there ought to be something of a firewall between the regulatory oversight of Ofcom and the criminal law and civil law avenues, so that they function independently of one another, but where, of course, a court and a plaintiff in discovery may request information and evidence from the regulatory body and receive it in a particular case. If all the evidence is maintained, that can simply be handed over. I do not see quite yet, but I am open to persuasion, integration of the regulatory elements of the Bill with the criminal justice system.

**Baroness Kidron:** Thank you. Matthew, I was struck by what you said about the law. We have some laws that we could leverage. Instead of creating new harms and new language about harms, the regulator could pull in certain rights language and certain equality Acts, and so on. Would you be more comfortable if the construction was, "We are trying to deal with this upstream. We do not want to see vast abuse, and if you are systemically abusing people, under certain circumstances, we will reach out to existing and agreed rights processes and legal processes"? Would that make you more comfortable?

**Matthew d'Ancona:** Much more comfortable—I think you have put it brilliantly—in the sense that I am not really sure why a Bill of this sort should seek to reinvent the wheel of what constitutes harm, at least not in the first instance. Except for free speech absolutists, there are plenty of perfectly legitimate, legislated, in precedent or in common law, restrictions on speech that now really need to be put into action. The problem is a legislative structure that matches the technological revolution rather than identifying speech that is harmful.

The only caveat, which has been mentioned once or twice, is the sheer scale of, to put it crudely, a Twitter pile-on, which really has no immediate precedent in human history. It may be that we have to revisit what constitutes speech that is harmful, but I do not think this Bill does so in sufficient detail. Just to insert virality as a cause of emotional harm, almost as if it is a given, without much discussion behind it, is a big step in the curtailment of speech. I like your formula very much.

**Baroness Kidron:** I think I am going to stop while I am winning.

**The Chair:** The last question is from Darren Jones.

Q142 **Darren Jones:** I want to ask the witnesses whether there are any objections to moving away from a high number of statutory duties of care to an overriding duty of care. You only need to answer if you have an objection to moving to an overriding duty of care. I accept your silence as agreement, which is helpful. In doing that—and this is not necessarily my view—if we decided that the whole issue around exemptions for democratic content and democratic processes needed further thought, and therefore took that out of the Bill but instead put in it an obligation on Ministers to deal with it by a certain date, would it fundamentally undermine an overriding duty of care if we took out at the start of the implementation of this Bill some of the exemptions around democratic content and democratic processes? Again, I will accept silence as acceptance.

**Professor Richard Wilson:** No, it would not because there would still be all the existing provisions about existing illegal content—terrorist content, child sexual abuse, threats, defamation. All of that still exists, plus we have not mentioned the Racial and Religious Hatred Act, which is salient in British law. There are already existing statutory mechanisms to define illegal speech. What about just getting those implemented?

**Darren Jones:** Unless there is anything further that anyone wants to add, not for the sake of filling time—

**Silkie Carlo:** I would be concerned if by merging duties of care you mean merging the duty of care regarding lawful and unlawful content, for example. I think we have covered that ground, but that would be my objection.

**Darren Jones:** Understood. That is everything from me, thank you.

**The Chair:** Thank you to all our witnesses not only for your very interesting contributions and discussion this morning but for your forbearance during our technical break.