



# Digital, Culture, Media and Sport Sub-Committee on Online Harms and Disinformation

Oral evidence: Online Harms and Disinformation,  
HC 234

Thursday 4 June 2020

Ordered by the House of Commons to be published on 4 June 2020.

[Watch the meeting](#)

Members present: Julian Knight (Chair); Kevin Brennan; Steve Brine; Philip Davies; Clive Efford; Damian Green; Damian Hinds; John Nicolson; Giles Watling.

Yvette Cooper, Chair of the Home Affairs Select Committee, was in attendance.

Questions 114 – 231

## Witnesses

[I](#): Dr Megan Emma Smith, consultant anaesthetist, Royal Free London NHS Foundation Trust; and Thomas Knowles, advanced paramedic practitioner, NHS 111.

[II](#): Leslie Miller, Vice-President of Government Affairs and Public Policy, YouTube; and Derek Slater, Global Director of Information Policy, Google.

[III](#): Monika Bickert, Head of Product Policy and Counterterrorism, Facebook.

[IV](#): Nick Pickles, Director of Public Policy Strategy, Twitter.



## Examination of witnesses

Witnesses: Dr Smith and Thomas Knowles.

Q114 **Chair:** This is the Digital, Culture, Media and Sport Sub-Committee on Disinformation. Today's hearing is on Covid-19 and the impact of disinformation on the pandemic and the public. We have two panels today. We are going to have a panel of frontline NHS workers, and then the second panel is going to be social media companies. Before I move to the first panel, I am going to ask members to declare any interests. Do members have any interests they wish to declare? No. Fine, thank you.

Our first panel consists of Dr Megan Emma Smith, a consultant anaesthetist at the Royal Free London NHS Foundation Trust, and Thomas Knowles, an advanced paramedic practitioner at NHS 111. Good afternoon and thank you for joining us today. I am going to kick off by asking you what experiences you have had of misinformation and disinformation during this pandemic.

**Thomas Knowles:** I have had lots of frontline experience via 111, insofar as we are, by and large, a first point of contact for people seeking health information, people who might not want to engage with other more formal methods. There have been lots of instances where online misinformation has mirrored my experience, my patient interactions and things that people are concerned about when they contact us with respect to things like the use of various medications, whether or not things are greater or lesser risks at different times. There has been a palpable and very clear delineation between the kinds of concerns people have at different times when accessing health services when different forms of misinformation are circulating.

**Dr Smith:** I agree with much of what Tom has said. There has been a lot of confusing and conflicting medical evidence out there, and lots of misinformation. I work at a slightly different point in the process from Tom. He is at the front end and I am very much at the later stage of a patient's passage through their unfortunate Covid experience, so I see them when they are extremely sick.

What I have seen are a lot of patients who are not presenting to hospital. They are presenting very late on in the illness because, in some cases, they have been afraid to come to hospital or they have believed online messaging that suggests the illness is not as serious as it really is—that it is just like a bit of a cold, a bit of bad flu—or they have tried alternative remedies, such as gargling salt water. There have been all sorts of things out there. By the time they come to me and I have an interaction with them, they are unbelievably sick, and the patients I have seen have all required intubation. Many of them, within half an hour to an hour of coming through the doors of the hospital, have had to be intubated and put on to ventilators. That is my experience of how it has impacted on my patients.



## HOUSE OF COMMONS

Q115 **Chair:** Dr Smith, what failings have you seen that prompted you to write to the social media companies?

**Dr Smith:** Specific ones?

**Chair:** Yes.

**Dr Smith:** I, and our EveryDoctor members, represent 25,000 UK doctors. As I have said, we have seen multiple bits of misinformation that relate to things like home remedies, and they can be either things that are directly harmful, such as President Trump's disinfectant-related suggestion, or things that are—

Q116 **Chair:** Sorry, just to stop you there, you have seen patients come in that have been effectively saying to you that they have followed that advice? Is that right?

**Dr Smith:** No, not in relation to the disinfectant. What I can say is that doctors that we represent—particularly GPs and frontline doctors, so emergency department doctors—have seen patients who have said, "I did not want to come to hospital because I was afraid." One of our members, for example, has direct experience of the Asian and Pakistani community up in the north-west where, via WhatsApp, there is a lot of information circulating suggesting that, if you go to hospital, doctors will not look after you, and in fact what they will do is effectively kill you: they will give you an injection, make you comfortable and leave you to die. Lots of people have not come to the hospital because of that.

**Thomas Knowles:** There have certainly been many circumstances in which I have had personal conversations with people who have stopped taking blood pressure medication, or who have stopped taking particular kinds of pain relief that they may have been prescribed long term for various risk factors on the basis of exaggerated information, at the very least, in that respect. As Megan has rightly pointed out, there have been lots of bits of information circulating that have eroded trust in the health service, which is a major issue. Again, there have been lots of people where I have had to have quite extended conversations, trying to convince them that we are on their side, because the information that they have received has been otherwise, and the suggestions from information they have read or have been exposed to have been otherwise.

Q117 **Chair:** I presume part of that is also the fact that some people believe that, effectively, they may go through some sort of selection process, in terms of age and that sort of thing, as to whether or not they receive proper treatment. Is that correct?

**Thomas Knowles:** Yes. Certainly some of that information has been around processes that exist and are well established, like DNACPR choices, like people's concerns that they will perhaps be mistreated or risk-stratified because of their ethnicity or cultural background. That certainly has played into a lot of the structural harms that we see day to



## HOUSE OF COMMONS

day in the health service. Concerns have in some cases run along those lines, yes.

Q118 **Chair:** Dr Smith, to what extent are your concerns shared by other healthcare professionals?

**Dr Smith:** Certainly a lot of my colleagues are concerned about the misinformation that is circulating both now and prior to Covid. As I said, the doctors that we represent are deeply concerned. Right across the board, we represent doctors ranging from me, as an anaesthetist, through to GPs and doctors in all specialties, and they are all getting similar stories from their patients and are having to deal with them, whether that is in the community or in hospital. It is a big problem, and it is a problem that I think is being experienced by lots and lots of colleagues.

Q119 **Chair:** When you say, "lots of colleagues", is there any way that either you or Mr Knowles can quantify that number?

**Dr Smith:** I do not think I could do that here, no.

**Chair:** Do you have any percentages or that sort of thing?

**Dr Smith:** I do not think so.

**Thomas Knowles:** I agree with Megan insofar as I do not think I would be able to place a number on it, but it is anecdotally regular enough that it is a conversation I have had multiple times a week. It is something where as soon as we asked our cohort, our peers about it, we had a significant response quite quickly. While we cannot quantify it, it is something that is very much present and urgently felt.

Q120 **Chair:** Is it an everyday experience? Would that be fair?

**Thomas Knowles:** I would say it has tapered off a little bit now as general fear around Covid has tapered off, but certainly there were days where I was having multiple calls per day dealing with elements of misinformation.

Q121 **Chair:** I imagine that is incredibly distracting when you are trying to save people's lives.

**Thomas Knowles:** Part of the role of 111 is to provide that level of health information and general health interaction. What I would say is that it speaks to how much of a burden misinformation has the potential to place on health services. I can speak to one person for 10 minutes and have an influence on that one person's experience of healthcare. The Committee is probably familiar with the pandemic documentary that was circulating on YouTube, and one version of that had 40 million views within 48 hours. That is 25,000 people in 10 minutes. I cannot speak to 25,000 people in 10 minutes, so that level of exposure is why I think so many of us are so concerned that we need to take action to identify those clear harms that people are experiencing as a consequence.



Q122 **Steve Brine:** Hello, Dr Smith. Hello, Mr Knowles. Thanks for coming along and joining us. What areas of the health service do you think have been most hit by this misinformation? Obviously, Dr Smith, you said the EveryDoctor UK organisation represents people across primary and tertiary care. Which areas of the health service have been most put under pressure by this?

**Dr Smith:** You can almost divide it into two broad categories. You have heard from Tom about the experience of the out-of-hospital, frontline element of the NHS, and they have clearly had a huge amount to deal with in terms of patient interactions.

From my perspective within the hospital system, the impact has been huge, and that is primarily because we have seen so many late-presentation patients. At the point in time when they come through the doors of the hospital, because they did not want to come to hospital, they are so, so sick. They are unbelievably unwell. I think it has had a massive impact in that sense, but at both ends of the health service.

**Thomas Knowles:** If you will forgive me for saying so, I almost feel as though the question presupposes that the focus needs to be on areas that have experienced harm. My major fear is the burden it could have the potential to place on public health, in terms of people's trust in any practices that are put in place, but also just in terms of people's engagement with health services.

Referring specifically to your question, it is difficult to split off whether demand is Covid-related or specifically misinformation-related, but I would refer back to what I said before in that we saw a big spike in calls and my anecdotal experience is that lots of those calls had a component of misinformation related to them.

Q123 **Steve Brine:** I am interested that you said you spend a lot of time trying to convince people that you are on their side, as you put it. If you look at the "clap for carers", or if you asked the public at any time, let alone at the moment, whether people who work in the health service are trusted professionals, I think you would get a very high trust level from them. What you have said is worrying.

**Thomas Knowles:** I would absolutely agree. By and large, the vast majority of the public do have a significant element of trust in the NHS, which is very welcome—that is what we would like to see. We would like to think that we are providing a service that engenders and deserves that trust. However, the reality is that in public health measures in particular, if we look at vaccination, it only takes a relatively small number of people to be mistrustful, and to be concerned about whether or not we are to be trusted and whether or not people should be following our advice, before we experience a set of circumstances that harms all of us and places all of us at risk.

Q124 **Steve Brine:** Have you noticed any different faith groups in particular? I



noted in our notes that Facebook and other social media platforms are spreading content that is targeted at UK Muslims, saying that a lack of body bags means they would have to be cremated, which caused a lot of upset. Do you have any experience of that?

**Thomas Knowles:** I cannot speak to that specifically, just because it is not an area of healthcare I typically deal with at the moment.

**Steve Brine:** Dr Smith, is there anything that you wanted to add?

**Dr Smith:** No, there is nothing that I have come across specifically. One of the things I would say is that, in terms of the nature of the work and people staying away from hospital, normally in my out of hours on-call work I am looking after emergency surgical patients and anaesthetising them for emergency operations. That work dwindled to almost zero during the first wave of this pandemic. It is not that those patients were not sick at home; it is just that they did not come into the hospital, and I think a lot of that is to do with fear about coming near the hospital. They are starting to come back now. I think that is also reflected in the non-Covid-positive excess deaths at the moment.

Q125 **Steve Brine:** Sticking with you, Dr Smith, on that fear of coming into the health service, we hear of concerns about strokes, about cancer presentations and about cardiac presentations. Are we storing up huge problems down the line? Do you think members of the traditional media—we are going to talk to Facebook and Twitter next—have any responsibility for misinformation about Covid-19?

Let me give you an example. The health trust in the area I represent has never once run out of PPE, yet if you were a constituent of mine or if you were in my part of Hampshire watching the national news, you could quite easily be forgiven for thinking that the health service was overrun, that staff were unprotected and that if you went into the health service you would face exposed staff without PPE. Do you think members of the traditional media have any responsibility for misinformation around Covid-19?

**Dr Smith:** I think they do. Any media outlet and any news-sharing organisation or fact-sharing organisation has a responsibility. In part, the differences that you are describing are probably down to regional differences. My own personal experience, and the experience of lots of doctors in London, was of being right up to capacity. When I say capacity, I mean over double our normal intensive care capacity, so of patients in non-normal ICU scenarios, of being extremely close with PPE and, at times, of not having enough. I think there were regional differences. London was hit first and hit harder than many parts of the country, but I completely agree.

Almost to flip that on its head, I would say that the print media and the television media are regulated and have obligations. If some of this misinformation appeared in their pages or on their screens, steps would be taken. Why should these other platforms be any different? I



## HOUSE OF COMMONS

completely appreciate that they do not write the lies, they do not compose the lies, but they do facilitate the distribution of them, and that is what we have to get rid of.

Q126 **Steve Brine:** But you do take my point? As a former Health Minister, it is a national health service, but there are thousands of outposts across its different sections. What is true in one area and what is true in London is not necessarily true everywhere, yet Twitter does not differentiate and say, "This is a specific tweet relating to London."

**Dr Smith:** It depends. You can look around Twitter or Facebook and find information that is specific to an area or is generic for the entire United Kingdom. I have not dug into that on an individual basis, but it depends who is tweeting, where they live and what information they are tweeting, which is kind of the nub of the issue, isn't it? The tweet is only as good as the underlying information.

