

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

Oral evidence: Online Harms and Disinformation,  
HC 234

Thursday 30 April 2020

Ordered by the House of Commons to be published on 30 April 2020.

[Watch the meeting](#)

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

Questions 1 - 113

## Witnesses

**I:** Stacie Hoffmann, Digital Policy and Cyber Security Consultant, Oxford Information Labs, Professor Philip N Howard, Director, Oxford Internet Institute, and Dr Claire Wardle, Co-Founder and Director, First Draft News.

**II:** Alina Dimofte, Public Policy and Government Relations Manager, Google, Richard Earley, UK Public Policy Manager, Facebook, and Katy Minshall, UK Head of Government, Public Policy and Philanthropy, Twitter.



## Examination of witnesses

Witnesses: Stacie Hoffmann, Professor Philip N Howard and Dr Claire Wardle.

**Q1 Chair:** This is the first meeting in this Parliament of the Sub-Committee on Online Harms and Disinformation as part of the Digital, Culture, Media and Sport Committee. This hearing has been convened in order to get evidence concerning Covid-19 disinformation. We will be hearing today from a panel of experts who have been studying the subject closely and we will be asking questions of the social media platforms later in panel 2.

First, I want to check with members to see whether or not any of them wish to declare any interests relating to this session. There are no interests to declare. Thank you.

We will crack on with our first witness, who is Professor Philip Howard, Director of the Oxford Internet Institute. Good morning, Professor Howard. I am going to kick off with the first question. What is the situation in terms of the social media companies and Governments? What are they doing to combat disinformation and what more needs to be done, in your opinion?

**Professor Howard:** We know that there is a significant amount of disinformation about Covid. Over the last couple of weeks we have learned that misinformation reaches 1 billion social media user accounts around the world and much of this content is generated by state-run agencies and much of it—almost all of it—is in English. It is targeted at people in the UK, people in the US, English language users.

It is hard to measure the impact. We don't really know how many people believe the things that they are reading—the low-quality information they are getting about Covid. In a recent study we did with colleagues at the Reuters Institute here at Oxford university, we found that roughly 25% of the population believes Covid was created in a lab and released either accidentally or purposefully on the population. The Oxford Internet Institute is trying to generate weekly memos, weekly briefings on what themes are being covered by the accounts pushing this information. Overall, the themes involve trying to degrade our trust in public institutions, collective leadership or public health officials.

The content that generates from authoritarian regimes is about how great Russia and China are doing in advancing the science and solving the problems and sending aid to the UK and other countries. Again, it is difficult to know how much of this stuff people internalise but there is significant volumes of it. In this case, the myths about Covid are deadly.

**Q2 Chair:** Is it specifically Russian and Chinese?

**Professor Howard:** They are the largest generators of the state-backed content. Some of it is generated by the far-right extremists in our country in particular but most of the content comes overwhelmingly from CGTN and Russia Today, the state-backed media agencies in those countries.



Q3 **Chair:** Is there a difference between the types and focus of the disinformation that is coming from Russia and that coming from China?

**Professor Howard:** The Chinese response is mostly about trying to make sure we don't start calling this the Wuhan virus. The Chinese tend to respond with misinformation by volume, so they will call up 100,000 fake accounts or set up tens of thousands of fake users. If any of us looked at these users on Twitter or Facebook, we would know that they were just created last month and have no real people behind them. The Russian style of misinformation tends to involve accounts that have been around for a while—long legends with a tweet about sports and a tweet about soap operas and then start talking about politics or politicising health news, so their style is a little different.

Q4 **Chair:** What you are suggesting is that there is a long antecedent for the Russian bots and, as I understand it, there are times of the day that they come online. They almost work standard office hours. The Chinese is probably much more sophisticated in its propaganda scope.

**Professor Howard:** I think it is much harder for us and average users to tell the difference between a clear bot that has no picture and no history and some of these accounts that get into all of our social media feeds and start sharing information. One day they start waking up and spreading conspiracy stories about Covid and that is how the content leaks into our social media feeds.

Q5 **Chair:** The Government set up a counter disinformation unit to deal with, they say, 10 instances of disinformation a day. I understand there is no new funding in place. What is your view of our governmental response so far to the disinformation relating to Covid?

**Professor Howard:** My sense is that it has been strong, and it needs to be stronger and continue to be strong. It is unfortunate but the major platforms, the social media firms, do not really share data. The best data we have is months old, it is not quite adequate and does not cover all the features that these social media platforms provide. The misinformation initiatives that the Government have are very important because you have the authority to collect and collate and analyse information in the public interest and the firms don't act in the public interest. Independent researchers like myself, at OII, or investigative journalists, don't have access to the same levels of information—the levels of information that we need to help fight this.

Q6 **Chair:** As someone who has dealt with this for three or four years, that is a constant refrain. What information should the platforms be sharing with academics like you and also with Government and the wider community?

**Professor Howard:** First of all, we need more information about the limited pools of data they release. Sometimes they will release small chunks of accounts that they have verified as Russian but we will have no histories. We will have no history of the interaction those accounts have had with citizens in the UK. We will have limited ability to retrace what



those accounts actually did, the messages they pushed out. Also this information tends to be released weeks or months or years after the fact. We are in a situation where the misinformation is changing daily, the major themes change from week to week. They depend a lot on what is happening in politics in other countries. But getting access to a representative sample of this data in real time is what we need.

**Q7 Chair:** I was quite struck by some figures—I think it was in one of your reports—about the percentage of the false posts staying up, not being taken down by the social media companies. As I understand it, it was YouTube at 27%, Facebook at 24%, Twitter at 59%. I imagine that Twitter has a greater volume of posts and that may account for some of it but apart from that, why do you think there are issues specific to Twitter?

**Professor Howard:** I think it is because on Twitter it is particularly easy to repost or to change slightly the content. YouTube has the same problem. If a user flags a video as being patently false, YouTube takes a day or two to evaluate that. It makes a decision about whether it is propaganda based on its own rubric of what counts and what does not count and then makes a decision. The person who put the content up can edit the title slightly and reissue it. One of the problems is that the platforms do not have consistent standards for deciding what to take down and in some ways it is up to Government and public agencies to compel, to encourage consistent standards.

**Q8 Steve Brine:** Good morning, Professor Howard. Thank you for joining us. Could we talk about celebrities and influencers, which is one of those dreadful words that has entered the English language, being taken in by misinformation online and then spreading it and the conspiracy theories that go with it? Of course, they are influencers, many with great affluence to themselves from being influencers, because they have a lot of people following them. What are your thoughts on the threats that they pose in spreading disinformation?

**Professor Howard:** In some ways the influencers, as we call them, are sort of the gateway drug. If a prominent Hollywood star or a prominent political figure says things that are not consistent with the science or the public health advice, some people will go looking for that stuff and they will spread it. That is how misinformation develops. Those human influencers are often the pivot point that takes a lie from something that bots just share with each other to something that passes in human networks. As far as I can tell, the best solution to dealing with that kind of misinformation is for users to flag things to the platforms and for the platforms to act with some kind of warning label. You don't want to silence entirely the social media accounts of most of these figures, especially the prominent political figures. That is a very political act, but labelling some of the bad health instructions that come from political figures as misinformation is a useful step.

**Q9 Steve Brine:** Political figures, of course, are one thing but then you have general celebrities, don't you? I would loosely class David Icke in that



category, but you have other celebrities. Do you think they have a social responsibility to be even more careful than anybody else about spreading information online?

**Professor Howard:** Certainly. That is a great observation. Anybody that has a long follower list needs to be attentive and sensitive and aware of what they are sharing. One of the good findings that we have had in our research is that for the most part when it comes to the health crisis people do trust the NHS, and they will go to public health officials first for health information. A number of celebrities are doing constructive things with their follower lists. There are definitely a few bad actors who have passed conspiracy theories out but on the whole most people most of the time will not see that stuff. The audience for that stuff is fairly small.

Q10 **Steve Brine:** Is the technology merely the vehicle that has made it worse or has this always been the case? Has your research taken you into historical perspectives? You will remember that in the 1960s John Lennon said that the Beatles were more famous than God. That was spreading misinformation. Of course, he had a different platform in which to spread it but it still led to a worldwide response.

**Professor Howard:** Certainly. There has always been political misinformation and health misinformation. The difference now is that it is targeted and it is distributed along particular networks. There are sensitive communities that get more misinformation than others. Its spread is much more dependent on our own information habits, our own ability to keep trolls out of our accounts and out of our social media feeds. It moves so much faster. It will get to millions of people daily.

Q11 **Steve Brine:** You touched on other states with the Chair earlier. How might other states be exploiting the current crisis to spread disinformation? What are the recurring themes of the misinformation, the disinformation that is emanating from foreign sources? Are you able to tie any threads between them?

**Professor Howard:** Yes. I think there are two or three broad themes. State-backed media in other countries use social media to promote the idea that democracies are weak, their leadership is weak, democracies are not handling the crisis very well, that soft authoritarianism, a light touch from the police state, does lock everybody in place until the virus is done for. Democratic institutions are failing us and preventing us from dealing with the health crisis properly. The second theme has to do with how authoritarian governments are spending more on science and sending aid to the democracies that are struggling. Then there is the conspiracy sensationalist story that maybe the virus was created in a lab in the west and is either targeted at our own minorities or targeted at foreign governments.

Q12 **Steve Brine:** What do you think is the motivation behind that? What would encourage them to do that abroad during a pandemic? The logical consequence of what you are saying is that people in the west would say,



## HOUSE OF COMMONS

“Okay, clearly we should give up on democracy and turn to being an authoritarian state”. I am guessing that the students at Oxford, who I am sure you interact with regularly, don’t suddenly see it that way. Looking at the motivations behind them and the realistic expectations behind other states acting this way, what do you think they are hoping to realistically achieve?

**Professor Howard:** There are two actors. Let’s talk about China and Russia, and we can talk about Iran if you like. I think the motivations are different. The Chinese want to prevent us, stop us from thinking of this as a virus that originated in China. They want to make sure that we deal with the issues internally and don’t notice how the Chinese Government repressed the original medical research about the nature of the crisis. For China it is about a little bit of face saving and getting the story out there for the English language audiences.

I think the motivation for Russia is much more about whittling away at our trust in elections and public affairs and getting us to use less trustworthy news sources for our trust in elections and public affairs and for our news. I think the motivations are very different and we may see this next time there is an election. We probably won’t have an election for a while but there certainly will be one in the US. One of the common strategies for authoritarian governments is to peg the failure of public service at election time. We are getting ready to study the US 2020 election. We fully expect to see stories about how Covid means the election has been cancelled and you shouldn’t bother anymore. Any messaging that helps discourage voters in the UK from showing up or being good citizens is powerful content for the Russian Government.

Q13 **Damian Green:** Good morning, Professor Howard. I pick up on your point about the Chinese Government, who are clearly at the heart of a lot of the misinformation for obvious reasons. Can you identify whether they are doing it all from inside China or do they have a network of people around the world using servers so that it does not look quite as obvious as it might do otherwise?

**Professor Howard:** I think the answer is that there is a homegrown base of staff. Sometimes they are military units that have been retasked to doing social media, sometimes they are simply public-private partnerships, state-run agencies that employ people to plant stories. The state-run media, CGTN—there is a range of state-run media properties—will generate these stories and they have their own social media teams. I think your instinct is right, there is a wider network of people, Chinese citizens, who are living abroad. I wouldn’t say that they are paid in service of the state but they are in part an audience for this content and they reshare along their own social networks: students studying abroad, not all of them but some of them, recent immigrants, people who are Chinese speakers, mother tongue speakers. The same is true for the Russian networks. The stories begin with organisations based in the country and there is a wider diaspora of people who share the content and pass it along.



## HOUSE OF COMMONS

Q14 **Damian Green:** Can you identify them? Presumably there are lots of Chinese students at Oxford at the moment. Have you seen that phenomenon personally?

