

Home Affairs Committee

Oral evidence: [Hate Crime and its Violent Consequences](#), HC 683

Tuesday 19 Dec 2017

Ordered by the House of Commons to be published on 19 Dec 2017.

[Watch the meeting](#)

Members present: Yvette Cooper (Chair); Rehman Chishti; Mr Christopher Chope; Stephen Doughty; Preet Kaur Gill; Sarah Jones; Tim Loughton; Stuart C. McDonald; Will Quince; Douglas Ross; Naz Shah.

Questions 1-170

Witnesses

[I](#): Dr Nicklas Berild Lundblad, Vice President Public Policy, Europe, Middle East and Africa, Google; Sinead McSweeney, Vice President Public Policy and Communications, Europe, Middle East and Africa, Twitter; and Simon Milner, Director, Public Policy, Facebook.

Examination of witnesses

Witnesses: Dr Nicklas Berild Lundblad, Vice President Public Policy, Europe, Middle East and Africa, Google; Sinead McSweeney, Vice President Public Policy and Communications, Europe, Middle East and Africa, Twitter; and Simon Milner, Director, Public Policy, Facebook.

Q1 Chair: Can I welcome our witnesses? Thank you for coming, we are very grateful for your time. This session is part of our ongoing inquiry into hate crime. We set out an interim report in May looking at issues around online hate crime in particular. We took evidence from your organisations back in March¹ and we are glad to see you here again. Thank you for coming. Please introduce yourselves with your positions in your organisations and tell us briefly the actions that you have taken since the publication of our report. We are tight on time. We have a lot of issues to cover, but you are welcome to set out in further detail in written evidence other actions that you have taken in the last 12 months. I will start with Mr Milner.

Simon Milner: I am Facebook's policy director for the UK and some other countries in Europe, the Middle East and Africa. I welcome being able to come back to the Committee, having appeared before you in March and on the back of your report on 1 May. Quite a lot has happened since then in our company, both of our own volition but also in concert with the other companies on the panel. I might highlight three, but I am happy to answer questions about them as we go on.

One of the questions you asked me last time was, "How many people do you have working in your community operations team?" I was not able to give you a definitive number, but two days after your report came out we put out our public number and we said we had 4,500 people reviewing reports and content and working in our safety teams. We also said we would be adding 3,000 to that group by the end of 2017. We have actually surpassed that, so we have above 7,500 people now working on reviewing content. Our chief executive, Mark Zuckerberg, recently said that although overall we have around 10,000 people working on safety and security, we will be doubling that by the end of 2018. So that is one highlight that is new since I last appeared before you.

The second initiative I would like to highlight is the launch of the online civil courage initiative for the UK, which was launched in London in June by our COO, Sheryl Sandberg. It is a partnership with the Institute for Strategic Dialogue that involves supporting and working with many different organisations representative of and working in communities

¹ Correction made: the previous oral evidence session with these organisations was mistakenly referred to as having taken place in February, when it was actually held in March. These references have been corrected throughout the transcript.

around the UK to try and encourage really positive uses of Facebook to counter online extremism.

The third thing I would like to highlight is that within the last few weeks we have launched our safety guide designed to keep Muslims safe on Facebook. That involved an event here in the Houses of Parliament. It was something we did in partnership with Imams Online and they are distributing that report. I welcome the fact that Mr Chishti was involved in that launch. We very much appreciated his being there.

There are other things I could talk about, but I am conscious of time.