Q127 **Steve Brine:** Finally, do you think that some of the sources of misinformation we have seen around this have been motivated by money and monetary gain?

**Dr Smith:** I do not personally have any experience of that, but I know that Tom does, so it may be good to ask him.

**Thomas Knowles:** In doing some of the research I have done for appearing before this Committee, I have certainly come across people who, while not necessarily directly selling products associated with Covid, for example, are painting a picture of themselves as perhaps holistic health practitioners or as people who claim to be practising integrative healthcare. They use all sorts of euphemisms and all sorts of sciencey language while just peddling mistruths, so people who are pursuing anti-vax agendas or agendas that promulgate 5G conspiracy, those kinds of things. As a sideline, they sell health-related products, quack cures, things that could very easily draw an implication between what they are selling and the lies they are peddling.

They are very clever, lots of these people, not to cross the line into deliberate false advertising, and that is what is so insidious about lots of this misinformation. It is very deliberate, and it is very carefully constructed. Some of these people have very large followings. There is a registered nurse in the UK who I have come across in the course of this, who is one of those people that is mis-selling health, using the veneer of trust, which other nurses have deservedly earned, to manipulate the public. It is all well and good my turning around and saying that about this person, but she has 11,000 followers on Facebook. As before, I cannot contend with that on an individual level. I can contend through the regulatory process of her registering body, and I intend to do that since I have discovered what she is up to, but on a more foundational level, only the platforms that are hosting her content can meaningfully take action against her.



## HOUSE OF COMMONS

**Steve Brine:** You keep ploughing your furrow on that one, because it sounds like good work. Trust us, we will get to the platforms next.

Q128 **Chair:** Mr Knowles, on that point, do you recognise that the social media companies are making money from this?

**Thomas Knowles:** Yes, absolutely, and that is one of the reasons why I think it is so important from an ethical and moral basis, insofar as the use of algorithmic generation of where people are seeing their content. If somebody watches something that might be a bit flat-earthly, maybe they will be interested in something a bit homeopathic. It is effectively a monetisation of the content that is published on those services. When we are looking at monetisation of content that actively engenders a social harm, that is actively damaging to public trust and to public health, I think we are looking at a moral obligation. That cost cannot be borne by the public purse when we are looking at organisations that are turning over billions of pounds a year.

Q129 **Chair:** Effectively what you are saying is that people on the frontline are having to clean up the mess that is caused by disinformation, which is carried by these platforms that are in on the game. They are making money from this. That is basically the crux of the matter, and it needs to stop.

**Thomas Knowles:** Sorry, I missed the very first part of that.

**Chair:** Sorry, we have a bell going in the background, which makes it more difficult. The point I am trying to make to you is that the nub of what you are saying is that you are finding that this disinformation is creating enormous difficulties for you on the frontline, and it is more than galling that you are looking at these platforms and they are themselves making money from this, and you think that in itself is a moral outrage.

**Thomas Knowles:** I absolutely do, and it is not just that they are making money from it. I could almost live with them just making money from telling people things that are not quite true. I object to them making money from people being harmed by that, and I object to them having the capacity to damage public health in such a convincing way.

If we look at, in the past, the furore around MMR vaccination and the false link to autism, we saw outbreaks of measles, mumps and rubella as a consequence because we lost the effect of the vaccine on a community level. Just thinking about Covid, vaccination is one of our ways out of this pandemic, and what we cannot have is social networks reducing themselves to merely being bastions of free speech, when what they are doing is profiting off a system that places everybody, as I have said previously, at risk of increased harm.

All of that cost is borne at present by the public purse in terms of NHS services. The resource in terms of policing it is currently borne by DCMS in terms of its Disinformation Unit, and all we can do at the moment is turn around and ask nicely for them to remove things. Some social media



networks have done that and, in doing so, I think they have acknowledged that this problem exists as an online social harm. What we are asking for is for that to be fixed in a legislative and regulatory framework so that there is no meaningful recourse for social media platforms to remove themselves from that social responsibility.

**Q130 Giles Watling:** Thank you, Dr Smith and Mr Knowles, for coming today. My take on this, just touching on Steve Brine's issue about trust levels, is that of course we all recall the doctor who ridiculed the MMR jabs and the damage that caused. That came from a doctor, and therefore it was a trusted source, which is why it spread. Would you agree that perhaps there is a mental health issue, that people pick up on that and want to take it further because it makes them feel valued and informed in some way, and that somehow we need to reach out to those people who spread this misinformation further?

**Dr Smith:** I see. You mean someone who has taken the information from a doctor and is then propagating it?

**Giles Watling:** Precisely.

**Dr Smith:** I had not thought about it, but I think it is a valid point, certainly. You will have seen from the Avaaz study—and it is one of the reasons that we as a group signed the letter—that you have to take two approaches. One is, as soon as you have identified what the misinformation is, to correct it. We are not saying clamp down and get rid of free speech, but you have to go back and correct it, and hopefully that will go at least some way towards helping those sorts of people who might be getting a bit of a kick out of putting themselves forward as a pseudo-expert, and/or hopefully it will at least flag up for those to whom they are proffering this misinformation that it is not accurate and is not true.

Doing it that way is almost like treating a medical problem, but then you almost need to vaccinate the system by taking it out of the algorithms, stripping it out. Avaaz talks about it as "detoxing the algorithm," but we completely agree. If it is sitting there, it spreads faster than any virus we have come across, and it is across the globe in a matter of minutes to hours, as Tom was saying in relation to the YouTube pandemic video. That is something we cannot compete with at all. We do not have the time or the resources, especially not at the moment. We are too busy looking after these sick patients.

**Q131 Giles Watling:** You are putting it firmly at the feet of the social media platforms in that statement.

Moving on, Elon Musk tweeted back in early March that chloroquine was worth considering, then he said that hydroxychloroquine was probably better to use. His Twitter account has 35.5 million viewers. How do we take that on? How do we get the message out? I feel there is an issue here that if we politicians rightly stand up and say, "Look, this is fake news," there are people out there on the social media platforms who say,



## HOUSE OF COMMONS

"You are saying it is fake news because that is a Government conspiracy," and we cannot win. The message needs to come from people like you. Do you agree?

**Dr Smith:** Absolutely, I completely agree. That is where the question of truly independent factchecking comes in, and it needs to be almost specialty-specific. It is no good having an economist factcheck some medical information, or vice versa. You need to have proper dedicated teams, and I would suggest that it is the social media platforms who should be funding those teams, given that they make such massive profits generally, and also that they make such massive profits from so much of this misinformation.

I think we all accept that sometimes it can be tricky, and sometimes it can take time to establish that something is, in fact, inaccurate and is misinformation, but as soon as that is established you have to go back and correct the record and you have to stop it being propagated any further. Unless there are proper formal factchecking entities that can do that, it is very difficult. I totally take your point. It puts politicians in an incredibly difficult position because it is an easy and, if I can say so, slightly low blow to come back at you and say, "You are politicians, you are the Government. Of course you have a vested interest in this." It needs to be objective and it needs to be independent, and demonstrably so.

**Giles Watling:** It means that we have to get you guys front and centre. Thank you very much.

**Dr Smith:** Possibly. We have quite a lot on our plates at the moment, but happy to help thereafter.

**Giles Watling:** Very good. Do you have any comments on that, Tom?

**Thomas Knowles:** I would agree broadly with what Dr Smith said there. I would add two caveats just to clarify what we are talking about. The first is the example that you gave of Mr Musk talking about hydroxychloroquine or related drugs being something that might be worth considering. It is important to be clear that what we are talking about is objecting to things that are demonstrably untrue, where they are demonstrably counterfactual. It is possible, in the course of whatever framework arises from this, that Mr Musk's comments would be considered a matter of opinion or a matter of consideration. It could very validly be said that hydroxychloroquine is worth considering. What we are trying to address are aspects of fact. If he had turned around and said, "Hydroxychloroquine is curative of Covid" then that is clearly wrong.

The other important thing that we need to focus on is this idea of retrospective correction. The platforms need to be ensuring that anybody who has seen that content through their feed has it flagged as being incorrect, so it is tagged with a screen or a flag or whatever so that people can refer back to what it is that was wrong, rather than just being told facelessly, "By the way, this is the correct information." That way we



discredit an uncreditable source and we make sure that people are then provided with factual information, as Megan says, to correct that record.

We also need to be quite careful to make sure that we are eliminating information that is harmful. It might very well be the case that information that is wrong does not necessarily confer a great deal of harm. In that case, it is reasonable just to flag it and say, "This is not true. We do not have any evidence for this." However, if information is actively dangerous, then I know from speaking with Megan previously that both of us would support that just being flatly removed as content and not even allowing reference to be made to it in future, while also retroactively correcting and flagging it for everybody who has witnessed it prior to deletion by the platform.

Q132 **Giles Watling:** Again, it is for the social media platforms to deal with?

**Thomas Knowles:** Yes. Just on your earlier point about mental health and people possibly having delusions of grandeur, or whatever, that lead them to post things, there may be something in that, but I think it is possibly a secondary issue. Certainly, if possible, we should be reaching out to those people with mental health interventions if that is appropriate. I do not have any information on how relevant that might be in terms of volume. There are definitely people who are deliberately spreading counterfactual information and deliberately profiting from it, so we need first and foremost to focus in the main on just making sure that information is not harmfully circulating. I take the point generally speaking around mental health conditions, but it is possibly a slightly tangential issue, and that certainly would be an issue for healthcare if it is one that is present.

Q133 **Clive Efford:** Apologies for being late in joining. If you have covered what I am asking, feel free to step in. I first came across Miracle Mineral Solution about 10 years ago, when the parent of a child with autism came to me, really concerned about the way it was being marketed as a cure for autism. I then discovered all sorts of things—it was being marketed as a cure for cancer and all sorts. I pursued a number of avenues. I pursued the agency responsible for registering medicines, and they said, "It is not a medicine, so we cannot deal with it." I went to the Food Standards Agency, and they said, "We cannot deal with it because it is not a food." The regulators back then, before all this on social media, were finding it very hard to come down on this miracle cure.

How do we deal with this when the regulators themselves say it falls between all their stools? Who has responsibility for stepping up and saying, "This is common or garden bleach that you pour down your drains, and you should not be pouring it down children's throats"? How do we deal with that situation? As much as we know it is wrong, what can be done?

**Thomas Knowles:** I would personally take the view that you are asking about two slightly separate things. Absolutely, there are regulatory



## HOUSE OF COMMONS

frameworks that exist around medicines and around other substances for human health. For example, you cannot advertise prescription-only medication in the UK, but you can advertise over-the-counter medicines. What we are talking about is holding social media organisations responsible for regulating fact, rather than regulating whether or not these products should be available.

It is perfectly possible, for example, for somebody to start talking about a medicine that exists and is widely used for a wide variety of purposes—paracetamol, for example—and then to make counterfactual claims about that. There is no crime in doing that in the UK, as long as it is not advertised as such formally, but people could have a conversation about it on Facebook, on Twitter or wherever and spread misinformation about it perfectly legally, but we still think the social media organisations have a responsibility for policing that.

Yes, I certainly would agree with you most heartily, and I would be more than happy to try to work with you on any of this if there is anything along those lines I can do to help. The sale of things like bleach for the treatment of autism is abhorrent, and it obviously is not something that we should be supporting or in any way advocating, but it is a separate issue. It is not an issue of regulation of whether or not it is a medicine. It is a question of regulating whether or not it is true.

**Q134 Clive Efford:** Yes, but if you cannot get the agencies responsible, it throws up a big challenge. I do not want to be an advocate for defending Facebook or Twitter, or anyone—they can do that themselves—but in these circumstances, when you challenge these authorities, they can deal with it. The Genesis II Church was deliberately set up as a church as a device to get around the law, and there is evidence of him saying that that is why he set it up and advising people in this country that that is the route you should go down if you want to sell this MMS here in the UK. In fact, it was only trading standards that got to grips with it at a local level, location by location. This throws up a problem when these bogus companies are manipulating the system. How do we expect social media companies to police it?

**Thomas Knowles:** What I would say is that social media companies are the ones who have produced and managed the system that is being manipulated. If they are able to control their algorithms to promote content in such a way that people have repeated access to false information, and if they are able to host an enormous amount of data while employing some of the best and brightest data scientists in the world to enable them to do that, I struggle to believe that they would not be able to manage their algorithms sufficiently that it deprioritises those outlooks.