**Professor Howard:** I experienced it when I was working in the United States. I haven't seen it here as much. It is difficult for us to know who is working specifically for the Chinese state and who is simply a fan of its content and likely to pass it along to their own friends and family.

Q15 **Damian Green:** Another area that we have been discussing is influencers. It would be wrong to let this session go without mentioning President Trump and disinfectant. One of the things that we observe is that despite, no doubt, the best efforts of the world's media, they have not yet identified a single person who has actually drunk bleach as a way of stopping Covid-19 after what the President said. Apart from the fact that this might encourage one to feel better about the basic common sense of the American people, which is good news, does that slightly suggest that when somebody who in this instance can be described as an influencer says something quite wild, that people don't accept it?

**Professor Howard:** I think you are right, but this depends a lot on levels of education and levels of trust. In the UK and Europe most people would know not to drink bleach and would double check that one before acting on it. There would be a very small portion of the population who are fully trusting of everything that President Trump says and might call it in. There have been reports of governors who say that their state health systems are taking lots of queries now about drinking bleach but, you are right, I don't know that anybody has actually done it.

The better scenario would be to have political leadership that was giving consistent advice and backing up the public officials who say, "Stay at home", not calling for the liberation of states where the public officials have asked people to stay at home. Consistent messaging would do more to help in this crisis.

Q16 **Damian Green:** On the basis of consistent messaging being a good thing, how would you assess the British Government's performance in that regard over the course of the crisis?

**Professor Howard:** I think it has been strong. A lot will depend on how well the app goes. There is a lot of coverage about how the NHS's contact tracing app will work. I think it is likely that that piece of technology will be a very important point of contact for citizens. It will be a very important demonstration of the Government's and the NHS's ability to organise a working tool that will help people manage their own health in a way that respects their privacy but also protects the public. Consistent funding for studying the problem of misinformation, for equipping the Home Office with the right knowledge at the right time and getting this app right—the balance of privacy and public health—are the two things that I think are going to be important for the Government's activities in the next few months.



**Q17** **Damian Hinds:** Professor Howard, you have talked a lot this morning about Russia and China. Could we nail down some parameters? In your report you list a lot of different states with some activity of their own but, to be honest, some of the amounts of money listed, the numbers of people, look really quite small, in some cases \$10,000 or less. Could you try to pin down your best estimate of the proportion of activity that comes specifically from Russia and China in international disinformation? Setting aside domestic propaganda, or what you might call straightforward propaganda, what is your estimate of how much disinformation specifically is coming from Russia and China?

**Professor Howard:** About all topics or Covid in particular?

**Damian Hinds:** Give two answers if it is more illuminating.

**Professor Howard:** Okay. For Covid I would guess that it is two-fifths China and two-fifths Russia and the rest is from many other states. Often that small fraction is content directed for their own countries or for neighbouring countries. There is misinformation about Covid in Arabic but that is mostly governments in north Africa and the Middle East generating content targeting each other, not targeting English speakers. The vast majority at this point comes from Russia and China.

In the UK in particular there is probably a more significant domestic, internal supply and this is always difficult for us to be able to talk about. The far right ultra conservatives—not capital “C” conservatives, ultra conservatives—extremists, white supremacists in this country generate a fair amount of content and that leaks into mainstream conversation once in a while, and other countries would not have that challenge.

Broadly, over all issues, I would guess it is maybe a quarter Russia and a quarter China now. China’s only other care about what we think has to do with protests in Hong Kong. They do not really do much in other issues on social media. Those are my broad estimates.

**Q18** **Chair:** Thank you, Professor Howard. We are now going to turn to our second witness, Stacie Hoffmann, digital policy consultant at Oxford Information Labs. Good morning, Stacie. You have spoken in the past about terms of service and the fact that social media companies love trumpeting their terms of service and how they do all they can to enforce them. Are they wrong in this approach and if they are, what is wrong?

**Stacie Hoffmann:** I think it is important that we take a historical look over the last few years. We have been studying this since 2018. When we first looked at 37 different terms of services from the social media companies, we found that in the period up to and post the 2016 election cycles they all had the ability to counter misinformation that was being shared on their platforms but they were not enforcing their own terms. That means that they were not taking down content, they were not suspending accounts, they were not placing notifications on things. In the first instance their reaction was very slow and it was really reactive to what was coming out in the news.





## HOUSE OF COMMONS

The first big changes we started to see were at the end of 2017 and then into 2018 when things like Cambridge Analytica became front-page news for Facebook. For Google there was a holocaust deniers issue that was being highly populated in their search queries. They responded to what was happening then and started to enforce their own terms and then slightly modify them as well. I think what we have seen since about 2018 to now is a lot of testing and refining of different tools that they can use. Elections and politics have been a really big focus for them and that has been globally. They started a lot of their initiatives in the US or in Europe and those were populated and replicated in different countries around the world.

Covid is a different type of stress test for these platforms because we are starting to see what they can do in some senses. I think it is true that they are enforcing their terms more now than they were before but they also do not necessarily have a lot of terms specific to misinformation, disinformation or false news. They are using other tools in their toolbox to do this. For instance, spam is a very common policy that is used against misinformation because that helps to take down bots or automated accounts that are helping the spread of junk news. Impersonation is another long-standing policy that they use and that can be useful. For example, when you are looking at electoral jurisdiction, if you have somebody claiming to be the leader of a Tennessee GOP right-wing party in the States but they are actually a Russian troll, which is what happened on Twitter, you can start using your impersonation policy to take that down.

They have announced a number of different things to enforce their policies. One is human content moderation. Facebook has hired thousands of people to help with their human content moderation. Before all of this came about, they were really hush hush about their use of people and they tried to focus on technology for that use.

Q19 **Chair:** What you are saying there, effectively is that they turned away from any well-known methods by which to do this. Are they now seeing the value of people?

**Stacie Hoffmann:** It is not that they didn't see the value before. They did have a network of human content moderators but they liked to promote the idea of technology being able to identify misinformation or harmful content online. They didn't want to talk about human content moderation. This might be because it is a very sticky subject to have somebody sitting at a computer screen looking at sometimes horrific and abusive content. It takes a lot of oversight and a lot of care for your staff. I am not quite sure why they were not talking about it but they definitely had plenty of human content moderators before and they have expanded that now.

Q20 **Chair:** I think you said that the terms of service are so broad that there is not anything specific to deal with disinformation. Is there a need for it?



**Stacie Hoffmann:** It depends on the platform. Facebook has a section on fake news in its community guidelines. It is not very clear about what they take down versus what they reduce the reach of. They tweaked algorithms. A huge response for them since 2018 is that they are all taking up these algorithms or AI to help this issue. They started tweaking their algorithms to reduce the reach of content, but it is really unclear at what point they start taking down content. The terms of service that are directly related to misinformation or junk news tend to be very high level and confusing whereas they can employ other types of policies much more effectively. That is what we have seen so far.

Q21 **Chair:** It is a common refrain, isn't it? If you infringe copyright they are very quick on it but if you are spreading fake news and disinformation across all social media platforms there does not seem to be that sort of focus on it to that extent.

**Stacie Hoffmann:** Right.

Q22 **Julie Elliott:** Welcome, Stacie. I want to build on what you have just said in answer to the Chair there. How effective do you think the social media companies have been in the challenges posed by Covid-19? It is more of an infodemic than big news. It is a slightly nuanced thing that they are dealing with. How effective do you think the various platforms have been in dealing with that?

**Stacie Hoffmann:** They have been very proactive related to Covid-19 versus other areas of disinformation. That is my first reaction. We saw in late February, early March the first notifications of actions that they were taking or plans that they were developing. Instead of waiting for the infodemic to explode, which happened by the end of March, they were taking some earlier steps, which was very welcome. It is becoming clear that public health information is being treated differently than political information or political speech or other types of speech on these platforms. For example, Facebook has a policy that they do not put any oversight on political speech and they allow political ads on their platform. However, they did take down a post by Brazil's President Bolsonaro because it went against World Health Organisation guidelines on Covid.

We are seeing them relooking at their own policies and applying them differently in this space, I think because it is a public health emergency. We are seeing them develop tools in their toolbox and it is really interesting. Common things that they have done include giving free ads to authorities, the World Health Organisation and local health organisations, NGOs. They have put in more restrictive ad policies, so you can't advertise things like respirators or face masks or hand sanitiser. They seem to have increased their takedown and notifications as well, at least that is what they want you to believe. I don't have data on that to see if that is true. They are promoting authoritative content in a few different ways. They have portals that you can go to that they have



curated content for and then they have the different tweaks to algorithms that they have been working on.

It will be interesting to see if they use the policies and practices that they put in place for Covid—which I definitely welcome; I think it is a step in the right direction for these companies—going forward, if they stay in place post-pandemic, if they refine them to be used in times of electoral processes or in everyday speech online. A good example of this is that on YouTube, Google has taken down advertisements from videos that are spreading junk news or false information about Covid-19. In our most recent study, which we have yet to publish, we found that Google and Amazon are the two biggest ad providers for junk news purveyors and those are the websites that are getting those click-throughs to try to gain money as part of a round ecosystem. We have known that there is a plethora of websites since 2016 that have been key purveyors of junk news. Why do they still have ads? Why are they still making money when we have known this? If YouTube can do this for videos now, can Google implement that in the wider ecosystem of misinformation going forward?

**Q23 Julie Elliott:** It sounds to me as if you are saying that they are attempting to use their mass influence on people in a positive way to try to help in this pandemic. Am I right in thinking that is what you were saying there?

**Stacie Hoffmann:** Yes, they definitely are. They are starting to be a bit more bold in the tools that they have and implementing them and doing so quickly as well.

**Q24 Julie Elliott:** There are lots of differences between the actions of the various big social media platforms. Which ones do you think have been doing it better? What do you think are the best tools they have been using that perhaps the other companies could take on board from them?

**Stacie Hoffmann:** If you look at Google, Facebook and Twitter, Google is kind of an outlier. It has a much bigger umbrella of what it provides and, of course, Google search is one of its key businesses. It focuses on technical tools instead of policy tools to respond to these issues, but it has not implemented a report function. It has one in YouTube but not in its search.

**Q25 Julie Elliott:** Do you think it should have one?

**Stacie Hoffmann:** I think it should. Why not? If I am searching for something about Covid and something like Breitbart is the first website that pops up in a search—this has not happened but just say it did—I would want to be able to report to Google that I don't think this is the right content to be showing me in this search or some kind of feedback loop there that it does not have in its search function. On the other hand, Google has tweaked its algorithms—and we know this from studies that we have done and that we have seen—since 2016 to help reduce the prominence of junk news or misinformation in their searches, but those also come back up. It takes about not even a year for the reach of those



## HOUSE OF COMMONS

websites to go back up again. It is an iterative process, and I think that these companies need to acknowledge that this is an iterative process that will be changing.

There is another good example from Facebook. Most recently it has implemented a tool where it is going to inform people who have engaged with misinformation on its platform. This was just a couple of weeks ago. This is important because it feeds into what is called an inoculation theory where you start being proactive and you start showing people what misinformation is and educating them about it so that they can more easily identify it going down the road. It will tell people if you have shared, if you have commented on or if you have liked a piece of misinformation. This is particularly important in a period like the pandemic where what was misinformation today might not have been misinformation two weeks ago.

At the same time, they are not telling people who saw or who read that piece of content. Facebook knows how long somebody stays on a piece of content. That is one of the datapoints that they collect, so it knows that even if you didn't like it you at least reacted or had some kind of engagement with it. They could also be notifying those people.

They are taking good steps, some more than others, and I think they can always go a little bit further.