Sinead McSweeney: I am vice-president of public policy and communications for Twitter in Europe, the Middle East and Africa. I have been with the company for five and a half years. In terms of progress to date, 2017 has been a hugely significant year for Twitter in the area of safety and abuse. I think I have probably seen more happen in the past 12 months than in the cumulative time before that during my time at the company. That started with the CEO stating clearly and publicly, very much emphasising internally, and making it a priority for teams that safety was our No. 1 priority for the year. I know that Nick Pickles touched on some of those issues in March, but since then we have had a steady drumbeat of changes across the product itself, our policies and our processes. In fact, we have had 25 separate changes over the course of the year, and that has resulted in us taking 10 times more action on accounts than we would have done previously. We have introduced more user controls, various quality filters, safe search, and our reporting flows are better. We are looking and constantly working for more transparency in explaining how we develop our policies and why we design the product in certain ways, and we have also been giving updates directly to the people who report to us about what actions have been taken.

Touching on those areas more specifically of interest for the Committee, we have seen huge advances in how we are addressing terrorism and terrorist use of the platform, with 95% of terrorist content now being removed, based on our own proactive measures, using technology and otherwise. You will have seen some policy changes that we announced in late October, early November, and we began enforcing those in recent days. Those deal with violent extremist groups and the use of hateful symbols and hateful imageries on the platform, as well as any statements or actions that seek to glorify violence against protected groups. Those are probably the main issues, but there are others than we can get into in more detail.

Dr Berild Lundblad: I work on policy at Google for Europe, the Middle East, and Africa. I have been with the company since 2007. Since the report came out we have looked closely at the recommendations made by this Committee, and across those recommendations we have tried to make good progress. The first one was about people—being more transparent about how many people we have, and to increase the number of people we have. We said that this summer we had 4,500 people, and we will arrive at 10,000 people by the end of 2018, addressing and taking action

on content. We were also asked to look more into how we can use our own technology, and we have started to deploy machine learning and artificial intelligence to address different kinds of content—everything from comments to violent extremism. We are now at the point where the work that is being deployed by those machine learning technologies represents roughly 180,000 people working 40-hour weeks. It makes people five times more effective when it comes to the final human judgment, because you need to combine humans and machines to make this work. We will keep on developing that to ensure that it becomes more efficient.

This Committee also asked for transparency around how we measure those things, and what metrics and yardsticks we use to hold ourselves responsible. In spring we have committed to publish a first transparency report detailing flags, action taken, turnaround times, and so on. Furthermore, we were asked to review our policies, and we are reviewing them across the horizontal spectrum of our different issues at this point. We are making progress and trying to find new states and ideas for how we can address content, and that obviously tries to come very close to our guidelines.

Finally, a large piece of progress that we have made, which I think is important and should be mentioned, is that together with our colleagues we have founded something called the Global Internet Forum to Counter Terrorism. We believe that the issue of content is ultimately an international issue and it needs to be addressed internationally, and furthermore—this is from a UK-based initiative in March—we believe that this issue goes beyond the three companies before you today. We also need to address the many smaller platforms that are not here today. We need to talk to the start-ups who do not know how to deal with this issue. One key part of what the Global Internet Forum does is outreach to smaller companies and other platforms. This year alone, since the Committee's report came out, we have been addressing 68 of those companies and bringing them into the fold of the Global Internet Forum. We have held four seminars and we are making significant progress around the common technology: a hash database that allows us to share violent extremism and terrorism videos among all the members of the forum in order to try to bring this to an international level as well. There is more, but again I think we will get to that.

Q2 Chair: Thank you very much. I welcome the action you have taken since last year. I want to follow up a few of the issues we raised with you. First, we asked you about whether you were searching proactively for abuse, hate crime and illegal content and we got very mixed responses last time. Will you give a brief answer about what your organisation now proactively searches for, as opposed to reactively taking down things that are reported to you?

Dr Berild Lundblad: We are proactively searching for content in order to identify, for example, pieces of content that violate our policies—that can be all of our policies. We first identify new risks we find, then we have people searching for examples of those risks proactively. We then use those examples—there can be many of them, for example up to 2 million

in the case of violent extremism—to train the machines to achieve scale. There is a clear proactive element.

Chair: Thank you. That is a very different answer from the one we had from YouTube last time, so that is welcome.