It is worth saying that I do not think anybody has any expectation of perfection in this. We are mostly focused on content that has experienced virality and has become widespread. That is where we and certainly Avaaz would like the focus to be, on content where thousands, tens of



thousands or potentially hundreds of thousands of people are seeing it. Yes, absolutely, there are lots of holes in the regulatory framework, and I completely agree that things fall through the gaps between different agencies in a number of ways. It is almost both demeaning to them and letting them off the hook to suggest that social media organisations, as they stand now, would not be able to deal with it if they had to.

**Clive Efford:** Tom, I will accept your challenge. If you want to contact me, I will take up the cudgel against MMS again.

**Dr Smith:** Give it a go.

**Chair:** That concludes our first panel. I want to thank Dr Megan Emma Smith and Thomas Knowles for your evidence today. It has been very helpful and very interesting to hear from the frontline. We are going to take a few moments' adjournment now to set up our second panel, but please stay on the line.

## Examination of witnesses

Witnesses: Leslie Miller and Derek Slater.

Q135 **Chair:** Welcome to the second panel of the Digital, Culture, Media and Sport Sub-Committee on Online Harms and Disinformation. We are looking at Covid-19 and disinformation. We heard in the first panel about the effect of disinformation on the work of frontline healthcare professionals. In this panel, we will be hearing from the social media companies. The first guests are Leslie Miller, Vice-President of Government Affairs and Public Policy at YouTube, and Derek Slater, the Global Director of Information Policy at Google. We will then have representatives from Facebook and from Twitter. For this part of the session we are being joined by Yvette Cooper, who is Chair of the Home Affairs Select Committee. Welcome, Yvette.

I am going to kick off to you, Mr Slater. How have you changed your search algorithms to ensure the right information is there and not health disinformation, and how are you ensuring that public health messages are at the top of your searches?

**Derek Slater:** Thank you, Chair, and thank you all for the opportunity to be here today. The core of our business is delivering quality, trustworthy, relevant results, and nothing could be more important than the case of information about health. This is something we have worked on for quite a long time, and we benefit from interactions like this and from co-operation with Government in continuing to improve. For a very long time we have worked to ensure that, for health information, we are raising up high-quality, authoritative sources and lowering low-quality, misleading information. That has continued to be true in this case under Covid-19.

In March we launched, with the NHS, knowledge panels in Search, which raised up, across a number of different bits of medical information, authoritative sources around prevention, treatment, symptoms and so



## HOUSE OF COMMONS

on. We have also created dedicated sections in Google News and on Search for authoritative news sources around Covid. We have also worked with the Government to promote the campaign to help keep people safe, displaying that on our homepage. I believe that campaign received over 40 million impressions.

Q136 **Chair:** The next question is for Leslie Miller. Academics have told me that Chinese state media and Russia Today content is now being routinely repackaged on your site. What are you doing to address more complex issues of disinformation resulting from this style of infiltration?

**Leslie Miller:** First, I just want to say that I am sorry we cannot all be together, and I very much appreciate your willingness to do this over video conference at a time that works for those of us in California and for you. Thank you very much. It is important for me to hear your concerns and to answer your questions.

On this first one—and this is not just specific to YouTube—across the company, over the last several years, we have focused on how we can better address and tackle state-sponsored disinformation. We just put out an updated blog post last week that references our Threat Analysis Group. TAG devotes its time to reviewing, monitoring and conducting intel as it relates to Government-sponsored disinformation campaigns. Some of it can just be spamming material, and other types of material can be about trying to push a point of view. Essentially what we do is we look at the information and determine if it is co-ordinated. If it is meant to be deceptive and things of that nature, we take action on it. On YouTube, based on this process, we have removed more than 1,000 channels for those reasons.

Q137 **Chair:** Don't they just pop up again though? This is the point about repackaging. What are you doing in that respect?

**Leslie Miller:** We are diligent. We work all the time. This is not something we let up on. It matters for us across the globe to make sure that we are monitoring and taking action on this. You are right, trying to stay ahead of these bad actors is a challenge, but it is something that we are committed to doing and that we have made some progress on. There is always room for improvement, but it is something we work on day and night.

Q138 **Chair:** Would you say that your approach has changed towards Chinese and Russian actors?

**Leslie Miller:** The way we look at this is not specific to looking at any one country or any one country's Government. Instead, what we look at are the signals that suggest there are these deceptive practices, and we are taking action on them, regardless of where they are from.

Q139 **Chair:** My final question before I hand over to Yvette Cooper is for both of you. YouTube has a more granular report function than Google Search, allowing users to report specific videos and state why they think content



is illegal, misleading or, for example, in breach of copyright. What is the reason for this disparity between your two businesses? Has Google thought about taking a similar approach to YouTube to identify misinformation, disinformation and low-quality search results?

**Derek Slater:** Search is a reflection of the web; it indexes the web. It is not content we directly have control over or are responsible for, but we certainly take action on illegal content when we are sent notices for removal. As I said at the outset, we work very diligently to raise up authoritative sources and down-rank things that are low quality or misleading. We rely on a range of different signals to do that well. We have on every search page a place for people to send feedback, and we then take that feedback into account.

On particular features like the knowledge panels I referenced, we also have specific forms where people can send feedback to say whether something is inaccurate, misleading or low quality. That is something we experiment with. This, of course, is a supplement to a much broader process that we use to improve our search ratings over time, working with not just our employees in-house but also with search quality raters around the world who look at page rankings and at different pages and help guide and inform our efforts of what looks to be a high-quality result versus a low-quality result, based on public guidelines.

Q140 **Chair:** You say you expand on it, but why don't you just do what YouTube does in that respect? It is far from a perfect example by any stretch of the imagination, but why don't you just take that same step?

**Derek Slater:** Again, we have experimented with and implemented a number of different features. What we see is that while the general feedback we get on a page can be helpful, it is not as helpful or as informative as relying on the search quality raters that I mentioned. The search quality raters are trained and tested on a set of consistent detailed guidelines, whereas ordinary people might have very different senses of what is low quality versus high quality, so it is not as strong a signal. We are also looking—

Q141 **Chair:** If it works for YouTube, why doesn't it work for you?

**Derek Slater:** What we are trying to do on Search is come up with solutions that are scalable. It is not enough. If we get a signal or a feedback on one result, for one query, that is not going to help us solve at scale for the trillions of queries and hundreds of billions of websites that we are indexing. We want to come up with solutions that work at scale. If you do searches for Covid-19, you will see that we are working at scale to try to create those high-quality, authoritative experiences.

**Yvette Cooper:** Thank you to the Committee for having me, and thank you to the witnesses. My question is for Leslie Miller from YouTube. What are you doing to make sure that your algorithms are not promoting on the homepages material that is misleading or damaging to public health? Not just when people put in a search, but on the homepages.



## HOUSE OF COMMONS

**Leslie Miller:** Thank you for the question. We met over a year ago, so it is nice to see you again, albeit on the video conference. We do several things and, if you do not mind, I would like to give a bit of a broad picture. For YouTube, we—

**Yvette Cooper:** Briefly, if you can, just so a follow-up—

**Leslie Miller:** I will try to be as brief as possible. We rely on four pillars as it relates to responsibility: raise, remove, reduce and reward. The reason we approach it this way is we realise that it cannot only be binary of what stays up and what comes down. Instead, we have to look at our systems and the platform holistically. For example, as it relates to what we are raising, I can speak specifically to Covid. We have put systems and processes in place whereby we are raising the authoritative information on, for example, Covid. We launched info panels. Those info panels are based on the World Health Organisation and/or local authorities, such as the NHS in the UK. That info panel is available, and you can see it when there is a search on that topic. More broadly—

**Yvette Cooper:** What I am interested in is how you are preventing misinformation from being promoted. I can see that you are pushing authoritative sources. Just to give you an example that I raised with Ministers a couple of weeks ago, I searched for David Icke on YouTube, then I went back to the homepage. I then searched for 5G, just “5G” in, and went back to my homepage. When I went back to my own homepage on YouTube, the top recommended video for me, before I got to the authoritative sources on anything, was a conspiracy theory video titled, “Sickness from 5G cell towers, technology is killing us very slowly.” That is a misinformation video, and it is the top thing coming up on my homepage.

This week I have done the same thing. I searched for David Icke on my YouTube account. I then went back to the homepage, and the third or fourth video recommended to me on the homepage was an anti-vaccination video. I had not searched for anything to do with vaccinations. There was nothing in the video when I looked through it that mentioned David Icke. I do not know why I was having promoted at me an anti-vaccination video. Why would that be?

**Leslie Miller:** I cannot speak to that specific example, but what I can tell you is that we have moved in the direction, where there are videos or content that violate our policies, to work to remove them. Specifically as it relates to borderline content, if it is content that may come up to but not cross the lines of our policies, we work to reduce the availability of that. For example, in the UK, 99% of the time when somebody is searching for coronavirus or an iteration of that, the top 10 search results are from high-quality, authoritative sources.

In addition, I know there have been questions in the past as it relates to, “What is being recommended to me?” On the recommendation-driven watch time of this type of borderline content, it reflects less than 1% of



## HOUSE OF COMMONS

the totality of the watch time of content that is being recommended. As I was saying before, we use different levers to reduce the availability of borderline content while raising more authoritative—

**Yvette Cooper:** Yes, but it is not working on something so obvious as this. We heard from the health professionals in our previous panel about some of the misinformation and how damaging it is as a result. In both of these cases, this was not something I was searching for. It is not the results of a search; it is what was promoted to me on my homepage. That is what I find so shocking about this. YouTube has decided what to put on my homepage. It is not the result of something I have searched for, something that has slipped in at maybe ninth or tenth down a list of searches where the top eight were very sensible. It has come up there on my homepage, relatively high up, so that when I go back to home, this is what I am being encouraged by YouTube to watch.

Surely that is utterly irresponsible of YouTube, and I have repeatedly raised this issue with you and with many of your colleagues. I raised this with you in December 2017, in March 2018 and in April 2019 in Committee sessions like this as to what you are doing to stop your algorithms actively promoting misinformation or dangerous information to people.

**Leslie Miller:** I do not know if there was a question in there. I do appreciate that you, over the years, have shared your perspective and held us to account. What I can say is it relates to information that does not violate our policies. I can reference the David Icke channel, for example, as this is something that was raised with us. We looked at his content and we determined, first, that certain videos should be demonetised, but then we realised that our content policies as they relate to Covid were not going far enough and we expanded our content policies under the harmful and dangerous policy to include content that contradicts medical or scientific facts, such as that there is a cure for Covid or that 5G is a symptom or the means by which Covid is transmitted.

With David Icke, first we demonetised. We kicked him out of the partnership programme, then we started removing the videos, because we expanded our content policy to adapt to the dynamics that were happening, and then we ultimately terminated his account. There is always more to do in this area. As it relates to this type of borderline content, if it is not violative, we have made more than 30 changes in the last 14 to 16 months and we are always looking to improve in this area. I look forward to continuing to hear more from you on this.

**Yvette Cooper:** I feel like this is *Groundhog Day* and I am raising the same thing each time. Can I ask you if you could write to us, to both the DCMS Committee and the Home Affairs Committee, on what specifically you are doing to deal with the homepage recommendations? Not the searches, not the stuff that goes so far that you take it off completely, but what are you doing to prevent YouTube, on the homepage



## HOUSE OF COMMONS

recommendations, pushing information that is false?

**Leslie Miller:** Will do.

Q142 **Giles Watling:** This is for Mr Slater. Would you say it is the case that algorithms have more difficulty in identifying videos and images than they do text?

**Derek Slater:** I do think there are additional complexities when it comes to video and images, but it is hard to make a direct comparison.

Q143 **Giles Watling:** If that is the case, why does Google say that automated moderation is more effective than relying on user reporting?

**Derek Slater:** When it comes to how we enforce our policies for content that we host on our platforms, we use a combination of automated tools and manual review, both proactive enforcement of our policies as well as reactive responding to flags. It is not a matter of one or the other. In fact, both are essential. Automated techniques can be very useful in looking at broad sets of data at scale and identifying patterns. What they are not as good at is identifying particular context, and that is often very important or always very important when it comes to speech issues. Something that might be violative in one context, such as hate speech or incitement to violence, might be part of a legitimate news report or documentary in another context. We want to be sensitive to those distinctions, and that is why human review is always a very important part of the puzzle.

Q144 **Giles Watling:** The human review, I should imagine, will be more accurate when we are talking about videos and images.