**Q26 Kevin Brennan:** Good morning, Stacie. I was quite interested in what you were just saying about advertising. It strikes me that there is a very uneven playing field here between social media companies and traditional broadcasters. Do you think it is fair to say that we have seen a transfer of advertising away from broadcasters who have to follow standards and regulations in the presentation of the news that often they are required to present us with, particularly public service broadcasters that rely on advertising, and social media companies who effectively have taken that advertising revenue away from those traditional sources and use it, in effect, to finance fake news?

**Stacie Hoffmann:** There is definitely a regulatory gap between what we see on TV or in print news and what we see on video, not streaming sites such as Netflix but things like YouTube or Facebook or Twitter. As you go from traditional media down to something like YouTube there is less and less regulatory oversight. That is absolutely true.

**Q27 Kevin Brennan:** Do you think there should be, effectively, the same regulatory oversight for advertising on social media platforms as there is on Ofcom-regulated broadcasters?

**Stacie Hoffmann:** I think there should be appropriate regulation, definitely. It is not a case of being able to just copy and paste the current regulations on to online platforms, because it is a different type of media. It is a completely different ecosystem as well, but there definitely should be the same expectations and the same kind of levels of restrictions or



expectations on the actors involved as there is in current regulations for traditional media. It is very tricky. When we start talking about regulating online spaces, it is a bit trickier but that does not mean it cannot be done.

**Q28 Kevin Brennan:** Have you seen any evidence that platforms provide any funding or support for quality journalism to try to counteract some of this information on their platforms, before or during this pandemic?

**Stacie Hoffmann:** Yes, definitely. Before the pandemic there was a lot of focus put on to this, and during the pandemic they have increased their funding or support for quality journalism and fact checking. They did this in a few ways. One is supporting journalism itself. Google has an initiative to try to help journalists change their business models and attract more subscribers to their services. Fact checking is a huge area for these companies. They are trying to help traditional journalists by bringing in fact checkers to support and push people towards more well-known news sources. Facebook probably relies on fact checking a bit too much. It is taking responsibility and liability off itself and trying to put it on fact checkers, which is not the best route but that is another way.

There are literacy programmes as well. Google had a programme in the UK that supported media literacy for people so that they could identify fake news or how to engage with news, things like that. Of course, electoral outreach has been a big one as well. During electoral cycles, they have reached out to governments and campaigners to help them learn how to use platforms and media better and things like that. They have put a lot of money behind this, and there has been a huge amount of money during Covid. Twitter, for instance, has put some money behind two journalism projects that are focusing on minority and women journalists to try to support them during this difficult time since they know that the amount of funding will have gone down.

**Q29 Clive Efford:** Ofcom's research has found that most people disregard the misinformation they see. What would drive people to disregard information that they get from authoritative sources or to believe information that they get from sources that don't appear authoritative? What concerns do you have about that?

**Stacie Hoffmann:** Countering disinformation is very complicated. There has been research that shows that once you start engaging with the material itself you come out at an equal footing. You end up with somebody who probably won't believe the misinformation, but they won't believe the fact either. They will be in a middle ground and that is very worrying because that in itself will diminish people's trust and reliance on authoritative sources. A huge push right now by these platforms is pushing authoritative sources for information.

You need to get ahead of the curve. I think that is the key here and that goes a bit to the inoculation period. I have not done a huge amount of research in this area but it is about educating people beforehand about



how to engage with media online and how to identify something that is looking for an emotional response versus something that is trying to engage you intellectually. That is what a lot of the disinformation does. It elicits a very strong reaction one way or the other but we do know that the algorithms are rewarding negative reactions.

**Q30 Clive Efford:** We had questions to the previous witness about influencers and people in positions of authority that may influence people. Do you think there is a requirement on them to be more responsible about the messages that they relay and their sources of information?

**Stacie Hoffmann:** This gets into a very sticky area, particularly if you are looking at policy or regulation. At what point do you start taking a normal human being, who just happens to have thousands of followers on Twitter or Facebook, and start treating them differently than me or somebody else? But when you have a platform that reaches that far that quickly, I think there is a level of responsibility that needs to be taken on board. This might be something that is down to the platforms to enforce, it might be something that the UK looks to educate people about or put in guidelines about who influencers are and what our expectations are of those people. It gets quite tricky, but I think that we need to really look at that.

We talked a bit about public figures being influencers. When we did our research with the Oxford Technology and Elections Commission we found that having a public registry of public officials, official campaigners, something like that, would be very useful for all of the parties involved. It would help us have oversight of those actors and hold them accountable to how they are using social media. It would help social media companies know how to impose their own policies on particular users and it would help fact checkers fact check what is being posted. Most importantly, it would give the public somewhere to go for reassurance that the account that they are following is the account of their local MP or of a campaign that they want to donate time or money to. It is a good check and balance for everybody. You can't necessarily apply that to influencers broadly but that is something that we can do on a local level, particularly in the electoral space.

**Q31 Clive Efford:** I want to mention the issue of President Trump and bleach. This issue has been around for many years. There are some very sinister individuals and organisations behind promoting ordinary, everyday cleansing bleach as a medicine. I came across it being promoted as a cure for autism, and parents were being encouraged to give it orally to their children to cure autism. Some people were taken in by this and children became very unwell. When it comes back to this issue about fact checking, what responsibility or set of guidelines would you advise for people in positions of authority before they repeat things that are so obviously wrong but also extremely dangerous? I found that the Food Standards Authority and the Medicines Regulation Agency couldn't do anything about the people who were promoting this. It was disturbing



## HOUSE OF COMMONS

indeed. When a politician repeats something like this, it gives oxygen to some very sinister operators. What would you say about that? What guidelines should politicians follow when they are using their positions of authority to promote something?

**Stacie Hoffmann:** This is going into a very sticky area where you can't necessarily take away free speech either. I think it is about being able to have authorities that people trust and this is where trust really comes into play. We need to reinforce those trusted relationships so that if somebody, an influencer, whether it is a politician or not, says something online, we have trusted authorities that we can go to and get that information from. There will probably have to be an increase in engagement from all parties as well. We can't put all of our reliance on fact checkers. Right now there is just not enough capacity there, so one way would be to develop capacity.

Another way is to look at when it comes to political speech—and if you also look at the way that we apply offline laws online—we need to stick to the laws and the regulations for how we address political speech and apply that online. There is a disconnect where regulators in general are being very cautious about how they approach the online space, but if it is illegal offline, it should be illegal online as well. We need to make it clear that they have the authority to take action when needed.

Q32 **Chair:** Thank you very much, Stacie Hoffmann, for your evidence today. We are going to move on to our third witness, who is Dr Claire Wardle who has appeared before us before. Claire is strategy lead at First Draft News. Good morning, Claire. Thank you for joining us. Have the social media companies had a good coronavirus war, in your view?

**Dr Wardle:** We have seen them take more moves, but we also have to recognise that on other forms of speech, you have two sides but there is nobody now saying to the platforms, "We want more coronavirus misinformation". They have been free to take steps on other issues—even health issues like anti-vax speech. They have taken more steps but we have to recognise that right now we are in the middle of a global pandemic and there are real world harms. When we say we hope they do the same when it comes to political speech, we have to be careful about the different types of speech and what we mean by harm. Yes, they have made steps; we have got a lot to learn; they still have a lot more to do.

Q33 **Chair:** Have you seen any in particular that have instituted good practices and any that you think have, frankly, been laggards?

**Dr Wardle:** An example is Facebook notifying people. As you said at the beginning of this, we have so little data. We are in the middle of a natural experiment right now, so what I would like to see is the platforms do more but then allow academics to test alongside them to see what the effects are. Maybe by notifying people that they are seeing false information, there is an effect that means that people trust all information if it has not been flagged. We have to be very careful.



All of them are doing different things but what we are lacking is transparency and oversight. We are seeing them take down more content but there is no available archive for academics to know what they are taking down. Right now one thing that Governments could do is require them to be transparent about what they are taking down and to have independent oversight committees to help make the decisions about whether or not these Silicon Valley decisions are ones that our society would stand behind. They are doing more, but we have less oversight than ever before, and I think that is something that we need to be really critical about.

**Q34 Chair:** Despite the lack of oversight, which is a regular refrain from yourself and also the other witnesses, how much disinformation, in your judgment, is coming from malicious foreign actors—gangs looking to monetise fake news and disinformation—or, just frankly, as in the case of President Trump, what is seen as rank stupidity?

**Dr Wardle:** We are certainly seeing state actors and also have to argue that we are seeing the US being part of this. There are global narratives pushing information that supports their particular position. We are seeing a huge increase in scams and hoaxes and people motivated by financial gain. We are also seeing domestic actors pushing this content. We have talked a lot this morning about disinformation but we also have to recognise that there is a lot of misinformation, which is people becoming nodes for this and sharing it with one another because they are scared. What I am seeing is more people sharing false information and conspiracies than I have ever seen before and I think we have to recognise the human element of all of this.

**Q35 Chair:** Much of the disinformation has been pretty low tech, which is quite interesting considering that we talked about the dangers of deepfake. Why do you think that is? Is it that sort of community issue in which people are so tense that they just want to share in something in a very low tech way? Why is it that we are seeing low-tech disinformation rather than deepfake?

**Dr Wardle:** Mostly it is why would you spend money on a deepfake when you can have as much impact with a WhatsApp voice note. I think what we have seen over the last two months is a recognition that simple techniques are very effective. It might be resharing an old photo from three years ago and saying it is actually from today. We have not talked very much about closed messaging apps so far this morning. That is where we are seeing a huge amount of this kind of content travelling and similarly in Facebook groups. We are seeing people saying, "This might not be true but I am going to share it with my friends and family just in case". When we see more and more people sharing this, there is this kind of everybody is scared thing and so you see friends and family sharing information that they really should not be but they say, "It probably isn't true but just in case, and I love you". That dynamic is critical in what we are seeing and how we are seeing people respond to the pandemic.





**Q36 Clive Efford:** Can you say something about what you have come across for the motivations behind people creating misleading information? What do they get from it? Is it profit? Is it notoriety? What are the motivations behind creating this stuff?

**Dr Wardle:** It is either, as we said, foreign influence or it is about motivations for money. I would put the people who are doing this in the bucket of social and psychological. It is trying to get away with it, "Can I get an influencer or a politician or a media personality to reshare something? Can I convince my community that there are scam artists, because wouldn't that be funny?" It is very difficult to say that there is one motivation. There are many reasons for this, but we have discussed conspiracies a great deal and I have never seen a time when there are more conspiracies circulating and conspiracies being shared in spaces that are much more mainstream. For example, we have been monitoring the 5G conspiracy since January and it was bubbling along, bubbling along, bubbling along, and then about three weeks ago you had Woody Harrelson in the States and Amir Khan and Amanda Holden in the UK and all of a sudden it turned into something that the mainstream media were talking about.

It is easy to dismiss conspiracies but we have to understand why they are taking hold. There is not a good origin story for the virus and so this information vacuum is allowing misinformation to circulate. The reason that people attach themselves to conspiracies is because they are simple, powerful narratives and right now people are desperate for an explanation of what they are going through. They feel out of control and conspiracies give them control because it gives an explanation that they are looking for. There is so much about all of this, about understanding people's psychology, both those who are sharing but also those who are consuming and making sense of these things.

**Q37 Clive Efford:** You are saying that this has generated a huge volume of misinformation around the virus, but is there a difference in the motivations for the conspiracy theories around Covid-19 to other conspiracy theories that you have investigated in the past?

**Dr Wardle:** No, it is very similar and unfortunately people who believe one conspiracy are more likely to believe others. Larger proportions of the population are losing trust in institutions, feel hard done by, and conspiracies basically say, "You don't know the truth. I'm telling you the truth". That is why we are seeing it. There is a responsibility for us to think about why conspiracies are taking hold, because what does that say about the underlying sociodynamics of societies? I think we have ignored conspiracies at our peril and now they are coming in force during this pandemic.