Sinead McSweeney: We are increasingly leveraging technology in order to help us do this work. That is more effective with certain types of content than with others. It has proved highly effective, for example, in the area of terrorism, and also child protection. It gets more challenging in the area of abuse, because obviously some of the words can be used within protected groups to each other. So we have to be careful that we are not actually silencing speech—for example, within a Black Lives Matter hashtag, some of the language may be used among the community. We have to be careful that we do not have false positives when we leverage technology. But there are lots of signals, particularly on an open public platform like Twitter, that we can also use proactively, which may not necessarily just look at searching key words.

Simon Milner: The two areas in which we have made most progress this year are around terrorism—we have recently reported that 99% of the content we take down involving ISIS and al-Qaeda comes through our own efforts of finding that content rather than content reported to us—and around nudity and pornography, which are not allowed on our platform. Those are the two main areas we have made progress on this year.

Q3 **Chair:** Thank you. To go into some of the specific areas, we reported to each of you some individual cases we had found that appeared to us to violate your community standards or potentially be illegal. I will take each of those in turn. On Twitter, we sent you a series of tweets including violent threats and abuse towards Jess Philips, Diane Abbott and Anna Soubry, and an Islamophobic tweet and an anti-Semitic tweet. Can you confirm that you have taken all of those down?

Sinead McSweeney: Yes, all four were removed.

Q4 **Chair:** In August, my office reported a whole series of similar tweets and many of them are still up. They were not reported by me formally, so you would not have known that it was the Chair of the Home Affairs Select Committee who was reporting them; you would have seen them as somebody reporting them through a normal mechanism. There are a series of violent ones, including violent threats towards the current Prime Minister and a former Prime Minister, and again very racist abuse towards Diane Abbott. I will ask the Clerks to show you those tweets.

I appreciate that I am asking you to look at these very quickly, however I think it should be pretty clear that they similarly violate your community standards. Perhaps you can explain why those tweets, similarly reported, have not been taken down.

Sinead McSweeney: Obviously I am not going to go into each tweet one by one, not least because some of the content will be unsuitable for the public record. There are a couple of elements to this. I do not review this



HOUSE OF COMMONS

content every day, so I am not going to categorically say that every single one of these tweets violates our policy, but my initial read would suggest to me that they do violate our policies.

I think you will have noticed that among the changes we have made in recent weeks is a difference in how we are now treating what are called bystander reports. Originally, given the volume of reports that can be received, one of the ways in which we triage—almost like an emergency room—is that reports from the target of abuse receive a higher priority and attention from agents. Those are surfaced within our systems. To date, reports from bystanders, while they were used in some ways to inform our technologies and identify abuse on the platform, were not getting the same priority from agents. That is why there are occasions when abuse is reported by somebody who is not the target of the abuse, that they would not have been actioned in the past, and that has now changed. We have realised that bystander reports are an incredibly important signal to us of what is going on on the platform, particularly in the context of political figures because there is a growing recognition that besides the generally growing and unhealthy hostility towards politics and the narrative that is abroad more widely, both online and offline about politics, that people are standing up for their elected representatives and it is as valid for us to take into account the views of a voter who sees their representative and their beliefs being attacked.

Q5 **Chair:** If somebody in this room reported one of those tweets today, how long would it take for it to be taken down?

Sinead McSweeney: I cannot say categorically. It would depend on what else was going on in the world. We are not talking anything more than a day or two.

Q6 **Chair:** The problem is that that is not people's experience at all. Instead, people's experience is reporting a whole series of things and just getting no response at all, including victims of serious abuse and hate crime—also reporting them and getting no response at all.

Sinead McSweeney: I appreciate that and it is a frustration that I have heard from many people, including, as you say, victims of abuse. That is something on which we have put our hands up in the past and said, "We haven't been good enough at this. Not only have we not been good enough at actioning, but we haven't been good enough at telling people when we have actioned". That is something that particularly over the last six months we have worked very hard to change and we are making sure that we are communicating with people where they make their report, in the app, we are not sending emails to here, there and everywhere. You will definitely see people getting much more transparent communication at the individual level and much more action. We are now taking action against 10 times more accounts than we did in the past.