**Derek Slater:** This can pertain across different media.

Q145 **Giles Watling:** Expert evidence has claimed that content that has been demoted by Google Search's algorithms often resurfaces to users within a year. What causes that to happen, and how can you address it?

**Derek Slater:** I am not familiar with the research you are talking about, but again what we strive to do with Search is raise up authoritative sources and demote and down-rank low-quality, misleading information. Those dimensions are laid out very clearly in our search quality rater guidelines, which our search quality raters, who are people all around the world, use to look at rankings of pages, as well as individual pages, to inform how we do our ranking. That is certainly something we have to keep improving over time. We have made improvements over the years.

Q146 **Giles Watling:** I gather it resurfaces, so that is something you want to get on top of, particularly if we are dealing with stuff that could be life threatening, like misinformation around the Covid virus.

**Derek Slater:** I would be interested to see the particular research and the examples you are talking about. Certainly if you do a search for a very particular item and you are looking directly for something, it may be higher in the rankings. If you are looking very particularly for, "I want to



find this particular theory” because the query suggests you want to find it, we want to be relevant and address that query because there may be legitimate uses, such as people doing research on these matters to try to learn and track these trends. We are sensitive that if you are not looking for it, we do not want to surprise and deliver that low-quality result. I would be happy to take a further look at that research if you would send it along.

**Giles Watling:** That would be good, thank you.

Q147 **Damian Hinds:** Can I just come back first to Yvette Cooper’s question? She asked why it is that an apparently unrelated or different type of content would appear on a homepage as a result of a previous search. Isn’t there a simple answer to that, which is that the algorithm looks at what other people who looked at content type A—in this case, David Icke—then went on to look at, which in many cases or some cases would be this anti-vax material? Isn’t the way the algorithm works at the heart of how echo chambers deepen and divisions widen? Isn’t it not only about what appears on the homepage but, specifically on YouTube, what comes next in autoplay?

**Leslie Miller:** Thank you for the follow-up question and, again, thank you for hosting this meeting. As it relates to the information that is available, be it through search and discovery or through recommendations, we have made a series of changes to make sure that what we are making available, depending on a number of signals, is high-quality, authoritative information, particularly as it relates to topics that are newsworthy or topics such as the pandemic. When we talk about borderline content, so this type of content, we have reduced its availability. I would be interested, and that is why I am happy to follow up Ms Cooper’s example. Because we reduced the availability and we are raising the more high-quality, authoritative content, it represents at this point less than 1% of the content that is on the platform. There are various signals that go into what comes up in search results, but we emphasise, particularly as it relates to newsworthy topics, topics such as Covid, that we are raising high-quality, authoritative content.

Q148 **Damian Hinds:** We look forward to your further correspondence. On news in general, it is not realistic for YouTube, Facebook or any other platform to be the ultimate arbiter of all that is true in the world, especially when there has been this explosion in the number of sources. Would you agree that it is much the best policy to put front and centre of news dissemination established news brands, which might be of the left or might be of the right, but at least you know where you stand? They have journalistic standards and, above all, they have a brand reputation to defend.

**Leslie Miller:** Yes, that is important. I would also say that, for platforms like YouTube, it has in some regards democratised the ability for people to share information and for there to be a diversity and plurality of voices. Particularly in a time of coronavirus, where people are looking to



## HOUSE OF COMMONS

do videos from their homes and hopefully—potentially—even earn an income by doing that, we want to be a platform that allows for that diversity of voices while they do not cross the line of our community guidelines. As it relates to these authoritative sources, when I mentioned earlier that, on the whole, when people are searching for something like coronavirus or something—*[Inaudible.]*

**Chair:** We appear to have a problem there.

**Leslie Miller:** When I say 99%, that includes *Guardian* subscribers—

**Chair:** Hello, Leslie. Sorry about that, we missed the last minute. You broke up. Would you be able to repeat the last minute, just your last point?

**Leslie Miller:** I just wanted to say that, in addition to allowing for the plurality of voices and potentially those earning income, we have focused on these authoritative outlets. What I was referencing is that, for example, *The Daily Telegraph* has now surpassed 1 million subscribers on the platform, in part because we are making sure that these high-quality, authoritative results are appearing when people are doing queries such as on coronavirus or Covid.

Q149 **Damian Hinds:** Thank you, I wanted to come on to that. You will recall that the Cairncross Review recommended a news quality obligation. Pending regulation, when it comes in, how will you make sure that quality news delivery is sustainable on your platform and ensure incidental exposure to it? What do you think is the fair split of ad revenue when high-quality, costly-to-produce journalism appears on your platform?

**Leslie Miller:** We work closely with news outlets. This is something that we have long been doing. I can say that even the BBC, for example, has close to 9 million subscribers, so we do work with the news outlets, not just for their content that is on the YouTube platform, but finding ways to make it easier for them to have video content on their own sites that is supported in part by YouTube. While I cannot get into the specifics of any one entity and the sharing of revenue, what I can say is that there are many people in the UK—not just those news outlets—who are making six figures. That growth, in terms of those creatives who are making six figures, has grown 25% since last year.

Q150 **Damian Hinds:** Isn't that revenue split at the heart of this? Leaving aside the BBC, which has a different funding model, who are the news outlets? Quality journalism is not cheap. I do not expect you to divulge individual contracts that you have with individual organisations, but what is roughly a fair split between the organisation that has employed journalists, done research, done factchecking and done the editorial, on the one hand, and the technology platform that has conveyed that on to people, on the other?



## HOUSE OF COMMONS

**Leslie Miller:** Again, I am not going to get into the specifics or reference what is a reasonable revenue split. What I can say is that we are close partners with news agencies, big and small.

Q151 **Damian Hinds:** What do you do to make sure that the delivery of that quality news is sustainable so that, five years and 10 years hence, enough revenue has been earned through these new media channels and those quality brands? You were talking positively about their role.

**Leslie Miller:** We continue to work with these news outlets. We have previously launched programmes, such as the digital news initiative and the global news initiative. We work with these entities to make sure that we are promoting their content in our top news shelf, the breaking news shelf, so we are exposing users to these outlets and helping drive traffic accordingly.

Q152 **John Nicolson:** Ms Miller, thanks for joining us. Do you agree with the Centre for Countering Digital Hate that, in a pandemic, lies cost lives?

**Leslie Miller:** I do not know the specifics of what you are referring to, and I am not going to comment on that suggestion. What I can say—

Q153 **John Nicolson:** It is quite a straightforward suggestion. The source is interesting, but ultimately it is a standalone statement. If lies are told that mislead people, can that cost lives? It is a very simple question.

**Leslie Miller:** Yes. One of the things we look at at YouTube is we do address the content. If it has the potential to cause real-world harm, we take action on that. 5G is a good example of this, where again we did not have ones that were specific to Covid and coronavirus. I was in the conversations with other executives, thinking through and making sure that our policies were adapting to these types of topics that could cause real-world harm, so we do pay attention to where there is—

Q154 **John Nicolson:** I am glad to hear that because it brings us nicely, or disastrously, back to the subject of David Icke, whom two of my colleagues have mentioned. You removed David Icke for violating YouTube rules about coronavirus information. However, he has appeared on a whole host of other YouTube channels and YouTube advertisements. He spouts anti-Semitic tropes, he spreads lies about vaccines and 5G masts, and this is why that matters because these lies make people sick and they lead to vicious attacks on telecom workers.

You said to my colleague, Yvette Cooper, and I quote you exactly, “We have worked to remove them, worked to remove the availability of that.” Let’s test that, shall we? He has posted on your site for 14 years. You promoted his content over 1 billion times and you are still doing it. Why?

**Leslie Miller:** On David Icke—and thank you for your question on this—

**John Nicolson:** My pleasure. We will get an answer, though.

**Leslie Miller:** Yes. On David Icke what we did, and what I was just referencing in the question before, is that we determined not only should



he not be monetised on our platform, but that the videos were violative of our content policies and so we terminated his account. I will say what that means is other channels can have David Icke content on their channels, but if it violates our policies we will remove it, so—

**Q155 John Nicolson:** Sorry, for those not from California “monetising” means being paid for. You are saying he can no longer earn dosh from spreading lies directly, but he can earn money from spreading lies on other YouTube channels. Yes, he can, because he can be paid by the others and of course, for the adverts you are running featuring David Icke, he is being paid for and earning money through YouTube. You are doing nothing about it, and you know exactly what you are doing. I think it is enormously cynical. Once again you have offered to write to us, which is you basically kicking all this into the long grass, playing for a bit of time, and this will go on and on and on because it suits your purposes to have David Icke on because he is clickbait.

**Leslie Miller:** I don’t know if there was a—

**John Nicolson:** The question obviously is, I suppose, that you kind of agree but cannot say so. You are right that it is a standalone statement, because I suspect we are not going to get anywhere. I do think that every time you appear before politicians, we are going to call you out on this and not simply accept a lot of empty blandishments that never go anywhere.

**Leslie Miller:** If you do not mind, I appreciate your saying that, and I want you to know that it is my job to hear from you and to share these perspectives with others at the company. On the specific example, I would be interested in seeing what the advertisements are, because if it is content that we have deemed violative, it does not matter what channel it is on, we will take action on it.

**Q156 John Nicolson:** Sorry, you think he is doing these ads for free?

**Leslie Miller:** No, no, I—

**John Nicolson:** Common sense tells you that, if he is doing adverts on YouTube, he is being paid for them. He is not doing them for nothing. It is all business. It is about spreading lies and earning money from them. I think I have explored this as far as I need to at the moment.

**Q157 Philip Davies:** I know time is pressing, so I will be brief and move on to a different theme. I direct my question again to Leslie Miller. I think we can all agree that something needs to be done about things where there is a clear factual inaccuracy. Whether you are doing that is a different matter, but hopefully we can agree that that should be the guiding principle. I want to ask you about things that are more matters of opinion. For example, can you tell me what your approach would be if people were advocating that there should not be any lockdown and that things should carry on as normal? What would YouTube’s approach be to people who made that case?



## HOUSE OF COMMONS

**Leslie Miller:** This is a great example, because we have had many discussions internally to make sure that we are providing the most relevant and helpful information as it relates to Covid. I would be interested also in your view, if you don't mind. There is a distinction between health authority guidance as it relates to social distancing and the things to make sure that people keep social distancing versus different authorities giving different opinions as it relates to lockdown, sometimes within the same jurisdiction.

Q158 **Philip Davies:** Yes, so what was the upshot of these internal discussions you had at YouTube? I appreciate you have had a discussion about it. What was the upshot of this discussion and, therefore, what is your policy?

**Leslie Miller:** Our approach is that we rely on local health authorities and their guidance, and where there aren't the local health authorities providing guidance we rely on the World Health Organisation. We do this in a number of ways where, for example, again we have done the info panels. We have made this information available in the top news shelf. We have done promotions on this, so having this source of authority, be it the local health authorities, like the NHS or the World Health Organisation, is the stand we have taken, making sure that we are pushing that information, particularly as it relates to social distancing.

Q159 **Philip Davies:** Just to be clear, I am probably a bit too simple for this, but if I uploaded something on to YouTube that basically argued, "All of this lockdown is ridiculous. We should not be having a lockdown" what would happen to that? That is what I am trying to get to. What would happen in that scenario?

**Leslie Miller:** If your content violated our policies around Covid, specifically on medical and scientific facts—so if you had suggested that there was a cure or that social distancing does not do anything, things of that nature that cross the line—it would come down.

Q160 **Philip Davies:** Saying that there should not be a lockdown, where does that fit on the scale?

**Leslie Miller:** That type of comment can stay up. If it is specific to social distancing and contradicts the general consensus from medical health authorities, it would be violative.

**Philip Davies:** Thank you. I know time is pressing, so I will leave it there.

Q161 **Kevin Brennan:** Mr Slater, I do not have time to pursue a line of questioning, but last month your witness from Google told the Committee, "Our advertising does not run on any websites that are peddling medical misinformation or are contesting the existence of the virus". Yet I have a report today from a campaign called Stop Funding Fake News, which is run by the Centre for Countering Digital Hate, and it names a number of websites—including ZeroHedge, Waking Times, Voice



of Europe, GreatGameIndia, Gnews and Global Research—which have Covid misinformation on them and are still having Google ads placed on them. How can you explain that? Will you commit to the Committee to look into that to make sure those ads are removed, and will you also give us a full list of Google Display Network websites so we can see which sites are being monetised in this way?