**Q38 Clive Efford:** Is there any evidence that people create what is misleading information in a sort of well-meaning way? Does that have any influence? Does that play a part at all?



**Dr Wardle:** Very early on—you probably received this too—many of us saw a kind of 10 tips about coronavirus that said things like gargle with salt, hold your breath for 10 seconds—those kinds of things. There does seem to be some evidence that those early WhatsApp messages could have come from China. Wherever it came from, that was disinformation but when it was shared by people very widely it became misinformation, which was people saying, “I don’t know, but it doesn’t do any harm to gargle with salt”. You saw people unwittingly becoming part of this information network and that was just people harmlessly trying to help one another because they felt so powerless.

Q39 **Clive Efford:** Did people innocently repeating those homespun remedies have a significant impact?

**Dr Wardle:** We do a lot of tracking of questions people are asking on social media and you very quickly saw people asking Google and discussing this on Facebook, “Can I hold my breath for 10 seconds? Does it count? Can I gargle with salt? Should I put garlic in the corner of every room?” There was definitely an explosion over two weeks of people desperate for these kind of home remedies, so you do see an impact of people then questioning whether or not it is true. Whether they believe it is a different matter. It depends whether there is quality information on the other end of those search queries.

Q40 **Clive Efford:** How widespread is exploitation like this, such as those seeking to make a profit by suggesting that they can put bleach in a fancy bottle and sell it as a magical mineral solution to all ailments, including Covid-19? How widespread is that?

**Dr Wardle:** It is widespread. We are seeing a lot of people trying to sell health supplements, like elderberry supplements, but we are also seeing people selling testing kits and saying that they have been FDA approved or CDC approved, and of course they are not at all. There are many, many ways to make money right now because people are frightened and when people are frightened they are going to open their wallets. Unfortunately, scams, hoaxes, fishing attacks, we are seeing a real uptick in that kind of content.

Q41 **Clive Efford:** I suppose taking that down and identifying it is a very tricky field because of the content. You have to be sure that what you are taking down is not scientifically backed or anything. There must be a great deal of research that has to go into that at times. Are we being effective in taking those down? Have you seen any evidence that these sorts of advertisements are not getting through or we are minimising the amount that did not get through?

**Dr Wardle:** I am sure you are going to hear from the platforms. They are taking down more content than ever before. The problem is the scale of this and the scale globally in multiple languages is eye-watering. Yes, they are taking down what they can but the constant game of whack-a-mole and the ease with which you can just create another website with



another way of paying for supplements means this is just a constant battle.

- Q42 **John Nicolson:** Talking of garlic, one of my correspondents says that my lighting makes me look as if I am Dracula appearing out of the gloom. Apologies for that, I will try to sort that for the next session.

Thank you for your evidence. I was fascinated to hear you say that this is the worst that disinformation has ever been, which is an extraordinary state of affairs given this crisis. Would it be correct to say that this is the first time in this social media age that we have ever seen political leaders engage in disinformation on this level—deliberate disinformation to the public—in particular, President Bolsonaro in Brazil and President Trump in the United States? Is there a precedent for either of them and the way in which they are trying to mislead their public?

**Dr Wardle:** I would argue in the last four years we have seen a number of leaders around the world around election time pushing information that has been false or conspiratorial or misleading, but I do not think we have ever had a moment where what they were saying could potentially lead to harm. That is why it feels so distinct. That is the difference between political speech during an election versus a pandemic. It is very troubling for us to hear leaders saying things that are going so against scientific advice.

- Q43 **John Nicolson:** Which of course puts the social media companies in a difficult position because on the one hand they want to champion free speech but on the other hand they do not want to allow messages to remain from political leaders who are advising people to do things that those political leaders must know might kill people. Again, it is an extraordinary position to be in.

**Dr Wardle:** What it says is we should not be having conversations about leaving something up or taking something down. We need more innovative responses to this so, for example, questions to do with: should the media be livestreaming President Trump's broadcasts? Is there a way that we could have a slight delay and some live fact checks? Is there a way that on social media we can have more labels that suggest this has been fact-checked? They do do that, but when it comes to politicians and influencers, we do not have a way to suggest to people that we know is effective that will slow down people's recognition or make them understand that this is not supported by science. Whether or not it will work is a different matter.

- Q44 **John Nicolson:** In other words, in a President Trump press conference on a daily basis—and we know that some of the American networks have been choosing not to go with President Trump's press conferences now—but you would advocate a banner along the bottom that says, "False. False. False."



**Dr Wardle:** We need to be able to annotate speech and at the moment it is either do we have speech or not and we just need more context and more explanation as the norm.

Q45 **John Nicolson:** Going back to a previous point that was raised by one of my colleagues with the first witness, President Trump notoriously suggested that people could inject bleach and there has been some suggestion that nobody takes him seriously because we all think he is the David Icke of American Government, but it is worth pointing out that *The Hill*—which is a well-respected news website that covers Capitol hill politics—has reported that two people in Georgia have drunk bleach following the President’s remarks. So there is a record of people following what he says. The suggestion is that the two people concerned had learning difficulties but nonetheless it is worth putting that on the record. There is evidence that people have followed what he said and that it has caused them great harm.

**Dr Wardle:** Yes, there are many elements during this pandemic where we have had information, for example, about treatments like hydroxychloroquine, which have been untested, that have led to people who do need to take that drug struggling to get access to that drug. Even whether or not stay at home orders should stay in place. There are so many elements to this where we have had conflicting advice from leaders around the world and that has led to the public being confused and, ultimately, when the public is confused, it creates spaces for misinformation to jump into.

Q46 **Steve Brine:** Obviously there is a lot of misinformation out there and there always has been when the internet has been there, but we have not had a situation where so many people are sitting at home with so little to do. Is there as much a case to be made here about human behaviour and the power of click, share, through boredom, than there is through any great conspiracy?

**Dr Wardle:** That is the distinction between disinformation, which is people who knowingly create false information and push it out to cause harm, versus misinformation, which is people at home who do not understand it is false and forward it because they think they are helping people. The key to understanding dis- and misinformation is psychology. It is also about understanding that people share because they are trying to be helpful, that people right now are disconnected and the only way they are connecting with each other is through information. All our research shows the most effective thing to do for the platforms is to build in friction. If you force people to think before they share, they are more likely to think critically before they share.

For example, WhatsApp has now decreased the number of people you can forward information to and yesterday a report said it had dropped virality by 70%. Yes, of course. That is what we need to do: recognise that as humans we have lizard brains—unfortunately, all of us. It does not matter whether we are from the left or right or whether we are highly



educated or not. All of us are susceptible to information that supports our world view or makes us think that we are helping our loved ones, and that is what is happening right now. People are inadvertently sharing false information believing they are doing the right thing.

**Steve Brine:** There is a quote for our report, Chair, “lizard brain”.

**Chair:** Yes, and there are some who perhaps have more lizard brains than others. Claire, thank you very much, excellent as ever. We are now going to take a short break for a few moments, just as we set up our second panel but please stay with us.

## Examination of witnesses

Witnesses: Alina Dimofte, Richard Earley and Katy Minshall.

Q47 **Chair:** We will now start our second panel of today’s session and our first witness is Katy Minshall from Twitter who is the UK Head of Government, Public Policy and Philanthropy. Good morning, Katy.

**Katy Minshall:** Good morning.

**Chair:** We have heard in our first panel of witnesses from the Oxford Internet Institute that Twitter is twice as poor at taking disinformation down as other social media platforms. They stated that false posts on Twitter always stay up almost six out of 10 times. Why is that the case? Is it just the number of posts? What are the issues that you are facing?

**Katy Minshall:** Thank you for inviting Twitter to participate in the hearing today. If you will allow me to preface my answer—which I will provide directly in a moment—with a note that our entire company is focused on tackling the challenge of misinformation on our service. It is worth saying that what we are seeing from the vast majority of UK users is the best of humanity, the best of British, with people using Twitter in all sorts of positive ways, be it to take—

Q48 **Chair:** I am sorry, Katy, I have to cut across you. The question is: why is it that you are performing more poorly than your peers at taking down disinformation, which is highly damaging? Could you answer the question?

**Katy Minshall:** Of course. We are proud to be the only major service to provide an open public API. Tens of thousands of researchers have accessed Twitter data over the past decade for the purposes of answering those key questions. We went a step further yesterday by announcing the launch of a dedicated API endpoint specifically about the Covid-19 crisis. I saw Philip Howard testify this morning that what we need is getting access to this data in real time and that is exactly what we are providing.



## HOUSE OF COMMONS

When it comes to this report specifically, what I understand is that it covers the period from January to the end of March. We broadened our rules to cover harmful Covid-19 misinformation on 18 March. It is also a relatively small sample—225 pieces of misinformation—when on Twitter the scale we are operating at is we have now challenged 3.4 million accounts that appear to be engaging suspiciously or manipulatively in the Covid-19 conversation. As with all research, and if the OII share that dataset publicly, we would be very pleased to review it and to learn from it.

Q49 **Chair:** You said a small sample but we come across disinformation and misinformation shared by very prominent figures, either politicians or celebrities. I want to put a point to you that we find minimal instances of Twitter, for example, removing verified user status—the blue-tick status—from celebrities or politicians who are going about spreading misinformation and potentially harmful content over Covid-19. It seems to me that it is relatively simple to get verified status, it is straightforward. It does not seem so simple to lose it. Are you being robust enough?

**Katy Minshall:** Interestingly, the first ever account we verified on Twitter was the Centers for Disease Control in the US, @CDCgov. The reason we took that step and verified that account is at the time it was providing information and advice in a different public health context. People wanted to know that that account was the CDC. That is what verification is. Nothing more, nothing less. It is the knowledge that Julian Knight is indeed Julian Knight.

Over the past few weeks and months particularly that is the reason why we have worked so closely with the NHS, with journalists, to verify new accounts where people want that assurance, be it a new NHS hospital or a journalist covering—

Q50 **Chair:** Sorry to cut across you again. You are telling me about verifying. You are not telling me about the action you take against accounts that are found to be spreading disinformation that comes from a verified source. Just name me instances during this time that you have taken a verified user status off any individual, a celebrity or a politician, who has been spreading misinformation?

**Katy Minshall:** Generally we do not get into public discussion about specific accounts. I can assure the Committee that if any account, be it verified or not, breaks any rules, they will be subject to enforcement action. As I expect this Committee is aware, in recent weeks and months that has absolutely included world leaders and high-profile accounts.

Q51 **Philip Davies:** Can you can tell me how many of the problems stem from bots and anonymised accounts on Twitter.

**Katy Minshall:** The key thing here is: what are our rules and how do we enforce them. When you come on to Twitter we will ask you for your full name and for your e-mail address or phone number. When you get on to



## HOUSE OF COMMONS

the service we will ask you to abide by a series of rules. One of those rules is that you cannot operate a fake account so something like a stolen profile photo or stock images, fake location data, so saying you are somewhere where you are not. We will also ask you to abide by our rules around platform manipulation; that you cannot try to amplify the conversation through the use of multiple Twitter accounts to try to make something trend.

We have been pleased to share publicly regular updates on the number of accounts that we are challenging, be they pseudonymous or using their name. We most recently, just last week, shared that we challenged 3.4 million accounts that appeared to be engaging suspiciously in the Covid-19 conversation. But the key thing here is, as we have heard this morning, what we are interested in is a small number of often high-profile individuals who share information that is then broadcast on TV or shared by news websites all over the world and the focus has been less so on accounts more broadly that we continue to challenge when they break our rules.

**Q52 Philip Davies:** I will just try again. Can you tell me what proportion of the problems are caused by bots or anonymised accounts? Do you know that?