Q7 **Chair:** You raised the issue about public figures and particular politicians in a democracy. We have obviously had a series of reports, including from Amnesty International and the Committee on Standards in Public



HOUSE OF COMMONS

Life, which identify particular individuals getting targeted by an immense amount of abuse. They could not possibly report all those cases themselves, particularly Diane, who has had misogynist and racist abuse, Luciana Berger and Anna Soubry who have had death threats, and so on. Do you take a proactive approach to those individuals to look at the scale of abuse they are receiving to see what needs to be done about it?

Sinead McSweeney: Firstly, we have dedicated teams who work with parliamentarians and offer them safety and security advice. Also, there are particular escalation paths for elected officials who are experiencing difficulty. We have also worked hard, again within the platform, where we see somebody who is getting a lot of at mentions and a lot of abusive content. We will increasingly communicate with them within the platform to suggest ways in which they can protect themselves within the platform and also looking at actions that we can take against the content.

Q8 **Chair:** For example, with Diane, if I sat and did a search on Twitter with Diane are you saying that I would not find a huge series of Twitter abuse at Diane because you have these systems in place?

Sinead McSweeney: No. Right now, I can't say what you would find at this minute. If you are cleaning a street, you can clean the street every morning but you can't guarantee that is still going to be clean at 10 am.

Q9 **Chair:** Would I only find abuse for the last 48 hours and nothing before that?

Sinead McSweeney: No, I can't guarantee that.

Chair: We asked one of your colleagues last time about a series of tweets. James Berry, a former member of the Committee, asked you about the phrase "filthy Jew bitch" and whether or not that was appropriate. Your colleague Nick Pickles agreed that it was unacceptable unless it was being criticised. James raised two particular tweets that he quoted. One, from 17 October last year, refers to "high speed picture of filthy Jew getting bitch-slapped." He asked Mr Pickles whether that was appropriate and whether it would breach the community standards. Mr Pickles said "absolutely it would" breach the community standards. That tweet is still up on your platform. Why?

Sinead McSweeney: I don't know the answer to that question.

Q10 **Chair:** I'm wondering what it is we have to do. We sat in this Committee, in a public hearing, and raised a clearly vile anti-Semitic tweet with your organisation, it was discussed, and it is still there. Everybody accepted—you have accepted and your predecessor accepted—that it was unacceptable, but it is still there on the platform. What is it that we have to do to get you to take it down?

Sinead McSweeney: I have already laid out the number of measures that we have taken, fuelled by conversations around this table and in many other countries across Europe. We are constantly looking at these policies and looking at these processes. I will take that tweet away and I will come back to you with a definitive answer as to why it is still there.



HOUSE OF COMMONS

- Q11 **Chair:** Part of the problem is that if nothing happens even when we raise it in a forum like this, it is very hard for us to believe that enough is being done when everybody else across the country raises concerns. James raised another one, from 15 November 2014, attacking Luciana Berger, again with the same phrase. That was the second time that that one had been raised, because it was also raised by the Committee previously, and it is also still up on your platform. I don't know how many times we have to ask you to come and put the same things to you.

Sinead McSweeney: My reading of our policies—my reading of the policies as we have changed them—is that that tweet would be in violation of our policies, so I will come back to you with an answer as to why it is still on the platform.

- Q12 **Chair:** By the way, Mr Pickles said: "I would...be happy for my colleagues to review those accounts and write to the Committee". He never did, so it would be helpful if you did so.

Sinead McSweeney: I will ensure that that happens before the end of the week.

- Q13 **Chair:** Let me turn to Facebook. We put two to you. One is called "F Islam" and has an image of a huge bomb on Mecca. That one, as I understand it, is still up. Is that right?