**Derek Slater:** I appreciate the question, Mr Brennan. To be very clear, we have strict policies for our publishers who run AdSense ads. That includes restrictions on dangerous or derogatory content, including harmful misinformation about medical cures, treatments or the transmission of Covid through 5G. We use automated and manual systems, and proactive and reactive systems, to try to address these. I am not familiar with the particular examples you are suggesting, but I would welcome your sending them to us. I want to be clear that this has been a very fluid situation where we have been having to, in real time and 24/7, look at our policies, re-evaluate them and see how we can improve. This dialogue is important to that, so I appreciate your question.

**Chair:** Thank you, Leslie Miller and Derek Slater, for appearing before us today. I know it is a very different time of day there, so thank you very much indeed.

## Examination of witness

Witness: Monika Bickert.

Q162 **Chair:** Our next witness is Monika Bickert, the Head of Product Policy and Counterterrorism at Facebook. Good afternoon, Ms Bickert.

**Monika Bickert:** Good afternoon. Thank you for having me here.

**Chair:** Thank you for appearing.

Q163 **Kevin Brennan:** Thank you for appearing before us, Ms Bickert. I was not on the Committee on the previous occasion you appeared before it, but I did look back at your evidence. Following up on that evidence, could you clarify for us whether, last time you gave us evidence, you knew about the Cambridge Analytica data breach?

**Monika Bickert:** I would have to go back and look at when I gave evidence and at what exactly was said. I was not—

**Kevin Brennan:** It was in 2018.

**Monika Bickert:** I can go back and look at the transcript and follow up with you.

Q164 **Kevin Brennan:** I would appreciate that, if you could. I realise I am dropping that question on you without your having had a look yourself, but if you could do that and let the Committee know whether, at that time, you were aware of the Cambridge Analytica data breach and when you had the opportunity to discuss it with Mark Zuckerberg, it would be



## HOUSE OF COMMONS

very useful to know in relation to subsequent events. I would appreciate it if you could commit to doing that.

**Monika Bickert:** No problem.

Q165 **Kevin Brennan:** Thank you very much indeed. How many engineers do you have working on online abuse, online harms and fake news at the company?

**Monika Bickert:** I do not have a breakdown of exactly how many are engineers. We have about 35,000 people across the company who are working on safety and security. I would say slightly more than half of those are content reviewers, meaning people who look at the content, assess our policies and make a decision, but also in that mix we have engineers, most of whom are on our integrity engineering team, and they are doing things like building mechanisms for people to report content to us. They are building the mechanisms by which we go out and find violating content. For a lot of areas now, whether it is child exploitation or terrorism content—even hate speech—most of the content we remove we find using these technical tools that the engineers are building.

Q166 **Kevin Brennan:** How does that compare with the number of engineers you have working on, for example, online advertisements and so on, as a number of people within the company?

**Monika Bickert:** I do not have a comparison on the staffing. I can say that 35,000 people working on safety and security in the company is, for our employee base, quite a large number.

Q167 **Kevin Brennan:** Is it likely to be more or less, do you think, working on online advertisements? Would you have an idea of that?

**Monika Bickert:** I do not, I am sorry.

Q168 **Kevin Brennan:** There has been a bit of publicity this week concerning some of your former employees who have written a letter to *The New York Times* in relation to the decision that was taken about President Trump's comments on Twitter and then Mark Zuckerberg's reaction to that. They wrote in the letter to *The New York Times* commenting on the leadership at Facebook, "They have decided that elected officials should be held to a lower standard than those they govern." They are right about that, aren't they?

**Monika Bickert:** I should be clear that I have not seen the letter. I can tell you that our policies, by and large, apply to everybody. We do have a third-party factchecking programme, and I can go through our approach to misinformation. I know it is—

Q169 **Kevin Brennan:** Don't do that this time. You have not seen the letter that the former employees have published that is on the CNN website and that is probably well known to everybody across the world? Are you aware of the existence of the fact that some former employees have written to *The New York Times* and that they also staged a protest this



## HOUSE OF COMMONS

week in relation to that decision?

**Monika Bickert:** I certainly know that there have been people unhappy with our decision, and that people outside the company—and even employees inside the company—have expressed that they did not like the decision, but I am not aware of a specific letter—

Q170 **Kevin Brennan:** The letter was signed by Meredith Chin, who is a former corporate communications manager at Facebook, Adam Conner, former public policy manager, Natalie Ponte, former marketing manager, Jon Warman, former software engineer, and they said in the letter, “It is our shared heartbreak that motivates this letter. We are devastated to see something we built and something we believed would make the world a better place lose its way so profoundly.” It is quite a shocking indictment from a number of quite senior former executives of Facebook in relation to their reaction to the company’s policy on the tweet that Donald Trump tweeted concerning shooting people. Does that in any way disturb you?

**Monika Bickert:** I think we have heard from people who have been disappointed with that decision we made last week. I can tell you that the policies that we have in place, the post did not violate our policies, and we have had those policies in place since I joined the company seven years ago. I do not know the people who wrote this letter, so I do not know when they were at the company, but I can tell you that these are longstanding policies.

Q171 **Kevin Brennan:** If anybody else had written the same tweet, you are saying they would have been treated in exactly the same way as President Trump, is that correct?

**Monika Bickert:** Our policy is that we allow people to discuss Government use of force. We think if Governments are talking about using force, people should be able to discuss that. I guess frequently there could be a safety reason that people would want to know what Governments are planning, but we think that certainly this is something that should be in the public discourse and, yes, that would be true of any discussion of Government force.

Q172 **Kevin Brennan:** You have had a virtual walkout of your employees this week, and you have had the letter to *The New York Times* from a significant number of former employees. Looking at it from the outside, it feels to me like there is something rotten in the state of Facebook. Am I wrong?

**Monika Bickert:** Again, I can tell you that our policies on this issue are longstanding. I have been in this job for seven years. I do not know the names that you just mentioned, but it sounds like they were with the company at some point. I do not know if they knew our policies then or if they disagreed with our policies then. I cannot speak to that.

**Chair:** I will just take my jaw off the floor over the fact that you did not read the letter to *The New York Times*.



Q173 **Clive Efford:** What is the motivation for someone to set up a fake account? Has Facebook ever investigated what the motivation is for doing so?

**Monika Bickert:** A big motivation is financially motivated scams. That is not the only motivation. You can imagine people trying to artificially amplify or make their content more popular by creating more accounts, but often what we see are scammers, people who want to mass create accounts so that they can flood people with links to lead them off site to ad farms or other ways to make money. This is an area where we have invested a lot of our technical tools to be able to stop accounts at the time of creation, so we now remove more than 1 million accounts per day at or near the time of upload.

Q174 **Clive Efford:** You admit to 5% of the accounts on Facebook being fake accounts. That figure seems very high, and others claim it is even higher. Is there something wrong with Facebook—the way Facebook is policing this—that there are so many fake accounts?

**Monika Bickert:** We are very aggressive in finding those, and in fact the numbers we are reporting are based on the aggressive measures we are taking to find and understand the prevalence of fake accounts. To drill down into that, we have started publishing reports every six months where we say, “Here is exactly how many fake accounts we are finding and here is what we think the prevalence is” so we are making all of those numbers public.

In the past couple of years, if you track those reports, you will see that there is a sharp increase in the number of fake accounts that we are detecting. That is because of the investment we are making in building those technical tools, so I think the tools have come a long way. Now, are they perfect? No. There are still going to be people who will successfully get around the barriers that we have erected, but this is an adversarial space and we are getting better with our technical tools all the time.

Q175 **Clive Efford:** Are there any circumstances where Facebook itself makes any money out of these fake accounts?

**Monika Bickert:** People do not pay to have an account on Facebook. Of course, fake accounts that are sitting on Facebook may see ad impressions, but our long-term interest and our overall business interest is in making sure we have a community that works where people are not getting scammed, where they are not coming into contact with fake accounts. People do not want to interact with fake accounts, so it is very much in our interest to remove those from our site, and we work hard to do that.

Q176 **Clive Efford:** In relation to non-paid-for advertising accounts on Facebook, you prioritise paid-for accounts over non-paid-for accounts. For instance, the World Health Organisation, which at this moment in time would be very important in terms of circulating scientifically backed information about the current Covid crisis, would not get priority on



## HOUSE OF COMMONS

Facebook. Is that something that you think should be looked at?

**Monika Bickert:** I do not know what you mean by “paid accounts on Facebook.” Accounts on Facebook are free.

Q177 **Clive Efford:** It seems that Facebook deliberately reduces the reach of non-paid-for content on Facebook.

**Monika Bickert:** Accounts on Facebook are free. You can pay if you want to run an ad. We are giving free ad credits to the World Health Organisation and other health authorities, including health authorities within the UK Government, and we are promoting content from those organisations. Just to give you some numbers on that, the UK NHS and gov.uk sites, we are directing people to those through our coronavirus information centre, which we are putting at the top of people’s news feeds. We are returning it in search results if they are looking for information. We are doing pop-ups, and any of that leads them to information from the WHO but also NHS and gov.uk. We have seen more than 3.5 million visits to the Covid resources on NHS and UK Government websites from Facebook and Instagram since January as a result of us directing to those resources.

With regard to the WHO and the UK Government, we are giving them ad credits to allow them to reach more, so the UK Government’s free advertising has allowed them to reach more than 40 million people with messages about Covid during this time.

**Clive Efford:** Thank you, Chair. I will let others come in.

Q178 **Damian Hinds:** Ms Bickert, are you still planning to press ahead with your plans for encryption in the next year?

**Monika Bickert:** We are still planning to implement encryption, but we are still in the investigative stages at this point. We are talking to Government authorities and to security and privacy experts, and we are trying to learn the best way to do this. I am former law enforcement myself. I was a criminal prosecutor for more than a decade, so I know that the conversations we are having with those constituents are very important, and we are trying to understand that feedback and make sure that we can bring end-to-end encryption to people to give them the privacy and security benefits, while also being mindful of the challenges.

Q179 **Damian Hinds:** Speaking of those challenges, could you spell out for us in black and white what are the risks of end-to-end encryption for your or any third party’s disinformation detection methods and, indeed, for detecting instances of child abuse?

**Monika Bickert:** With end-to-end encryption you do not have access to the content. Some of the enforcement that we do right now on, say, Facebook and Instagram involves us looking at the content and making an assessment. Some of the enforcement that we do focuses on behaviour, and that is the sort of enforcement that you can still do in an encrypted space. For instance, you can look at things like mass creation



## HOUSE OF COMMONS

of accounts, mass messaging. I spoke earlier about seeing fake accounts being created, and often they are spammers trying to get messages out there. On WhatsApp—that is an encrypted messaging service—we have developed the capability to identify when there is this sort of mass messaging or mass account creation going on, so that we are now removing millions of fake accounts and—

**Q180 Damian Hinds:** Sorry to interrupt, Ms Bickert, but to be clear, those things you could do now, I think. The introduction of end-to-end encryption only decreases the amount of detection you can do. It does not create those new methods. Is that right?

**Monika Bickert:** You are right that we will not be able to see content. I do think as we are working—

**Q181 Damian Hinds:** There may be other methods that you have as well, but it removes one important strand of detection.

**Monika Bickert:** It removes our ability to see content, but I want to make the point that we have an integrity team that is always focusing its attention on the adversarial space and how to get better at identifying abuse. Right now they are looking at ways they can get more aggressive and learn more about how to do things in the behavioural space.

**Q182 Damian Hinds:** Back to today, in correspondence with your colleague after our previous hearing, he wrote of your content warning labels, “100% of people who see content that has already been flagged as false by our fact checkers will essentially be told that context.” What is the time lag usually between a falsehood being originated and it being factchecked and those content warning labels becoming available?

**Monika Bickert:** It depends. We have situations where content is factchecked extremely quickly. It is something that goes very viral in the press, and it will be factchecked almost immediately. We have other stories that get sent to the fact checkers that the fact checkers never end up factchecking. We do—

**Q183 Damian Hinds:** To be clear, many thousands or, in some cases, millions of people could have seen these falsehoods before they are factchecked. A question that we put to your colleague previously was, therefore, why don’t you introduce or reintroduce the truth to everybody who has seen those falsehoods or at the very least—considering you can and do measure the linger time, the dwell time that users have on individual pieces of content—make sure that the record is corrected with those people who have obviously spent some time consuming those untruths?