**Katy Minshall:** The challenges with that phrase “bots automated accounts” is I would suspect a range of Government accounts, even this Committee, could use automation to some extent. Plenty of high-profile accounts make use of things like tweet scheduling. The phrase “bot” is not a particularly helpful one here. Increasingly, external research is showing that it has often been the case that the volume of so-called bots has been largely overstated with many inadvertently designating human accounts as bots and a paper from Harvard University last month—

**Q53 Chair:** Sorry, again you have not answered the question. Just tell us yes or no. Do you know the proportion that are coming from bots or not? Just yes or no.

**Katy Minshall:** As I said, accounts use a range of different automated measures and so it would be misleading to say a specific number.

**Chair:** What I will do is we will write to Twitter afterwards as well and we will ask you a series of questions related because frankly we got nowhere there.

**Philip Davies:** I quite agree with you and I will leave it there. We are pressing for time so I will let someone else have a bash.

**Q54 John Nicolson:** One of the big problems that you have is that you do not verify people’s identities, do you?

**Katy Minshall:** We are aligned with our industry peers in that we do not ask for a Government ID or equivalent when people sign up to the service.



Q55 **John Nicolson:** What that means is that those of us who use Twitter have absolutely no idea who we are talking to, do we? People have multiple false identities, they can be absolutely anybody, and what the research shows is that the majority of accounts that are spreading disinformation may well be fake accounts. Because you do not check who people are, they can continue to spew out the disinformation nonsense, that they are doing.

**Katy Minshall:** We are proud of the progress we have made over the past couple of years in—

Q56 **John Nicolson:** I do not know why. Sorry, you have used the word “proud” quite a few times. I hate to say it but it sounds slightly pre-prepared. That is American organisational speak. There is a lot of venom and poison on Twitter and one of the big problems is the lack of verification. Why do you not verify people’s identities? We are not saying obviously that you should publish the identities but simply verify them so that you know that they are real people.

**Katy Minshall:** There are three bits there. First, you want to verify them. Secondly, how can we do it? Thirdly, will it work? If I could take all of those in turn, first, do you want to do that? I have seen all sorts of positive news cases that can come from being pseudonymous. Shortly after I joined Twitter a couple of years ago there was an extensive conversation around the #WhyIDidntReport with people who had experienced sexual assault and harassment sharing those stories and enabling them to be pseudonymous, lowering the barrier of people sharing their experience.

Q57 **John Nicolson:** Sorry, but you are answering a different question to the one I asked. I understand why you do not want to publicise people’s identities so that they can be harassed and pursued; I absolutely get that. That was not the question I asked. The question I asked was: why you, as Twitter, do not verify people’s identities? I am not asking you to publish their identities. I am asking you to verify their identities to stop both the spread of disinformation and the huge increase in online abuse that we see all the time aimed at women, aimed at minorities, aimed at trans people, to an extraordinary degree at the moment. You could help stamp a lot of this out by verifying people’s identities before you allow them to use your platform.

**Katy Minshall:** There are still the issues of how are you going to do it and will it work? In terms of how you are going to do it—

**John Nicolson:** Have you tried?

**Katy Minshall:** —there is a question of which ID would you use. Is it a driver’s licence, is it a passport? There are large swathes of the population in the UK that have neither and they may be some of the most vulnerable communities with the voices that you would absolutely want to be present on Twitter. At this point in time, do people want companies





## HOUSE OF COMMONS

like the ones here today to be asking for more personal information from people? Is that something the ICO would be supportive of?

Then the final and perhaps most important question is: will it work? There is not lots of evidence to draw upon internationally but there is a case study. South Korea had a robust real-name identity policy several years ago that they overturned when it became apparent not only had it had less than 1% of an impact on malicious communications, but it had created all sorts of new privacy problems as well.

As someone who sees examples of abuse come across my desk every day, every week, often times some of the most egregious abuse comes from people who are not only sharing their personal identity but sharing details about themselves, such as their location or their employment. There are increasingly real-world consequences for that up to and including prosecution. Of course at any time the police can reach out to us through a specific portal for police that is staffed 24/7, to make requests for further information.

**Q58 John Nicolson:** Exactly, but again I would not recommend, and clearly you would not, and I cannot imagine any member of the Committee would recommend people who feel vulnerable passing on, on their Twitter account, any sensitive information. Of course they should not, so we are all agreed on that. What we are trying to pursue is a different issue, which is: should you at least attempt to verify identity? You are right, there might be some people who do not have passports or driving licences. A lot of people do; the vast majority of people have either one or the other. You could at least begin by asking for passports and driving licences. You could ask for council tax papers. You could make an attempt to try to verify identity. What you have acknowledged here is that you have made no attempt at all and the only study that you can cite of any country that has tried to do this has been South Korea. It is a very different society from ours. Chair, if the witness would like to answer this and then I have no further questions.

**Chair:** Would you like to answer the question please?

**Katy Minshall:** The only further comment I would share is that the biggest progress we have seen has been on being proactive using technology. So whether an account is verified or whether it has its real name identity or it is pseudonymous, increasingly we are able to use our technology to look at behaviours of accounts that are likely to be breaking the rules. One in two of the tweets we take down we have detected proactively regardless of the state of that account.

**Q59 Clive Efford:** A number of years ago I complained to the police because I was tagged in a tweet that contained antisemitic comments. When the police investigated it they discovered that this person had hidden their identity by going through several servers in different jurisdictions. That was several years ago. Is that still possible? Can someone hide their identity being traced in that way still on Twitter?



## HOUSE OF COMMONS

**Katy Minshall:** That would be a question better posed to ISPs and BPM providers. But absolutely we are constantly trying to keep up with the changing nature of technology.

Q60 **Clive Efford:** Is that not a way perhaps of identifying somebody that might be intending to not behave in an appropriate way on Twitter? Is that something that you look out for?

**Katy Minshall:** We take into consideration a wide range of signals. Something like if an account tweets at a number of accounts that do not follow them back in a very short period of time that is a signal of someone engaging in bad faith regardless of what they are tweeting about. Or if there is an account that is tweeting at exceptionally high volume on a specific hashtag, that is again another signal, combined with a wide range of others, that enables us to surface for review those accounts that we suspect are bad actors and are breaking the rules.

Q61 **Clive Efford:** Yes, but if someone sets their account up in that way is there any way that you would become aware of that and would you question what their motive is for going to that much trouble?

**Katy Minshall:** I hope you understand I would not want to get into anything further past this point because it could give insight into some of our defensive measures and how we do identify accounts that appear to be breaking the rules. I am of course very happy to follow up with the Committee in a different setting.

Q62 **Chair:** What you are saying is you do not wish to disclose to the Committee how you try to prevent this from happening; is that correct?

**Katy Minshall:** In this public setting, I would prefer if we could find a more private setting.

Q63 **Chair:** What we will do is we will ask for you to supply this evidence on a confidential basis to us and we will ask you in writing to do so; if that is okay? Will you be willing to supply that?

**Katy Minshall:** Of course.

**Chair:** Clive, sorry to interrupt.

Q64 **Clive Efford:** I will leave that there, Chair, but can I just clarify one or two things you said about blue tick accounts earlier on? You are categorically stating you would treat those accounts like any other account if they were carrying misinformation or disinformation in their content?

**Katy Minshall:** Yes, that is correct.

Q65 **Clive Efford:** The recent incident of people with bogus information about 5G—Amanda Holden, Woody Harrelson famously retweeted that nonsense. Did you consider at any time that you might shut down accounts like that? Woody Harrelson has 6 million followers, I think, and Amanda Holden nearly 2 million. Would you consider shutting down



## HOUSE OF COMMONS

accounts that big?

**Katy Minshall:** If any account breaks our rules in a way that warrants suspension, that account is going to be considered in that way, whether it is verified or not.

When it comes to some of the accounts that have been mentioned by yourself and others this morning, we have absolutely taken action, whether that is reducing the visibility of tweets or removing tweets across a range of Covid-19 harmful misinformation and suspending accounts as well.

Q66 **Clive Efford:** Rather than shut down the whole account that you can target a particular tweet from that account and make sure that fewer people see it; have I understood that right or do I have that wrong?

**Katy Minshall:** That is correct.

Q67 **Clive Efford:** That is correct. In those instances, when it came to retweets about misinformation around, for instance, 5G or injecting bleach or imbibing bleach, did you take that action against those sorts of bits of misinformation?

**Katy Minshall:** That is correct. I can provide the Committee with some information on our website about this. It predates the Covid-19 pandemic but it is something like #injectdisinfectant, and we blocked it from trending. Just this weekend we blocked a website and marked it as unsafe because it was harmful.

The key thing here though is—just thinking about what the Culture Secretary said when he was here last week—the single best thing we can do is drive reliance on reliable narratives, and we have been heartened to see the level of engagement with the reliable narratives that we are curating on Twitter. Something at the top of all UK users phone timelines right now there is what we call “a moment” and it is basically an image that takes up the top fifth of your mobile screen. If you click on it you see all the latest updates from a small number of authoritative UK accounts, like Public Health England and DHSC. The use of moments has increased by 45% and so what we are seeing is people are responding as we elevate and amplify authoritative information on Twitter.

Q68 **Clive Efford:** With that in mind, do you often take action against tweets from the President of the United States?

**Katy Minshall:** We have taken action on tweets that break our rules that absolutely include world leaders. If any tweet breaks the rules we will take appropriate action. It is probably also important to say that we do have—

Q69 **Chair:** Kate, sorry, you say “world leaders” so I know there is one in Brazil. Have you ever reduced Donald Trump’s tweets at all, reduced the profile of them? Have you taken any action at all against the President of the United States or any other world leader apart from President



Bolsonaro?

**Katy Minshall:** We have taken action against other world leaders around the globe, particularly in the past few weeks, when it comes to Covid-19 misinformation. But what is important for the Committee to also be aware of is that we do have a policy where there may come a time where it is in the public interest to be aware that a world leader has shared a tweet, but if it does break our rules, we will hide it with an interstitial that notes that it is harmful and does break the rules.

**Chair:** I did not detect an answer there. Clive, is there anything else you wanted to say?

**Clive Efford:** No, it is fine.

**Chair:** Thank you for your evidence, Katy Minshall. We will now move on to our second witness, who is Richard Earley, the UK Public Policy Manager at Facebook. The first member I am going to call is Kevin Brennan to ask his questions.

Q70 **Kevin Brennan:** Thank you for coming to represent Facebook, Mr Earley, today. Have you discussed the issues that we are looking at this morning with Mark Zuckerberg?

**Richard Earley:** I have not personally discussed them with Mark Zuckerberg, no.

Q71 **Kevin Brennan:** Have you ever met Mr Zuckerberg?

**Richard Earley:** I have seen him. I have not personally met him, no.

Q72 **Kevin Brennan:** We do appreciate you appearing before the Committee today but to be absolutely clear in relation to Facebook, how would you describe yourself in relation to being on the rungs of the ladder of the Facebook organisation?

**Richard Earley:** I am here today to talk to the Committee—and thanks for inviting us—because my role in the UK team is working directly with the UK Government, Members of Parliament, such as yourselves, and the civil service on issues around misinformation and health.

Q73 **Kevin Brennan:** In relation to the organisation, the question I was asking: on the rungs of the ladder how far above you would Mr Zuckerberg be in terms of layers of seniority?

**Richard Earley:** I am not sure I can give a very accurate answer to that question because there are a lot of layers at Facebook, but my role is speaking directly to the UK Government and where what I hear from the UK Government is important, that will be conveyed back to other people in the organisation by me or by—

Q74 **Kevin Brennan:** Have you discussed any of this with Nick Clegg in person?



## HOUSE OF COMMONS

**Richard Earley:** Nick Clegg works in our headquarters in California so I do not see him personally. I know that he has taken a considerable role in these issues. He is obviously our vice-president in charge of communications and policies, so that is my organisation. A number of the steps that I can talk about, which we have taken, have originated from our leaders in California being influenced by the conversations we have had with authorities like the NHS in the UK and elsewhere in the world.