Simon Milner: The page is still up. When we discussed this previously, I explained that people can criticise religion, including in strong terms, using language like that. However, the particular image you are talking about has come down, because that certainly appeared to us as if it was praising violence. That has come down, but the page itself has not come down.

- Q14 **Chair:** You have another one that we have put to you, which refers to schoolgirls and grooming.

Simon Milner: Yes. I understand why that was reported. It was concerning since, on the face of it, this closed group looked as if it may contain disturbing content. Actually, the group had no content. There were no posts in the group at all. None the less, we have taken that group down because of the inference associated with it, and the admins of that group are no longer on Facebook.

- Q15 **Chair:** That one was encouraging "Group pictures...of secondary schoolgirls age 11-16" and said: "Please only post pictures of girls wearing school uniform". This is obviously an issue that we raised with you last time. We raised specific concerns about the way in which different groups were being set up to post imagery of paedophilia and of child abuse. You said at that time that you were "learning from this particular...journey". You said that "this journey has brought" things "to our attention". You admitted that you had a problem with issues that had been raised with you by various journalists and said that you had addressed it, but clearly this has still appeared.

Simon Milner: Actually, just to be clear, the issue we had last time was that innocent images of children—these were not child abuse images—



HOUSE OF COMMONS

were being posted with very lurid comments underneath them. There was a fundamental problem in our review process, which has now been fixed, but as I explained, there were no posts in this group. The phenomenon that you rightly called us to account for last time was not occurring in this group. I recognise that you could not know that because you couldn't look inside the group in the way that we can.

- Q16 **Chair:** On the other one—the Islamophobic one—you say in your community standards: “Organisations and people dedicated to promoting hatred against these protected groups are not allowed a presence on Facebook.” So you don't simply look at individual posts; you also look at organisations. Do you think you are strong enough on Islamophobic organisations and groups and individuals?

Simon Milner: Is there a particular group you have in mind that you would like me to address?

- Q17 **Chair:** I am referring to that one generally, but I am interested more generally, because I know Tell MAMA has put a lot of complaints to you and other organisations.

Simon Milner: That page is not, obviously, run by a group and there is not a group affiliation associated with it. It is a page with incredibly small attention—less than a thousand likes. There is not a lot of content on it, so it is not obvious to us that this is a group that is using the page to foment hatred against Muslims. The content is disturbing but it is very much focused on the religion of Islam, not on Muslims.

- Q18 **Chair:** So you are confident on that issue that you have got the balance right?

Simon Milner: My colleagues who are the experts on this, who are making those judgments every day, have looked at that page and re-reviewed it, because your report included that there was one image that you particularly highlighted that should have come down and now has, but that the page itself is not, overall, violating our community standards.

- Q19 **Chair:** YouTube, we also sent you some images that included racist abuse and comments against MPs, particularly Diane, as I understand. Some you have removed, but you have left up something that is aggressively calling for a British MP and a British citizen to be deported because of the colour of her skin. Do you think that is okay?

Dr Berild Lundblad: No, we have done a run-through of the comments of the video that you sent us and removed several of them and we are still reviewing the more than 500 comments on that video in order to get it out. Let me tell you that one of the things we want to get to is a situation in which we can actively use machines to scan comments for attacks like these and remove them.

- Q20 **Chair:** For example, some of the ones that are still up: are you still looking at them? I am just going to pass you one of the ones—



HOUSE OF COMMONS

Dr Berild Lundblad: Our reviewers have reviewed the vast majority of them, but they are happy to come back if there are any specific ones that you would like to—

- Q21 **Chair:** I am going to just pass you one of them, which is calling for somebody to be put down, which, as I understand it, is still up on the site.

Dr Berild Lundblad: This is clearly one that I will submit back to our reviewers and make sure that you get a full answer on it. To me, although I am not one of our experts, it clearly seems to be in violation of our guidelines.