**Monika Bickert:** There are some situations where we do. I would say our goal is to get people accurate information and make sure they are not being misled by misinformation. Let’s say you are using social media and something is in your news feed. It is false, but you do not necessarily see it or pay attention to it and then you receive a notification. Right now, when people see something false around Covid, which I can get into, we



## HOUSE OF COMMONS

send them a notification saying, "So and so, you saw information that has been marked false." The research is not clear yet on what the effect is of showing somebody that. In other words, they might not have seen or paid attention to the false information in the first place.

**Q184 Damian Hinds:** That is why I specifically asked about dwell time. You know how long people have spent with individual posts through the algorithm so, in theory, it is possible to distinguish those who have seen it but clearly did not read it and those who probably did read it and consume it.

Can I ask you finally about one separate but related thing? I think you have been trialling, on Instagram in particular, the downgrading of metrics like likes in a number of countries for quite some time now, and also trialling the reintroduction of friction, so asking, "Are you sure you want to post this? Are you sure you want to forward this?" What are the results of those tests so far?

**Monika Bickert:** On the likes test, I do not have any information on that. I would have to look at that. I will say on the friction it tends to be very powerful. Right now, if content is marked false by our fact checkers, we put a label over it saying, "This has been ranked false. You can click through to see it, and you can also go here for more information" and in more than 95% of the cases when people see those labels they do not click through it, so that friction is very powerful.

On your point about the dwell time and notifications, this is an area where we are still learning. We are doing qualitative and quantitative research to make sure that we make a decision that is based on the data, but I will say that, with information that has been factchecked after somebody shared it, we do go back and tell the person who shared it. I am not sure that is 100%, but we do go back and tell people that they shared something false.

We are also learning from when we remove content for being false. Now our approach to misinformation has our third-party factchecking programme where we mark content with factchecks, and then we have a programme under which we remove certain types of misinformation. Specifically I am talking about misinformation that is likely to contribute to a risk of imminent physical harm. That applies with Covid. I can talk through the categories of that, and then a couple of other areas. With the physical harm misinformation around Covid, if we remove something and somebody in the past liked it, commented on it, interacted with it or shared it, we are sending them a notification. We have removed the content so we do not surface the content to them again, but we send them a notification saying, "You saw information that was false, click here to get the real facts" and we are seeing those efforts driving people to reliable health authorities.

**Q185 Damian Hinds:** Could you please get back to us in follow-up correspondence on those trials of de-metrification, and particularly tell



## HOUSE OF COMMONS

us, not only in the context of misinformation and disinformation, what you are learning about the effect on young people's self-esteem and wellbeing?

**Monika Bickert:** Yes.

Q186 **John Nicolson:** Good afternoon, Ms Bickert. Thank you for your time, but we are short of time, so may I request brief answers to a couple of quick questions, please? First, is it acceptable under your rules to post the following: "Bomb the Board of Deputies of British Jews"?

**Monika Bickert:** Hypotheticals are going to be tricky for me because we would always look at the context, so just like my colleagues from YouTube and Google said, if somebody said, "This was said today. Isn't it horrible?" then that is going to change everything—

Q187 **John Nicolson:** No, I can help. It was a standalone. It was somebody advocating doing that.

**Monika Bickert:** No, that is not okay. I know—

Q188 **John Nicolson:** When the Board of Deputies of British Jews reported it to you, this is how you replied, "We reviewed the comment and found it does not go against any of our community standards. We recommend you unfriend the person who posted it." It seems unlikely that the Board of Deputies of British Jews will have a friend who advocates bombing them, so I wonder why you would send such a grotesquely dismissive response.

**Monika Bickert:** That was an error, and we did remove that content. We never should have sent that message.

Q189 **John Nicolson:** Let me try another one, then. A photo of George Floyd, the African American whose murder has provoked such an outpouring, was posted on Facebook. His nose—he is an African American—was circled with a caption saying, "Are you trying to tell me this guy couldn't breathe?" Is that acceptable or not?

**Monika Bickert:** Again, hypotheticals are difficult for me because I do not see the content or the context, but if somebody is mocking the death of George Floyd or anybody else—

**John Nicolson:** They were. They were.

**Monika Bickert:** Then it would be removed by the clear policies.

Q190 **John Nicolson:** This is what you responded to that hate-filled, racist post, "It doesn't go against any of our standards, including hate speech." It is a grotesque response.

**Monika Bickert:** Any time that we make a mistake in our content review and our messaging does not reflect the decision that we should be making, I will be the first to say that our enforcement is not perfect. We do—



## HOUSE OF COMMONS

Q191 **John Nicolson:** It is far from perfect. Your senior staff are up in arms. You have had staff stage a virtual walkout. I am astonished by the way you said you had not read the former senior staff's letter to *The New York Times*. I recommend *The New York Times*. It is a very good paper, and it is the newspaper of record for most Americans, so well worth getting out the cuttings file.

The problem is that Mark Zuckerberg sets the tone, doesn't he? He refuses to take down Trump's racist and violent messages encouraging the police to shoot protestors. My suspicion is that you will never act seriously on hate speech, such as the grotesque hate speech examples that I outlined just now, and the awful, awful dismissive responses that you sent. You will not take it seriously until we, as parliamentarians, hit you with serious financial penalties. It worked in Germany, and I think we need to do it here.

**Monika Bickert:** There are a couple of things there that I would love to reassure you on. First of all, with hate speech, we definitely have become better at our enforcement. I will grant you that we still make mistakes but we are working with partners in the UK, a number of whom have been our partners for years, on better tracking and understanding where we are falling short and getting better. We are now at a place where 88%—I think that is the number—of the content that we remove for hate speech we are finding ourselves with our own tools, which means fewer people see it. We are not waiting for somebody to report it.

On the topic of regulation, we have for quite some time said that we do think that regulation can help advance our shared goals of keeping people safer. We have worked in an open dialogue with the UK on this, and we think the UK has a real opportunity to be a leader in online content regulation and we look forward to continuing that.

**John Nicolson:** It is worth noting finally in closing that, when the Germans hit you with huge financial penalties, you hired a substantial number of content moderators and the amount of abusive, racist and anti-Semitic postings dropped substantially because you did not want to be fined. Whereas here and the United States, where you are not fined, on grotesque posts like that about George Floyd, mocking his death, mocking his African American identity, you write back replies saying it does not go against your community standards. It is disgraceful and there is widely shared contempt, I think, of the way you respond to this kind of abuse.

**Yvette Cooper:** Ms Bickert, can I just go back to the issue about end-to-end encryption? The Internet Watch Foundation told the Home Affairs Committee yesterday, "We are very, very concerned about the intention to encrypt Messenger and the impact that will have on victims of online child sexual abuse. We are asking Facebook to give assurances that child protection will not be hampered and that children and victims will be protected in some way. As yet, none of us has seen any assurances across the whole child protection sphere. We are all very, very concerned



## HOUSE OF COMMONS

about the encryption of Messenger.” What assurances can you give them?

**Monika Bickert:** Thank you for flagging that. This is something that we are taking really seriously. I mentioned earlier that my career was in law enforcement for more than a decade. Specifically what I worked on for a number of those years was crimes against children, and the child safety work at Facebook rolls up under me and Antigone Davis. She has been leading a lot of our efforts to understand what we can do. First of all, what are the concerns from the law enforcement and child protection community, and how can we work with our product design and engineering teams to be responsible and make sure that we are still keeping children safe? I do not have the answers for you on this yet. I can tell you that we are continuing to engage and understand the concerns, and we are committed to making this a safe experience.

**Yvette Cooper:** If people want it end-to-end encrypted, if an offender uses it to send an image of a child being sexually abused or being raped, what will you be able to do about it?

**Monika Bickert:** I do not have the answers for you on that yet. Like I said, we are still in the consultative phase. These are early days for us, but we can absolutely keep you in the loop as we are developing our plans.

**Yvette Cooper:** Does that mean you are not definitely going to go ahead with end-to-end encryption?

**Monika Bickert:** Our plans are to go ahead with end-to-end encryption, but we are still developing how we roll that out, how we implement it and how it looks.

**Yvette Cooper:** You are going to go ahead with end-to-end encryption, even though you do not have a way to protect children from abuse and from sexual abuse using that end-to-end encrypted platform? You are definitely going to go ahead with it?

**Monika Bickert:** We are committed to finding a way to make sure that, however we go ahead with our encryption, it is a safe platform for children and for everybody else.

**Yvette Cooper:** If nobody can see the content, and this content could be an image of a raped child, how can it possibly be safe?

**Monika Bickert:** It is true that we will not be able to see the content unless it is reported to us. Identifying content that has not been reported to us is, of course, very important, and we are looking at how we can build this product with those concerns in mind.

**Yvette Cooper:** Could you use AI, for example, on encrypted content to identify images that the Internet Watch Foundation has already identified and flagged?



## HOUSE OF COMMONS

**Monika Bickert:** I cannot give you answers on how we are likely to build this product. I can just tell you that we are working hard to understand what the options are, and we are in consultation with law enforcement and other safety groups to understand the best options for keeping people safe.

**Yvette Cooper:** They do not think they are in consultation with you. The Internet Watch Foundation and the British National Crime Agency told us yesterday quite how worried they are and how difficult they are finding it to get any communication with senior Facebook executives on this. They are not convinced that you are taking this seriously at all.

**Monika Bickert:** We are, in fact, in regular communication with law enforcement and safety groups, including the National Crime Agency, with whom we have met a number of times in the past. We have UK law enforcement outreach personnel who have been meeting with the UK Government, including the National Crime Agency, but we will make sure that we keep those lines of communication open and that they know how to reach us.

**Yvette Cooper:** It would be very helpful if you could keep us in touch with this, because I do not know if you recognise quite how serious it sounds, that you have made a decision to go ahead with something and you do not seem to have any idea how you are going to solve this massive problem about how to protect children.

**Monika Bickert:** These concerns are very, very serious. We are devoting a lot of time to this engagement and to explaining product options. Thank you.

Q192 **Chair:** Finally, there is a pattern here with Facebook in terms of your engagement with outside agencies. Yvette was talking about the National Crime Agency, the Internet Watch Foundation and so on. You do not share your data with truly independent academics and organisations working to assess and improve the reliability of information disseminated online, often at little or no cost to you. Given the widespread concerns, not just today but over the years now—there is a body of evidence—why is it that you are not willing to be more open with third parties?

**Monika Bickert:** We do work quite closely with independent researchers and we do work—

Q193 **Chair:** Sorry to cut across you there, Monika. You work with only a very narrow framework of independent researchers. You do not, for example, work with many very highly accredited independent agencies. The question is why are you so nervy about third parties?

**Monika Bickert:** We do work with external parties. We have to be very mindful of privacy concerns and provide data to people only in circumstances where we are sure it complies with our terms and with users' expectations. We work with researchers in a number of areas, including when we produce data. I mentioned the six-month reports we



put out. We even work with researchers to make sure that the way we are doing our research and the way we are producing our data passes muster, so we absolutely work with external researchers and take their feedback seriously.

Q194 **Chair:** Other social media companies are more open in this respect. What is Facebook's problem with outside scrutiny?

**Monika Bickert:** On the contrary, I would say we welcome it. We make resources available to researchers in a number of different areas. We have researchers look at our data practices. We publish a comprehensive transparency report where we give information about Government requests and how we have responded. We publish a community standards enforcement report where we lay out not just how much we have removed but what the prevalence is of content on our platform, so we do a statistically significant review of what content there is on the site that we are missing. We publish those numbers, and we also publish how much of our efforts at removing content are driven by proactive detection, meaning how much of the stuff we got to before anybody reported it to us. Therefore we are very public. We welcome those conversations and, indeed, get a lot out of them.

**Chair:** That is not the perception of almost the entire academic community, but thank you for your evidence today, Monika Bickert. It is very much appreciated that you were able to join us. Thank you.

## Examination of witness

Witness: Nick Pickles.

Q195 **Chair:** I am now going to call our third witness, Mr Nick Pickles, the Director of Public Policy Strategy at Twitter. Good afternoon, Mr Pickles. We are sorry to keep you waiting.

**Nick Pickles:** Morning, Chair. How are you doing?

**Chair:** Very well, thank you. It is not quite as bright here as it is there. Now that you are imposing editorial scrutiny on President Trump's tweets, are you a platform or a publisher, or do you recognise what many people believe you are, which is a hybrid of both?

**Nick Pickles:** You are absolutely right to highlight that the existing framing of businesses does not work anymore. What we did last week and what we have done many times is to seek to provide our users with more context. We do that by providing them with tweets that were posted to Twitter by other agencies. They might be journalists, experts, academics, third parties. When we provide this context it is not Twitter speaking, it is us curating and sharing tweets shared on our platform by other people.