Q75 **Kevin Brennan:** If you do bump into them at any time please do remind them about their past invitations from this Committee and also the invitation for Mr Zuckerberg to appear before the International Grand Committee because there is a feeling abroad that sometimes Facebook feels it is a bit too big for parliamentary scrutiny even at that international level. Do you know whether or not Mr Zuckerberg has met with the Prime Minister to discuss this or any other issues?

**Richard Earley:** I do not believe Mark Zuckerberg has met with the Prime Minister, but I did want to comment on your point about Facebook's appearance before Committees. I fully recognise, and the company recognises, the importance of the issues that we are talking about today and indeed, the value of inquiries like the one this Committee and other Committees have been running. As a company, we have always appeared before Committees of this Parliament when we have been invited and I am here today because of my role working on the specific topics at hand.

Q76 **Kevin Brennan:** We are grateful for you coming along. Do you agree that Facebook is the platform of choice for social media manipulation?

**Richard Earley:** I would not recognise that statement.

Q77 **Kevin Brennan:** It is from Professor Howard, who was our first witness, from his 2019 report on global disinformation. It had five key findings and that was one of them so I am surprised you do not recognise it.

**Richard Earley:** I have not had a chance to catch up on this morning's evidence but certainly—

**Kevin Brennan:** Sorry, just to stop you there. It is not from this morning's evidence, it is from the 2019 report on global disinformation that Professor Howard, one of our witness's earlier on, wrote rather than from this morning's evidence. You were not aware that that was one of the five key findings?

**Richard Earley:** I certainly recognise that there is misinformation on our platform and disinformation. We see this manifesting in a variety of different ways, from a variety of different sources, and the approach that we take leads back to the question of who is the correct authority to decide what is misinformation and what is not. I can talk about the ways that we approach the content we see on our platform but certainly I would not—



## HOUSE OF COMMONS

Q78 **Kevin Brennan:** We can come to that in a minute. But you would not agree; he has it wrong when he says that Facebook is the platform of choice for social media manipulation.

**Richard Earley:** I am certainly not an expert in this kind of research to be able to comment on a comment like that but there is very clear evidence we are one of the largest social media platforms in the world.

Q79 **Kevin Brennan:** Can I commend his report to you anyway perhaps to catch up on it?

There is another report that came out recently. Are you familiar with this one from Avaaz, the non-profit organisation that monitors the internet entitled "How Facebook can flatten the curve of the coronavirus infection"? That came out two weeks ago. It sampled 104 pieces of coronavirus misinformation on Facebook and found they were shared 1.7 million times with 117 million estimated views. Obviously this is just a tiny sample of misinformation—just a tip of the iceberg—but Avaaz found it had been viewed more times than Facebook's own Covid-19 information centre. Is Avaaz not absolutely right when it concludes that Facebook is the epicentre of coronavirus misinformation?

**Richard Earley:** In this case I am familiar with Avaaz, and they are a partner of ours that we worked with in a number of context, including in connection with the European—

**Kevin Brennan:** Are you surprised as partners they describe you as the "epicentre of coronavirus misinformation"?

**Richard Earley:** I would expect our partners to challenge us on the work we are doing where they see fit. I am grateful that they do so. I would say—and I want to be very up front with the Committee—that producing reliable accurate information on these kinds of very fast-moving topics is a challenge. I am not able to give statistics that would counter or mirror the ones that Avaaz has provided there. What I can talk to the Committee about is the actions that we have taken on misinformation where we find it on our platform.

Q80 **Kevin Brennan:** We can come on to that in a moment but there are a couple of other points I want to press you on before we move on because we are quite short of time. The Competition and Markets Authority's interim report on online platforms shows that Facebook make 50% returns on digital advertising, which it concludes is consistent with the exploitation of market power. In effect, you are using the less regulated online environment to make money out of advertisements that are around this misinformation that is being described by these reports, around fake news, while Ofcom regulated news outlets cannot compete on that unlevel playing field. Why are you not spending more, given these massive returns and the profits that the organisation makes, on promoting reliable news and on compliance measures as others in the news business have to?



**Richard Earley:** We are spending an enormous amount of money on tackling the issues of not just misinformation but also broader safety and security issues as they appear on our platform. Last year we spent more money on safety and security than was the revenue of our company at the time we went public in 2012. We now have more than 35,000 people working at Facebook across various teams, not just content reviewers but also looking into threats where they emerge, developing new policies to help tackle problems like this. That investment is reflected in some of the steps we have been able to take to counter just some of the issues that you mention.

On the question of advertisements, Facebook has taken a number of steps to increase transparency and control around advertisements, which go further than many other platforms and many other types of media. In the UK, for instance, it is not possible now to run an advert on a political or social issue, including health, including coronavirus, without verifying your identity with us, confirming that you are based in the UK and also providing a disclaimer that shows where the money behind the advert is coming from and also providing—

Q81 **Kevin Brennan:** Can I just take you up on that point about political advertising? A constituent of mine sent me an ad recently that he complained about that was misleading and political and it was about Covid-19. It took three days to take it down after his complaint but he wanted to know how many people it had reached. He wanted to know how much was spent on it. He wanted to know who had paid for the advert, but he could not find any of that information. Do you publish that kind of information where there is an ad that is political in nature, with misinformation, and that is taken down?

**Richard Earley:** Our ad library, as the Committee might be aware, contains for seven years a directory of all the ads that have run on our platforms that touch on these issues, including those that you have mentioned. Naturally, due to the adversarial nature of the space we find ourselves in, sometimes people are able to get round our systems—our human reviewers or our automated systems—and content can appear on the platform for a short time. I am glad that it was taken down.

My understanding is that where ads that are related to political or social issues run on the platform without a disclaimer we leave them in the ad library but I am obviously not familiar with that specific case.

Q82 **Kevin Brennan:** So my constituent should be able to get hold of that information by looking it up in the ad library?

**Richard Earley:** I would be very happy to check on that particular case because there are different steps we can take towards different types of advertising.

Q83 **Kevin Brennan:** I will send you the details.

Finally, because time is short, do you support the establishment in law of



## HOUSE OF COMMONS

a duty of care for Facebook in relation to what it publishes on its platforms?

**Richard Earley:** Safety and security is a paramount concern for Facebook.

**Kevin Brennan:** I accept that but just answer the question: do you support the establishment in law of a duty of care?

**Richard Earley:** We have been very public about our belief that further regulation from Governments, including in the UK, around the areas that we have just been discussing would be helpful.

**Kevin Brennan:** That is a yes, is it? You do support in principle a duty of care?

**Richard Earley:** With regard to the specific term “duty of care”, this has a number of certain legal meanings in English common law at the moment and also in the context that the Government have described in connection with the online harms paper. As we have not yet seen the online harms Bill, which would give further details of how this duty would operate, I am not able to comment on it right now, but certainly there are a number of aspects of the UK Government’s proposals around online harms that we support.

**Kevin Brennan:** But you are not emphatically against the idea of a duty of care, in principle?

**Richard Earley:** It very much depends on how it is framed. There are a number of complex issues related to how you —

**Kevin Brennan:** I can see I am not going to get an answer to that so I will pass back to you, Chair.

Q84 **Steve Brine:** I am astonished you have never met Mr Zuckerberg or Mr Clegg. You are the UK public policy manager for Facebook. I would have thought that would give you much greater credibility when talking to Parliament or to leaders. Maybe it is something for your diary secretary.

On the 5G issue, I just put “5G” into Facebook while I was listening to the last bit of evidence and here is a post from within my own constituency—and it is a matter of public record, clearly, where I represent—which says this month, “Any 5G masts where we live? I need to go out and take them out. They cause the virus, you know.” It is still up there. That was 26 days ago. That seems to be inciting a group to go and cause criminal damage; it is still up there on Facebook. The early evidence we heard from Professor Howard talking that Facebook was a host to, and rife with misinformation; maybe he was right, wasn’t he?

**Richard Earley:** Obviously I cannot see the content that you are referring to—

**Steve Brine:** Put “5G” into Facebook and my constituency and you will see it right there. It is all over it.





**Richard Earley:** There are a number of things in what you describe that I wanted to pick up on.

First, even before we started to remove content from Facebook, which falsely linked 5G technology with the symptoms of coronavirus, it would already be against our policy to incite or encourage people to commit acts of violence against businesses, including supposed 5G masts. But on the specific claim around 5G causing coronavirus, there is a slightly different approach that we take towards misinformation in general on Facebook and misinformation that can cause real world harm.

When it comes to general misinformation, our usual approach is that we tend to agree with those who say that it should not be for companies like Facebook to decide what is and is not true. Therefore we have, since 2017, built a network of third-party fact-checking partners around the world. We have three in the UK, Full Fact, FactCheckNI—Northern Ireland—and just earlier this month we added Reuters to that list. They are able to act on instances of content that they feel are misleading, they issue a rating on that, where they choose to do so, and then we take action to show it to fewer people.

There is a separate approach that we take towards what we fully recognise as some types of misinformation that can, if they are believed, lead to real world harm. We already had policies about this before coronavirus but in the context of the coronavirus pandemic, we have been applying those policies in connection with misinformation related to the coronavirus where if it was believed, might lead people into harm.

In connection with 5G, it is fair to say that just a few months ago no one would have predicted that misinformation around 5G and coronavirus would have had the potential to cause real world harm but we know that the installation of 5G has been controversial for a long time. The installation of previous telecommunications infrastructure like 4G and 3G produced similar sorts of opposition at the time they were rolled out.

Q85 **Steve Brine:** Yes, It just did not have any basis in fact. I am going to stop you there, because time is so short.

You have taken action against groups that are spreading misinformation, and I appreciate that on 5G. You have taken action against people who are organising events to undermine social distancing measures through the event creator on Facebook. How do you identify those malign actors, and what sort of transparency do you provide to them when they are having their content removed or does it just disappear? Take the lid off and explain to us, very briefly, how that works.

**Richard Earley:** I think the 5G case is a very instructive one because we started to receive reports from our work with Government, the media, NGO partners and also as flagged by our users about this sort of content. As I said, we believe that it is not Facebook's place to decide what is and is not true, so we worked with international and national health



## HOUSE OF COMMONS

organisations, including in the UK here, to come to the decision that we should remove this sort of 5G content. We started then removing it on the basis of where it was flagged to us by users or where others flagged it to us.

Usually when we start to take down a piece of content we start applying our policies in a certain way. We in some cases try to develop technology that can help us to proactively surface that content. This is really challenging, it is really cutting-edge technology that we have been able to build in other contexts and were able to apply to the 5G situation.

I think we started taking this content down on 12 April and for about 10 days or so we have been proactively finding and removing this content; as you mentioned, both content that makes these claims as well as groups that are dedicated to spreading this kind of harmful misinformation.

There is an additional step we needed to take, which is to make sure that if people repeatedly share this kind of misinformation—which is, as you said, harmful—we can take action against them. When somebody breaks our rules for what is and is not allowed on Facebook—our community standards, which include this line about not sharing harmful misinformation—and they repeatedly do so, we can take action against them. In order to do that, we need to make sure we are transparent with our users about what kind of content is against our rules, because naturally it would be inappropriate of us to remove people from the platform and deny them the ability to use Facebook and our other platforms.

**Q86 Steve Brine:** Mr Earley, does your technology allow you to stop them resurfacing under a different name—for instance, by tracking IP addresses? You must have usual suspects who probably continually attempt to resurface and spread misinformation—as briefly as you can.

**Richard Earley:** Once someone has been warned appropriately and either banned or removed permanently from our platform, we do have systems in place to attempt to prevent recidivism—people coming back. I also fully accept this is very challenging. There is a limit to the amount of information that we can see and there are also people in this space who will continue to try to get around our rules. However, we remove millions of fake accounts every day through the detection we are able to carry out and we apply that to actors who we know are trying to circumvent our rules as well.