- Q22 **Chair:** Okay. Again, given that we know that particular individuals are getting a huge amount of targeting, I ask the same question that I asked Twitter. Do you have proactive support, for example, for individual MPs who receive a huge amount of racist and very threatening abuse?

Dr Berild Lundblad: We have proactive support for safety and security for parliamentarians, but we do not yet have the kind of proactive support that you describe, that would mean that we reach out toward—that we put in place policies that protect these MPs. This is a very valid question. One of the things we do is look at ways in which we can use machines in order to address comments, and we would naturally expect that we can train machines also to find particularly inflammatory commentary around specific individuals and figure out how we work with that.

- Q23 **Chair:** The last time we heard evidence from YouTube, one of the videos we put to YouTube was one from National Action, a banned organisation, banned by March of last year. It was a National Action propaganda video that we put to YouTube, which you and your colleagues accepted should be removed and took down. However, within a month, I found it again on YouTube. Several more months, I found it again and each time raised it with YouTube. It was just simply being posted on different channels. I raised it eventually with the chief executive of YouTube and then still found it again a month later on YouTube. I finally raised it with Kent Walker, the senior vice president of Google, this autumn and finally—your staff will probably be relieved to hear—I cannot currently find it on YouTube. However, that process took about eight months, with me as the Chair of the Select Committee raising it with the most senior people in your organisation, before YouTube managed to use the most basic technology that you use for your copyright issues all of the time to remove a banned video. Why did it take that much effort and that long just to get one video removed?

Dr Berild Lundblad: I can understand that is disappointing. What we have done now is made sure that we are looking at how to remove these videos more along—they are sometimes manipulated, so you have to work out how they have manipulated them in order to take the new versions down. We are now looking at removing them faster and faster. We have removed 135 of these videos, some of them within a few hours with no more than five views. We are committed to making sure that this improves over time.



HOUSE OF COMMONS

As we have started to roll out our machine learning in the last six months, we have improved. Although I really am sorry about the individual example, I think we will be closing that gap with the help of the machines, and I am happy to review this in due course. But clearly, as you say, it is disappointing.

- Q24 **Chair:** But you use that technology to remove Katy Perry videos, or people using music from different artists. You use it, and have been using it for a very long time. Why has it taken you so long to use it on issues around terrorism or extremism?

Dr Berild Lundblad: I think we have seen a sea change this year, thanks to the harsh and rightful criticism of this Committee, where we are looking at how we can deploy this machine learning, and use this technology, in a better way. We are now actively looking at ways in which we can use it for new categories, such as this content.

- Q25 **Chair:** I could not yesterday find that particular video on YouTube, but I could find it on both Twitter and Facebook. Do you guys not share?

Simon Milner: I am not aware of the particular video you are talking about, but I am happy to look into it. You didn't raise it with me last time.

- Q26 **Chair:** No I didn't, partly because I want to know whether or not you share information. It is a National Action video; it is linked to a march in Darlington in 2016 and is clearly a propaganda video with vile statements being made by this banned and proscribed organisation. Now, you have all referred to this Global Internet Forum where, in theory, you share imagery around terrorism. This is a proscribed and banned organisation, so clearly this particular image—despite me raising it repeatedly with YouTube and Google: I could not have raised it more often with you—has obviously not been shared with Twitter and Facebook or, if it has, the system isn't working to find it.

Simon Milner: I am happy to address that. The hash database is of terrorist images involving ISIS and al-Qaeda only; it does not involve other organisations. At the moment, that is very much focused on those global terrorist organisations that have been particularly cited by Governments and law enforcement and intelligence agencies as the main purveyors of terrorist content online, so that is where we have focused thus far.

I don't know the video in question and, of course, I am happy to look into this. One of the issues one often finds with this kind of content is that it can also be used by news organisations or people fighting extremism to highlight the activities of organisations like National Action. Without the context, I don't know whether or not that was the case here. Indeed, terrorist content is sometimes used by news organisations. Again, I don't know the detail, but I just want to caution that context really matters, therefore understanding the context in which this video has been shared on our platform could be extremely important.