Q196 **Chair:** You still have not answered the question. Does it mean that you recognise that you have much more skin in the game, you are more than



## HOUSE OF COMMONS

just a post-it, you are more than just a platform and you have real editorial responsibility?

**Nick Pickles:** Sorry, just to be clear, I totally agree with the premise. I do not think the traditional dichotomy of platforms and publishers works anymore. You are right. There is a middle ground that protects the important business functions that companies like ours have—content moderation, providing users with context—that does not work in the same way as a newspaper would do. I think that hybrid is the correct way for regulatory discussion to go.

Q197 **Chair:** In relation to President Trump’s executive order, will this have an effect on political speech in this country or elsewhere?

**Nick Pickles:** You rightly highlight the importance of his executive order and the regulation it proposes. It would fundamentally limit the ability of platforms to keep their users safe. Something that is not often discussed is that the law being discussed—section 230 of the Communications Decency Act—provides protection for companies to do content moderation. It is sometimes interpreted as other things, but that is a critical function. If you undermine the ability of companies to moderate content free from liability, it restricts our ability to do our jobs and, worse, it probably acts as a barrier to entry for smaller companies who do not have the same kind of resources to devote to these issues as some of the larger players. Twitter in this space is definitely not the biggest and definitely not the smallest. We are somewhere in the middle.

Q198 **Chair:** Yes. I think only President Trump could conclude that we need less oversight of social media. Has there been any discussion at all within your organisation of suspending President Trump’s Twitter account?

**Nick Pickles:** Whenever a tweet by any user is posted and reported to us, we consider it under our rules. As you will have seen last week, we did conclude that one of President Trump’s tweets broke our rules. We also concluded that we should use a policy that we announced last year, which said that in certain circumstances where public figures who are verified on the platform post something that breaks our rules, but we think the public debate about that tweet is important to protect, we would place a notice on the tweet saying it broke our rules and that we would allow the tweet to remain up, but you cannot retweet it and you cannot reply to it. If any user on Twitter continues to break our rules, we will continue to have discussions about any and all avenues open to us.

Q199 **Chair:** You do not rule out suspension of President Trump’s Twitter account?

**Nick Pickles:** Every Twitter account is subject to the Twitter rules.

Q200 **Chair:** That is a yes then?

**Nick Pickles:** As I say, every Twitter account is subject to the Twitter rules.



**Q201 Chair:** On the issue of anonymity, at present you have a huge swathe of harmful content pumped out by anonymous users, either individuals or foreign actors. You take down one piece of content and another one appears. It is a bit like—as one academic describes it—a game of Whack-a-Mole. Surely you must now recognise that the time has come for you to make some form of identity checks before allowing people to post, or is this a cynical decision, that anonymous accounts drive traffic and notoriety and that is simply good for business?

**Nick Pickles:** You raise an excellent point. One of the things that Twitter does have is the verified badge. We did open that up so that anybody could apply to be verified. One of the things we saw happen then was that some people felt that the badge said, “You are who you say you are” and other people felt it was Twitter endorsing a user. With that problem of different people thinking it means different things, we stopped verifying members of the public openly. We continue to do so in critical cases, but the question of whether we can give our users ways to give more context to demonstrate their identity in different ways is something we are investing in, and the long-term ability of people to share information about themselves to give you context of who they are is definitely something we are looking at.

We have already launched labels for candidates in the US election, and this idea of adding labels to accounts that might, for example, say whether you are affiliated with a state media agency, or even whether you are affiliated with a state, can add context to users without being the same as us checking their passport. I think that is an exciting area of opportunity for us to add context for our users.

**Q202 Chair:** Yes, but does ending anonymity damage your business model?

**Nick Pickles:** More importantly than that, anonymity and pseudonymity are incredibly important for the ability of people to speak out in times of crisis where their safety may be at risk. If you look at what is happening in Hong Kong, the United States and many areas around the world, the ability to use Twitter anonymously is one of the reasons why Twitter is used to bring first-hand accounts of what is happening in these places to the world. We think that is something that is not a business decision. That is a principled decision based on protecting the safety and human rights of those voices.

**Q203 Damian Green:** Returning to President Trump, he has been tweeting in his particular style for years now, so what changed last week that meant you applied a corrective to that particular tweet?

**Nick Pickles:** The corrective we applied, which was a link to a curated set of tweets from people, was under a policy that we introduced on 11 May this year, so it is very new. Some of the previous discussions you had with Mr Knowles and Dr Smith talked about this idea of sometimes you do not want to remove the content, but you do want to add context to help people understand and avoid being confused. The intervention



that we deployed last week is very new, and we have now deployed it thousands of times around the world, particularly focusing on Covid and making sure that people can get up-to-date, accurate information about Covid-19.

**Q204 Damian Green:** One of the criticisms has been that this new rule already appears to be being applied inconsistently, that you have not applied it to things he said about hydroxychloroquine. Are you now committing yourself as an organisation, every time a public figure says something that is easily challengeable, that you are going to put up one of these blue flags that say, "Go and look at the facts somewhere else as well"?

**Nick Pickles:** No. The discussion about medical information, as the previous sessions highlighted, is very complicated, particularly where there are pre-publication, non-peer reviewed papers being discussed, which different people may take different readings of. We have said publicly there are three policies this label will be applied to: our civic integrity policy—which particularly relates to the ability of people to vote, and in the particular case last week the risk of confusion about ballots and mail-in ballots was the reason why that tweet was labelled—Covid-19 and synthetic and manipulated media, so we are starting small.

We have also said, as I think previous witnesses from the health service called for, that we are focusing on the most viral and the most visible content first. That is why you may see these interventions more regularly on higher-profile accounts because they have bigger audiences.

**Q205 Damian Green:** How many people do you have working on this now? This is a whole new area that Twitter has never had before. How seriously are you taking this?

**Nick Pickles:** The actual product that we use, which we call Moments, is something we have had for many years. It is staffed by people with a range of expertise. Some of them are former journalists; some come from academia. We have had that product for several years. The change is taking those Moments and attaching it to a tweet to make it easier for people to find. One of the things we have seen during Covid is that this is an effective way of giving people extra information and extra context about things that might be disputed or unclear, without Twitter being the arbiter of truth. That line, where we can provide context to our users without telling them something is true or false, is good for our users and it is the right approach in terms of Twitter not wanting to regulate things such as political debates.

**Q206 Damian Green:** You made the point that you are now something of a hybrid, that you are not just a channel for other people's views, but you are not quite a newspaper, as it were, in your desire to avoid that kind of regulation. I genuinely cannot see how you can avoid or argue that you should not be regulated. Effectively it is a warning sign, isn't it, that, "This person has said something, and frankly we think you should go away and look at some facts"? The clear indication is that this is, at best,



unreliable, and at worst is dangerous and false. Once you start making that decision, you are behaving very close to the position of a media or newspaper editor and as such, around the world, you will end up under the same regulations as newspapers and old-fashioned broadcasters.

**Nick Pickles:** It is an excellent point. One of the benefits of platforms like is that they give many people a voice in a way that traditional media have not. For example, this week with the events in the United States and the tragic killing of George Floyd, we saw 10 million tweets using the hashtag #blacklivesmatter in one day. That was the most that movement has ever been talked about. That cannot happen in the kind of traditional media world. Indeed, one of the reasons that social media has been so critical to bringing attention to what is happening around police violence is because it was not being discussed by media organisations. This idea of how you protect the most vulnerable voices, which did not have a right for a voice and were not being heard before social media. How do you protect those voices?

At the same time, and I think this is the long-term challenge, we are moving beyond a world where the only decision available to companies is to leave something up or take it down. The world is far more complex than that dilemma of doing nothing or removing, and our approach is about making interventions that give people more context and allow them to make more decisions themselves, but also allow the public debate to happen around things such as, for example, when politicians make statements that some people may take one view or another about.

Q207 **Damian Green:** One final thought is that you could have a sort of correct the record tool, like Facebook's. Are you doing that? That would be another tool in that particular armoury.

**Nick Pickles:** The correct the record idea discussed by Mr Knowles and Dr Smith is interesting. I would caution, in recognising the conversation we had about peer review research, that there is a risk that interventions in this space make the problem worse. There was a paper in *Science* earlier this year looking at something similar in Brazil, using the World Health Organisation, and it did not work. There is a real risk that, if you intervene in some of the ways being proposed, you could make the problem worse. Particularly if people see information from an organisation that they do not like, that they may politically not trust, you can trigger a reaction that is not productive.

One of things that I think is very important is that a number of studies around correct the record are not peer reviewed. More of this research needs to be peer reviewed in journals, and the academic community needs to weigh in because right now there is a concern—I think you heard this from previous witnesses—that just showing someone information that they have previously missed could make the problem worse. If you get the source wrong, it could trigger a motivated reasoning reaction, which again might make someone more firmly believe the wrong information they saw. The science about the best way to do



## HOUSE OF COMMONS

this is still unclear, but we are closely following the science and are keen to learn from it about what good interventions could look like.

Q208 **Philip Davies:** Can I come back to the point about anonymous profiles, which the Chairman mentioned in his line of questioning? Is there more likely to be inaccurate information given on an anonymous account than on a named account?

**Nick Pickles:** I have not seen any studies that look at that one way or the other.

Q209 **Philip Davies:** You are not aware of any evidence at all? Is this not something that you, as a company, have looked into? Have you not commissioned any research yourselves to find out whether or not there is much more likely to be factually inaccurate information on an anonymised account than on a named-person account?

**Nick Pickles:** As you heard from the earlier witnesses, one of the challenges we have is that it is not just about accounts posting; it is how many people see. Far more people see those high-profile, celebrity accounts, so in some cases it is when more people see it, rather than an account that—

Q210 **Philip Davies:** If you could have a crack at answering the question, have you not made any attempt, commissioned any research or done anything to find out whether inaccurate information is more likely to be found on an anonymised account than on a named-person account? It is not a difficult question. Have you or have you not? I take it from the waffle that you have not, but perhaps you might just clarify.

**Nick Pickles:** No, we have not commissioned that research, and I have not seen any research from external academics on that topic either.

Q211 **Philip Davies:** Why not?

**Nick Pickles:** I do not know. I think it is a fair question to ask the research community.

Q212 **Philip Davies:** I am asking you. I am not laying the blame for this on the research community. I am asking you. It is your company. Are you telling me that, for all the brains at Twitter, nobody has thought, "Do you know what, there is all this controversy about anonymised accounts. Why don't we do some research to see if information is much more likely to be factually inaccurate on an anonymised account than on a named person account?" Has it never occurred to anybody in Twitter to do that, or do you not care?

**Nick Pickles:** We are focusing our research on questions such as what is the best way to intervene to give people more context. If Twitter and the academic community are not prioritising this research, one reason may well be because the problem is not as you have envisaged it. The problem that we are looking at is different. I am genuinely open and, if there is research out there that says this is an issue, I would love to read



## HOUSE OF COMMONS

it, but I think one of the problems is that people have focused on anonymous accounts as a disproportionate part of the problem than is actually the case.

**Q213 Philip Davies:** So you have no intention of doing this? You are not interested in doing it, and it does not bother you whether that would be the case or not?

**Nick Pickles:** We will continue to look at the problem. If someone shows evidence or we see evidence that this is a problem, we will direct research resources there. But one of the challenges here is that there is no evidence that this is a problem, and it is very difficult for us to preemptively say we will study everything where there is no evidence of a problem.

**Q214 Philip Davies:** Is there any evidence that offensive tweets are more likely to take place from anonymous accounts than named-people accounts?

**Nick Pickles:** Again, I am not aware that is true. One of the experiments we are currently running goes to this very question of whether we can provide our users with an opportunity to think again before they post tweets that perhaps might be reported to us. I am happy to ask the research team if the status of an account being verified or not is captured in that experiment. But again, I have seen high-profile examples of abuse and I have seen examples of abuse from accounts where people have hidden their identity, but I have not seen a conclusive study saying one way or the other. Indeed, I know the South Korean Government tried this and did not find a connection between anonymity and abuse.

**Q215 Philip Davies:** If I ask you to go away and commission some research from a reputable firm to see whether inaccurate information is more likely to happen from an anonymous account than from a named-person account, will you say yes to that or will you say no?

**Nick Pickles:** I will say I will follow it up, because the complexity of asking that question is, I think, a lot greater than just saying there are two very clear buckets, accounts with real names and accounts without. Twitter allows parody accounts, for example, which might use a real name but also not be the person. I am happy to follow up on this. If you are aware of studies that talk about this, I would be very keen to read them.