**Q87 Damian Hinds:** Mr Earley, I would like to start with a step away from disinformation for a moment following a report a few days ago from the NSPCC about the expected increase, sadly, in child abuse during this period and reports that there was a lack of moderators at Facebook due to staff shortages. Can I ask you what you have done already or what you will today commit to do in terms of redeploying people from other roles in the organisation to make sure that that highest of priorities can



be executed fully?

**Richard Earley:** Thank you. I completely agree that this type of content and this type of report are absolutely the most serious kinds of content we have to deal with at Facebook. It is our utmost priority to make sure we are able to handle these reports and act where we find this kind of content.

On the specific question around moderation and our community reviewers, I mentioned we have around 35,000 people at Facebook working on safety and security, and around half of those directly review content. In the extraordinary circumstances that we find ourselves with the coronavirus, we took the decision in March to send all our content reviewers home in order to protect their own safety. We have to balance here the need to ensure the safety of users in our community with our obligations and our responsibilities towards our staff. When we did that we took a number of steps to make sure that we minimised any negative impact on our ability to review content. The first piece of that was that we shifted responsibility for those most serious types of violation—that includes child exploitation material which you mentioned here, as well as things like suicide and self-injury material—to our full-time employees so that those queues will continue to be managed. We also had a large number of employees at Facebook who do not even work to review content in their daily roles volunteering to step forward to help make sure that we did not see significant negative impacts on that queue.

Secondly, we started to enable our content reviewers, who are contractors, to work from home. Obviously we need to put in place a number of security and privacy checks before we can enable this to happen but currently the majority of our content reviewers are now able to work from home and are reviewing content in that way.

Lastly, in the last week we announced that due to the evolving situation and the recognition that we are going to be in this “new normal” for a long time, we have started to open some of our review centres and allow contractors to return where they express a desire to do so. That is fully voluntary for the contractors. It will also be accompanied by new steps to ensure we protect their safety, including reduced building capacity and also ensuring that they have access to protective equipment.

One final thing I wanted to point out was that these are the ways we have ensured we can minimise the impact on our moderating team—our review teams. However, we also at the same time took some steps to increase the reliance that we have on automated ways of flagging and removing content. As you might know from our transparency report, we already remove the vast majority of problematic content that we find on our platforms through finding it ourselves. In the case of the child exploitative material you are talking about that is well above 99% and has been for a number of years.



## HOUSE OF COMMONS

The combination of using our technology, supporting our reviewers and reprioritising work where we can has been our approach to minimise the impact on this.

**Q88** **Damian Hinds:** I do not want to oversimplify the question, Mr Earley, but do you today have the same number of full-time-equivalent staff working on tackling images of child abuse and related matters? Do you have more, or do you have fewer?

**Richard Earley:** I cannot give you a specific answer on that because the amount, as I have just described, is changing all the time in terms of shifting the responsibility to full-time equivalents, bringing in additional staff, bringing online capacity at homes and—

**Q89** **Damian Hinds:** Can I ask that we follow up on this by letter and that you commit to what you are going to do so that people will not have to see these pop-up messages saying that there is a shortage of moderators available and you are having to restrict the amount of activity?

**Richard Earley:** I will be very happy to do that. We recognise that this situation is going to have an effect on the speed with which we process some types of content and that is why we are showing these messages to people, to make sure we are transparent with them about the impact that is having and how we are handling it.

**Q90** **Damian Hinds:** Thank you. May I come back now to disinformation and misinformation? Is it not the case that novelty sells and that Facebook's business model is based on user engagement? Set aside Covid-19 for the moment. We know that incorrect news travels further and faster than correct news. While I am sure that you personally and your colleagues would like to see a reduction in fake news and disinformation—you may have a moral imperative and moral motivation to do that—you do not have a commercial motivation to reduce disinformation and misinformation.

**Richard Earley:** I would challenge that. The way Facebook presents information to people when they log in is based on a series of algorithms. In 2018 we completely overhauled the way we select and choose information that we show to people based on a large number of—

**Q91** **Damian Hinds:** It is true, is it not, that novelty attracts attention and fake news will tend to travel further and faster because of its novelty value, because of its shock factor and because of the disgust sometimes it can generate?

**Richard Earley:** The way we choose content to show to people is based on what we think is the most likely to trigger—

**Damian Hinds:** What they engage with.

**Richard Earley:** —a rather jargony term, which is meaningful social interaction. In essence what that means is content we think people are



most likely to engage with. That is one part of the way we show people content.

In addition to that, I have spoken briefly before about the way we work with third-party fact checkers who are able to review what is flagged to them either by our users or by our proactive systems. Where they find and rate content as false we apply a demotion, so to speak, to that content so it is shown much lower in the newsfeeds of people who might have access to it. We also cover that false news, where it is found, with a covering—I cannot think of the word now—that includes a link to some correct information provided by the fact checkers. We think that is the right approach.

**Q92 Damian Hinds:** Yes, but that is by the time it has been corrected. Is it not true that in the meantime information that is novel, new, conflicts with established norms will tend to attract much more of what you describe rightly as the engagement of your users? Therefore it travels further and you have a restricted, if indeed it exists at all, commercial incentive to reduce the prevalence of disinformation and misinformation.

**Richard Earley:** I certainly do not want to deny the fact that we know misinformation exists on our platform. We are taking action against it in the context of our third-party fact-checking programme. In addition to that, a lot of the evidence and a lot of the discussion around misinformation that is being shared suggests that in many cases the most impactful step we can take to tackle misinformation is to connect people with authoritative true information. That has been a huge focus of the company and a huge focus of my work here. We already show people who search for coronavirus-related terms on Facebook links to the latest government advice. At the very top of every single person's newsfeed, so above any other content, we have a link to our information centre that links directly to the latest government advice, statistics and guidance from international health organisations. We have connected more than 2 billion people worldwide to that information, with over 350 million of them clicking through.

**Q93 Damian Hinds:** Forgive me, I know we are short of time. Not just thinking about the present crisis but more generally, of all the people who see something that is dis-informing—a piece of fake news—what proportion of them will go on to see a correction?

**Richard Earley:** As I outlined at the start of my evidence, it is very challenging for us to provide accurate meaningful figures in this area. What I can tell you is that we know that when people are shown a—

**Damian Hinds:** Give us a rough guess.

**Richard Earley:** —piece of misinformation with that content covering on it we know that 95% of them do not go on to click through it.

Also, earlier this month as you may already know, when we apply one of those—



**Q94 Damian Hinds:** Sorry, Mr Earley, forgive me. That was not my question. My question was for the number of people who see a piece of disinformation on your or another social media platform, but particularly yours, what is your estimate of the proportion of those who actually go on to see something that corrects it?

**Richard Earley:** Any person who has interacted with a piece of content that our fact checkers subsequently rate as false will receive a notification from us linking them to the debunk provided by the fact checker. That is the case for misinformation that we apply these actions to.

When it comes to misinformation that we remove for being harmful, again in the context of coronavirus just this month we have now started to show new messages to people who have seen harmful misinformation that has been removed from our platform directing them to the latest collection of myth-busting resources that is available at the World Health Organisation.

Again, when it comes to the question of balance I accept that this information is travelling on our platform but at the same time we are taking these steps to show people authoritative information. This is not just within Facebook. Within our company in the UK we have partnered with Public Health England to launch a WhatsApp coronavirus bot where people are able to seek the latest statistics, information and myth-busting guidance from Government. We know that bot has already sent over a million messages since it was created.

Between that, our work to connect people who search for coronavirus information with the latest information, and also with the large amounts of free advertising credit that we have given to the Government to enable them to directly reach the millions of people on Facebook, we are working in a number of different areas to make sure people can find authoritative information out there as well.

**Q95 Damian Hinds:** My final question, Mr Earley, is sort of a repeat of my earlier question. I realise you are doing all these things; I realise you are working with fact checkers and I realise in the current crisis you are promoting good, solid, positive public health information. My question is, for everybody who sees bad information how effective are all these things that you are doing? What do you know about the proportion who then go on to see that bad information being debunked, being corrected?

**Richard Earley:** As I said, in the context of bad content that our third-party reviewers have gone on to rate we not only show information about the content to people who still find it despite our actions to show it lower in people's newsfeeds, we also show this additional notification to people who have interacted with it previously. We are taking the same steps with—

**Q96 Damian Hinds:** Okay. And on WhatsApp?



**Richard Earley:** WhatsApp is a very different sort of service because it is a private communication system, much like text messaging or e-mails. In the context of WhatsApp there is a difference between the people who are using text messaging and e-mail to spread disinformation now and how you can do that on WhatsApp. Two key areas are first of all that we take significant action on WhatsApp to prevent people from sending bulk messages, and we ban millions of fake accounts a day on WhatsApp that are engaged in that kind of behaviour. Secondly, really important, in the last week we have extended the limitations that we place on the ability of people to forward messages multiple times in WhatsApp. We have now reduced the number of times that a person can forward a message they have received that has previously been sent to them through five different people. We have dropped the limit of how many people they can forward that to from five to one. Our initial research suggests that has reduced the number of those messages being sent by 70%, which is a huge impact.

Q97 **Damian Hinds:** To be clear, specifically on WhatsApp, you do not know what the reach is of these false information messages and you do not know what proportion of people who see those false information messages then get correct information to replace it?

**Richard Earley:** WhatsApp is a private messaging service, as I have said. However, we are able to use the metadata we have about WhatsApp to apply these labels so when people are seeing messages that have been forwarded more than five times before they receive it there is an indication, a double arrow, which shows this is multiply-forwarded information. The purpose behind that is to give people additional context and understanding of what they are seeing.

Certainly, as we are aware, there is misinformation being spread on WhatsApp. However, people are also using WhatsApp to connect with the latest information. That is not just through the WhatsApp NHS bot that I described, but we have a number of resources that we have created with the World Health Organization. We have created a WhatsApp information hub that gives tools, tips and advice for people as to how to spot false news, and we continue to work with Governments on those kinds of projects.

**Chair:** Thank you, Mr Earley. I did not particularly notice any facts in the answer to Damian Green's question. An observation here: as with other witnesses here today it does seem to me, frankly, as if none of you have supplied what I would call real, genuine, hard information about how you are specifically going about tackling Covid disinformation.

I will move on now to our third witness and hopeful to have more luck here. Our third witness is Alina Dimofte, the public policy and government relations manager at Google. I will hand over our first question to Damian Green.

Q98 **Damian Green:** Thank you, Chairman. Good morning. I want to ask you



## HOUSE OF COMMONS

some questions about your response to people complaining about fake news, misinformation—whatever you want to call it.

First of all, in terms of what you do proactively, does Google employ fact checkers or the equivalent?

**Alina Dimofte:** Thank you for that question. When it comes to this pandemic and tackling misinformation across our platforms we have been employing both proactive and, of course, reactive measures. When you look at YouTube, 90% of the videos that are related to Covid misinformation that we take down have been detected by our automated systems. That is not us waiting for someone to flag issues to us, but our investment in machine learning paying off and us being able to proactively identify and then take action against this material.

In relation to fact checkers, we are working in partnership with fact checkers. You heard from Dr Wardle earlier from First Draft News. We are one of their founding partners, so we have been working with them and funding fact-checker work across the world for many years. We know that during Covid-19, their work streams have increased by a lot, and they have more data that they need to check. That is why we have, on top of the existing support, announced a \$6.5 million fund that is being administered through organisations like Full Fact in the UK and First Draft in Europe. It is aimed at helping to create databases of fact checks, exchanging information between fact checkers and also, very importantly, helping train journalists on where to get the facts and how to partner with fact checkers as well.

Q99 **Damian Green:** Thank you for that. We have heard from one of our witnesses this morning that on YouTube users can flag content as harmful or misleading but you cannot do that on the search engine. Will you introduce that?

**Alina Dimofte:** We have introduced methods on search that ensure the quality of our ranking systems. I think when we are looking at social media versus search engines, these are very different systems but of course that does not mean that we should not take responsibility, we should get outside feedback on how our ranking systems work and whether users receive the best information available from authoritative sources. On search we do this through our community of 10,000 raters who are placed around the world and who look at search results, key words in search results, analyse the effectiveness of our ranking algorithms and give us feedback on how we can improve. We also conduct tests ourselves and roll out hundreds of improvements every single year. We know there is always more that we can do. We have updated our ranking very often. Where we can get external input into how we can serve our users better, we will do that.

Q100 **Damian Green:** You say you have 10,000 people. That is not quite the same as if somebody flags something as probably misinformation. What systems do you have for, as you say, dropping it down your ranking so





that effectively nobody will find it?

**Alina Dimofte:** Our systems work at scale and we do exactly what you are describing there. We make sure that when it comes to searches that are related to Covid-19, but also generally about medical disinformation or any newsworthy searches for that matter, we do our best to surface authoritative rankings.

What we have done in particular in response to this pandemic is to work with Governments and health authorities such as the NHS to create new experiences on search. There have been taskforces inside Google that have been building products from scratch for this. When you go to Google now and search for anything that is coronavirus/Covid-19 related you will be served an experience that has been developed in partnership with the NHS here that gives information about treatment, symptoms, prevention, even infection rates. We are absolutely committed to surfacing authoritative information because the No. 1 ask we have heard from Governments and from health organisations is that people need clear information—concise information for that matter—from one or two single sources about what to do. What we want to do is to empower people with this information. In the UK, we are working with gov.uk and with the NHS as the primary sources of information.

I heard the Committee ask earlier about how we are tipping the balance away from misinformation and onto these trusted sources, and I can share some facts.

For example, when you look at YouTube we are surfacing—for any search related to coronavirus—we are surfacing panels with NHS information. Those panels have been viewed over 40 million times in the UK and over 20 billion times globally. Of course we localise the panels so you might have WHO information or information from your local health organisation.

Q101 **Damian Green:** Thank you. YouTube is the only platform that does not have a publicly available API. Would not making that publicly available help researchers and help everyone know the scale of the problem?

**Alina Dimofte:** We are an open platform. Anyone can build APIs to integrate with YouTube and test the efficacy of our measures. For example, Oxford university has done a piece of research looking specifically into how we have responded to Covid-19 and how well our recommendation and search functions work on YouTube. I think Professor Howard was part of the research. What they found is that four-fifths of all the information that is surfaced on YouTube is coming from either news organisations or from Government and health organisations. We have indeed seen, and our own numbers show, that there has been, a huge increase in the consumption of news from authoritative sources. In the UK specifically we have seen a 65% increase in the consumption of publishers like *The Guardian*, *The Telegraph*, *BBC* and so on. *The Telegraph* and *The Guardian* have, for the first time, surpassed 1 million subscribers on YouTube.



## HOUSE OF COMMONS

Of course the NHS is an important source of information. We have done the panels, which I have explained, but we have also worked with them to produce content for YouTube. The NHS had some videos for YouTube but was not spending a lot of time on our platform previously. It is doing more, of course, and we want to support that. That is why on our homepage, when people open YouTube, we have promoted the NHS videos. That is why we have an increase of the views of these videos from about 6,000 views to over 1.8 million.

There really is a lot that we can do through partnership but, to the core of your question, we understand the value of openness and the value of having academics, researchers and fact checkers look into the efficiency of our tools and we want to learn from that research.

Q102 **Damian Green:** You are saying your API is available to researchers?

**Alina Dimofte:** There are integrations that people can do with our API. There is definitely more we can do and where fact checkers are telling us they need more tools, we will look at those requests and try to accommodate them.

**Damian Green:** Thank you.

Q103 **Philip Davies:** Research from the Global Disinformation Index—which is an independent not-for-profit organisation working with Governments, businesses and civil society—shows that Google is funding nearly 90% of coronavirus conspiracy sites by serving the advertising. You presumably also profit from selling this advertising. Is this not remarkably reckless? Why do you not check the sites on which you are serving advertising before you publish?

**Alina Dimofte:** This is a very important issue and thank you for raising it. Whenever we find these types of websites that are peddling misinformation or disinformation, we take quick action. When it comes to Covid-19 we have updated our policies to ensure that our advertising does not run on any websites that are peddling medical misinformation or are contesting the existence of the virus or the treatments that have been described by authorities like the NHS and WHO.

Q104 **Philip Davies:** Do you accept the information from the research from the Global Disinformation Index? Do you accept what they are saying is right and that you are responsible for a huge proportion of these sites by serving the advertising?

**Alina Dimofte:** I have to apologise. I do not know this piece of research and the methodology behind it. I can commit to looking into it and responding in detail.

However, what I can reassure you now is that for our platform that allows publishers to monetise we have specific misinformation and Covid-19 related policies that we have been enforcing rigorously. We have taken action to demonetise thousands of these pages. Whenever we see this we



## HOUSE OF COMMONS

take quick action and we under no circumstance want to profit from this type of information.

**Philip Davies:** Chairman, in the interests of time can I therefore ask that Google look at this report and respond to us in writing? On that basis perhaps we can move on to the other questions that need to be asked by other people.

**Chair:** Yes, thank you. We are obviously interested to see exactly what happens in terms of the money because it is all very good saying that you demonetise them, but what do you actually do? What happens to the money before you demonetise them? Do they just get to keep it? We can deal with this in writing.

Q105 **John Nicolson:** I would like to move on to the issue of problem gambling during the lockdown. There has been an explosion in problem gambling with people moving from sports to high-risk gambling such as poker. Do you think it is appropriate for Google to continue to serve gambling advertisements during this lockdown period?

**Alina Dimofte:** We have policies around the type of gambling advertising that is allowed and not allowed on our platforms. I can definitely go back to the teams working on this area specifically to see whether they have seen any issues and any trends related to the pandemic and whether more action can be taken. We will definitely take that on board and respond in detail. I am afraid I am not an expert in this area.

Q106 **John Nicolson:** I am glad to hear it because it is a moral question, I think.

Can you explain whether or not you target people with gambling problems or who are prone to gambling regularly with bespoke advertisements?

**Alina Dimofte:** Again, I am not really an expert in the gambling area. I am definitely able to reassure you that we will never want to target the vulnerable, those are the exact audiences that you want to protect when it comes to this kind of advertising.

Again, I would ask that I respond with more detail and be more precise in writing.

**John Nicolson:** Thank you. The Committee will publish that, I imagine, when you do respond.

**Alina Dimofte:** Of course.

**Chair:** Thank you, John. I have to say the number of letters we will be writing is increasing because you have come here today, all three so far, seemingly unable to answer quite basic questions. We will have one final try, which is going to be Clive Efford.

Q107 **Clive Efford:** Thank you, Chair. On Tuesday 7 April, the BBC reported that YouTube has updated its policies to delete 5G conspiracy videos in



## HOUSE OF COMMONS

response to the livestream of the David Icke interview that was watched by 65,000 people. The BBC reports that YouTube was aware of the video at the time it was being streamed. If YouTube was aware why did it not take action at the time? Why was action not taken sooner?

**Alina Dimofte:** When it comes to our policies, we are always looking at how we can evolve and adapt the policies, especially in an unprecedented time like this pandemic.

With the David Icke video, it was the first instance that we have seen where these kinds of 5G allegations were creating real-world harm and were being linked to the coronavirus in particular. That is when we took quick action to update our policies going forward and since then we have enforced them robustly to ensure that any kind of similar content has been removed. We have also made sure that we have changed our algorithms to reduce the propagation of other conspiracy theories. We, of course, make sure that every time a user comes to YouTube—whether they are searching for Covid-19 related keywords or they are watching a video, including that livestream—they would have been served with an information box underneath linking them to official information from the NHS.

Yes, we definitely understand we need to take quick action against content that disputes the existence of the virus or the spread of the virus as it is described by the NHS and the World Health Organization. We will take action quickly. From the videos that we have removed we have seen that 85% of those videos had less than 100 views. Therefore ever since we have updated our policy we have been able to take quick action and catch this harmful misinformation before it spreads, before it reaches a mass.

I was talking earlier about tipping the balance. You need to compare action taken on removing content with less than 100 views and then pushing NHS content—NHS information that has been viewed by tens of millions in the UK.

Q108 **Clive Efford:** Are you saying that the conspiracy theory around 5G did not exist until David Icke gave that livestream interview?

**Alina Dimofte:** No, we have definitely seen trends around 5G misinformation. What I was trying to explain is the link between 5G and real-world harm, like the attack on mosques that we have seen here in the UK where that type of misinformation evolved. We are very aware that this is not a solved problem. There is no silver bullet when it comes to misinformation. We all need to keep updating the way we tackle this, be humble in our approach and learn from where we need to do better.

Q109 **Clive Efford:** The thrust of the question is why was action not taken immediately, given that YouTube knew at the time of the livestream of the content? Why was action not taken immediately if you were already aware of the conspiracy theories around 5G? The question is: why the



delay?

**Alina Dimofte:** I understand your question. At the time that we first viewed the video it was not against the policies we had at that time. That is why we understood that our policies needed to evolve. Other platforms then followed suit and started following our lead in having more precise policies and removing more real-world harm content from their platforms.

**Clive Efford:** Sorry, did you want to carry on?

**Alina Dimofte:** I wanted to say very quickly that this is only part of our response. When it comes to YouTube, our response is around removing the videos, reducing the availability of borderline content and, of course, surfacing authoritative information together with partners, like the NHS, fact checkers and established news partners.

Q110 **Clive Efford:** YouTube gave the money that it made from advertising on that video to charity. However, it allowed the company to keep the money it made from the video's Super Chats. Why was action not taken against that? Is it possible for YouTube to take action of that kind, and if so why did it not do it?

**Alina Dimofte:** If I may, I should correct here, we have given to charity all the Super Chat revenue as well. We do not want, under any circumstances, to make money out of this type of content. Our business model is one built on trust. For 20 years Google has tried to organise the world's information and make sure that users are connected with information that is helpful and trustworthy. We know that we are dependent on users trusting us in finding the right information on our platforms. We absolutely do not want to make money out of misinformation.

Q111 **Clive Efford:** The question was specifically about the hosts making money out of the Super Chats. They were allowed to keep that money, is that correct?

**Alina Dimofte:** When we see repeated infringements, like we have seen in the case of the channel that you are mentioning, we are taking away the ability to monetise. Therefore this host will not be able to make money going forward.

Q112 **Chair:** Did they keep the money, yes or no?

**Alina Dimofte:** I will need to look into that specific detail.

Q113 **Clive Efford:** You see the problem is this becomes an incentive for companies to put out disinformation if they, through the controversy and discussion that they generate, are able to keep money from Super Chats. This is an absolutely important issue, is it not? Should YouTube not be closing down that opportunity for people to make money and gain from spreading misinformation?

**Alina Dimofte:** You are absolutely right and that is why a few years ago we changed our approach to monetisation on our platform. Before,



## HOUSE OF COMMONS

everyone would be able to monetise on YouTube. We have changed that approach to look at monetisation as a privilege. You need now to meet certain benchmarks of quality and trustworthiness in order to be included in our YouTube Partner Programme in the first place. When we see channels changing behaviour, which happens, when we see that they are no longer trustworthy—

**Chair:** Thank you. We are going to have to bring the session to an end now, unfortunately, as we have run out of time for broadcast. Thank you today to our three witnesses: Richard Earley, Alina Dimofte and Katy Minshall. We will be writing to all your organisations with a series of questions and, frankly, we will be expressing our displeasure at the lack of the answers we have received today and we will be seeking further clarity.