**Q216 Steve Brine:** If I were to copy a tweet that President Trump put out that you censured and tweeted it, would the exact same message go up on my tweet? Not if I re-tweeted it, because clearly if I re-tweeted it I would re-tweet it in the way that you had censured, but if I copied it and re-tweeted it, would it happen to me?

**Nick Pickles:** The process is that we first consider whether the tweet breaks our rules. We said this tweet did.



**Steve Brine:** It would be the same words. I am literally copying it.

**Nick Pickles:** The second stage in the process is we would ask whether there is public interest in people being able to discuss Steve Brine MP's view of what is happening in the United States. In the decision about whether that public interest is the same as it is in the view of the President of the United States, the public interest discussion is different. You, as someone who is not in the United States and who does not have command authority over the US military, have a different context. It is possible that we might remove it without providing a label.

Q217 **Steve Brine:** When the Chair asked you whether it is possible that you could remove President Trump's Twitter account, you did not say no. You said, "Everyone has to adhere to our policy." What would have to happen for President Trump's Twitter account to be removed from your platform? Is there a three strikes and you're out policy as part of the Moments policy that you have just introduced?

**Nick Pickles:** Every one of our policies is based on proportionality. In some cases, if the harm is severe, we would remove immediately. In some cases, if the harm is less severe, we might, for example, apply the label we did last week or require the user to take a timeout. It is entirely dependent on different policies, the proportionality of our response, and every Twitter account is subject to the Twitter rules. We demonstrated clearly last week that that does include the President of the United States, and we will continue to enforce our rules.

Q218 **Steve Brine:** There is nothing against him if he keeps being a repeat offender on posting stuff that breaches your rules? The Moments team would look at it and put up the censure message, as they did on the tweet about registering to vote? Let's assume that there was a presidential election coming up. I guess it is possible to say that, over the coming months, the heat in US politics is probably not going down, so the chances are that this President, knowing how he campaigns—it was successful for him last time, so no reason to suggest that he will not do it again—will probably not dial it down in the next few months, will he?

**Nick Pickles:** I think you are absolutely right that we are seeing a period in the United States and, indeed, in other countries where the political dialogue is heightened. I would say that during election campaigns people also recognise the value of their Twitter account, so they are likely to act in a way that is within our rules because of the value of using Twitter during a campaign to talk to the public. I think there is a balancing effect there.

Q219 **Steve Brine:** That is the point, Nick, isn't it? What you are saying is that President Trump will be so chastened by events on Twitter that he will say, "Okay, hands up, I have screwed up here" and that he will not post his inflammatory material during the campaign. Have you considered the impact that Twitter could have on the election to the most powerful office on the planet over the next few months?



**Nick Pickles:** Certainly the context of the user is something we take into account. That is why last week we applied this label and did not remove the tweets. This is an ongoing discussion, and we will continue to apply our rules, but I will not get into the intentions and mindset of any one user. All our users are subject to our rules, and we will continue to enforce them throughout the election.

Q220 **Steve Brine:** Finally, let me ask you a general point about the teams. You said that the new censure policy, the Moments policy, has been used. How many times did you say it has been used since it was created last month?

**Nick Pickles:** Several thousand times now.

Q221 **Steve Brine:** Have any Democrat politicians in the US, or politicians on the left elsewhere, been censured?

**Nick Pickles:** I want to distinguish. Two different things happened last week. One was the label that we placed on the tweets about civil unrest, which was a warning saying they broke the rules. That was different from the opportunity for users to find more context about a topic, which is not saying something broke our rules, but was us saying it has the potential to cause confusion about voting. I do not have a party-by-party breakdown of that data, but I would certainly say if we did see something—and I can certainly say that in the mid-term elections, for example, one of the things we saw is people saying, “If you are a Democrat, vote Friday; if you are a Republican, vote Thursday” or vice versa—we would take action on either version of those events, because they interfere with people’s ability to vote.

Q222 **Steve Brine:** President Trump was obviously making some threats about what would happen as a result of this, because he is not happy about it. If you were suddenly employed by the President of the United States instead of Twitter and you wanted to do anything to Twitter, what would you do to you? Is there anything? My question is: is there anything that he can do to you?

**Nick Pickles:** There are a number of leaps I would have to make to envisage that situation, which might take a moment. My role right now is speaking on behalf of Twitter. All our users are subject to our rules. We will continue to enforce those rules. I am not aware of an offer of employment as you describe being made.

Q223 **Steve Brine:** Could he do anything to you? Is there anything that he could do to you?

**Nick Pickles:** You have obviously seen that the White House has published an executive order. We have said very clearly that we do not think that is the way forward for the open global internet, for speech in America or for smart regulation. We have said it is a politically motivated and reactionary move. I note a court challenge has now been placed by



## HOUSE OF COMMONS

the Centre for Democracy and Technology against that executive order, and we will be following that closely.

Q224 **John Nicolson:** Let's pursue that for a moment or two. Are you aware of Bizarre Lazar's @SuspendThePres Twitter account?

**Nick Pickles:** I think I saw this two days ago, yes.

Q225 **John Nicolson:** He tweets exactly word for word what Trump tweets, immediately after Trump. This is not abstract, because what you did was suspend his account for violating your standards, yet his White House doppelganger rants away with racist and violent bile, almost unchecked. Why is there one rule for Trump using Trump's words and a different rule for Lazar using Trump's words?

**Nick Pickles:** I think this is the system working as intended. In both cases we said the tweet broke our rules. This is a public policy announced last year. We said if an account breaks our rules but meets the criteria of being verified, having more than 100,000 followers and being operated by a public figure, we may take the option that, in the public interest, we want that tweet to be available. One of those accounts meets those criteria; one of them does not. I would also note that the account was not suspended. We asked them to remove that individual tweet, which has the effect of locking the account until they do. If the user removes that tweet, they can come back on Twitter. This is the system working equally. Both tweets broke the rules; both tweets were actioned. One from a public figure was maintained to allow debate.

Q226 **John Nicolson:** That is an interesting spin on it. You could get around this, of course, by insisting that folk do not have anonymous accounts. That way it would be a level playing field. If you are vile and racist and the President of the United States, you can continue because you have a blue tick, but if you are an anonymous troll, your account is cancelled or suspended.

You said that you did not know of any research that linked particular levels of abuse with anonymity, so let me help. The Clean up the Internet campaign, which is UK based, conducted an in-depth study on this and found a clear link between anonymity and the spread of conspiracy theories on Twitter, and also online abuse targeting women, black people and other minorities, including trans people and gay people. Can I explore for a moment or two the language that is acceptable on Twitter under your rules? Is it acceptable under Twitter rules to call a gay man a greasy bender, a paedophile or a rape neighbour?

**Nick Pickles:** Without seeing specific tweets, it is difficult to comment, but if you have specific tweets using that language, I am happy to make sure they are reviewed under our rules. Those things do sound like something we would want to take a close look at.

Q227 **John Nicolson:** In what circumstances would it be acceptable to call a gay man a greasy bender, unless of course you were saying, "This is a



## HOUSE OF COMMONS

revolting tweet. How vile it is that someone said that?" But if you, in an unsolicited way, are aiming that at a gay man from an account or calling him a paedophile or a rape neighbour, how could that ever be justified under Twitter's rules?

**Nick Pickles:** I think you just highlighted that in the situation of someone condemning it, but it does sound like this is something that we should take a close look at. Without seeing the specific tweet, it sounds like something that we would want to look at as a potential violation of our rules. I am certainly happy to follow up with colleagues and to make sure that happens.

Q228 **John Nicolson:** I am a gay man. I have been called all three in the last month in a pile-on campaign of abuse. Twitter in the UK had all these tweets. They sent me a link explaining how I could mute the abuse from one particular account. What else can I do?

**Nick Pickles:** I am very sorry to hear that. I have not seen the specific tweets. I will follow up with my colleagues in the UK and follow up with you afterwards to make sure that has been thoroughly reviewed and to make sure that every opportunity to review them has been taken.

Q229 **John Nicolson:** I understand how to tweet and I understand how to block somebody, but that misses the point entirely. Accounts should not, as I understand it—I have read your Twitter rules on negation very closely—target abuse, and if they do, they should not be allowed on your platform. I raise this not just because it is me, but because I know lots of colleagues, particularly BAME colleagues, are regularly targeted with vile abuse on Twitter, which they find enormously distressing and that affects their mental health and their ability to work. There is a general consensus that when issues are raised with Twitter—and in a very similar way with Facebook—they tend to get back platitudes or, in my case, an enormously patronising response from one of your senior company executives, who appeared last time before this Committee, sending me a little link explaining to me how I can mute nastiness. It totally misses the point, doesn't it?

**Nick Pickles:** You have my commitment. I will personally follow up with you and your office. We will make sure these tweets are reviewed. We also keep our rules under review, so if there are things that are not being caught by our rules right now that we think should be, we are open to changing our rules, if needed. I think this is a case worthy of much more investigation.

**John Nicolson:** Thank you, because I have read your rules very carefully, and I think there is no doubt that these tweets breach your rules. I cannot see any defence under your rules. I think it is a question of enforcement.

Q230 **Chair:** Finally on disinformation, Nick. The Oxford Internet Institute tells me that Chinese media are reaching something like 1 billion social media accounts across all platforms each week in English—in English, crucially.



## HOUSE OF COMMONS

Are you comfortable with that? If not, what do you propose to do about it?

**Nick Pickles:** Thank you for raising this very important question. It was something I spoke about a few weeks ago. One of the effects of companies being better at catching information operations run through fake accounts is that states are now moving their activities into things like diplomatic networks and state media. That is a result of pressure we have applied. You may have seen that we have published the archives of information operations directed by states, including China, Russia and Iran.

One thing we did very clearly last year was an important step: we banned all state media from advertising. If you are a state-controlled media entity, something like Russia Today, you are not allowed to advertise on Twitter. You can have an account, you can speak to our users within our rules, but you cannot buy advertising. One of the themes that has come through today's hearing is the question of monetisation, and I think this question of how disinformation is monetised and how platforms play a role in that is one that has been significantly under-investigated. We raised this in our submission to the Australian Parliament's inquiry on foreign interference and with the US Congress last year. If we remove from the playing field bad actors that are seeking profit, it is easier to address the state actors.

Finally, Twitter is a public platform that is used by Governments every day. I have certainly had conversations with Governments around the world and with organisations like the NATO StratCom Centre of Excellence. A big part of the challenge here is public diplomacy. If foreign states are using media, using their diplomatic accounts to propagate a narrative, the most important thing is that Governments who share our values respond to those narratives directly. Content moderation is not the solution. This is a geopolitical public diplomacy matter. We work with Governments to improve their use of Twitter to make sure that they are able to fully engage in the debate.

Q231 **Chair:** Is there not an element in which content is repackaged and then promoted through Twitter? They may no longer be able to advertise on your platform but, at the same time, what they are doing is repackaging and pushing it through your platform anyway, so it is getting the wide scope and viewership that they desire.

When it comes to the monetising issue—and this is my final question—what is the way around it in terms of ensuring that you are able to stop these actors making money and basically seeing a benefit from pushing out so much disinformation?

**Nick Pickles:** We heard this in Mr Efford's question in the first hearing, I think. A range of different regulatory responses is needed here. Some of it is that there is an ad tech industry, companies you have never heard of that monetise these sites. They are the networks, the content



## HOUSE OF COMMONS

distributors, and they are often not the companies that you hear from in hearings. They have to be part of this. For example, it is not just monetising through advertising; it is sometimes monetising through selling products, so someone is on a website being sold a health supplement. I think that is definitely an area where regulators such as the Food Standards Agency and others should be intervening. You can take a very broad look at the problem, and lots of different actors have different roles to play. Our job, once it gets on to Twitter, is making sure they cannot buy reach and making sure that fake accounts are shut down.

To your earlier point about academic research, one of the things we have been very heartened by is that we took a decision to publish the full archive when we remove information operations by a state. It is a unique archive. It is the full tweet data. Researchers around the world are making use of it to understand the tactics of Governments, whether it be Venezuela, China, Iran, Russia and so on. That kind of data access is critical. If there are moves that parliamentarians and legislators around the world can make to encourage that kind of disclosure, I think we will all be better informed.

**Chair:** I am afraid time has caught up with us. It is very kind of you to join us this afternoon, Mr Pickles. Thank you, I very much appreciate it. That concludes our session.