- Q27 **Chair:** Yes, I think you will find the content is really straightforward and



HOUSE OF COMMONS

these are videos that should be removed. I don't think there is any question in these cases of the context being in any way acceptable, and these are banned organisations. Why do you not apply this co-operation to far right extremism or other kinds of terror threat? This is a banned organisation; it is a proscribed organisation.

Simon Milner: I honestly feel I have explained this. When it comes to Isis and al-Qaeda, they are the most extreme purveyors of this viral approach to distributing their propaganda. That is why we have addressed them first and foremost. It doesn't mean we are going stop there, but there is a difference between the kind of content they are producing, which is more often clearly illegal.

Again, without knowing the detail, certainly I have seen news reports about the behaviour of National Action, which will include some imagery of that organisation. If we were to say, "Let's just take out all images of anything to do with National Action," there would be a pretty good chance we would also be taking down important journalism. Of course, I totally trust your judgment around the context of this but I would like us to be able to review the post in question.

- Q28 **Chair:** It would be very helpful to have a response from you on that by the end of the week. I will pass you the link now because this is clearly a propaganda video built by National Action, and it has already been banned repeatedly by your colleagues at YouTube. It is incomprehensible that you would not be sharing this about other forms of violent extremism and terrorism, as well as ISIS and Islamist extremism, and the other problem for us is that this is an area that is easy for us to search because we understand the videos. We do not understand a lot of the Islamist and Arabic images and so on. It is hard for us to trust that you are doing this properly if we find all of these examples where it is clear that you are not. I did not give you a chance to respond: could you return to Twitter and respond to why you are not sharing imagery around far right extremism?

Sinead McSweeney: I do not have much to add to what Mr Milner said. I think that if we have learnt anything in the last year, it is that the world is becoming an increasingly challenging place, and that we have different forms of extremism that are emerging. We have learnt a lot from the sharing that we have done around terrorism, and—as Simon said—there is always room to improve those levels of information-sharing. The world in which we are now operating has become a much more challenging place in terms of the different forms of extremism, be this in thought, conduct or behaviours.

- Q29 **Chair:** Could I raise one more issue with you about the way in which these kinds of things are promoted? As a result of searching intermittently for that National Action video on YouTube, I have a series of recommended videos promoted to me when I now go on YouTube, presumably because of my interest in that National Action video. Included



HOUSE OF COMMONS

in that are a series of other frankly pretty vile videos about white genocide for example, and the recommended channels that come up to me from YouTube include Red Ice TV which—as I am sure you are aware—is a far-right vile channel that promotes quite a lot of racist material. Why am I getting recommendations from YouTube for some pretty horrible organisations?

Dr Berild Lundblad: I want to answer that in two steps. The recommendation itself builds on the kind of content that somebody has watched before. Clearly, this becomes a problem when you have content where you do not want people to end up in a bubble of hate, for example. One of the things we are doing as we review our policies as part of an ongoing project is to try and find states in which videos will have no recommendations and no impact recommendations at all. We are limiting the features, meaning that those videos will not have “recommended”, they will be behind an interstitial, they will not have any comments and so on. To achieve the scale we need, the way we will then address this would be to make sure that we use machine learning, identify these kinds of videos, limit their features and make sure that they do not turn up in the recommendations.

Chair: But you have decided not to put Red Ice TV into limited state. So for now it is not in limited state, despite the series of pretty horrible things that it has in it, and it has been promoted into my recommended list. It comes up on my recommendations. It is not simply that you have not removed it, and that you are making it hard or not hard for me to search for; you are actually actively recommending it to me. You are actively recommending what is effectively racist material into people’s timelines.

Dr Berild Lundblad: I will ask our reviewers to look at Red Ice TV and get back with guidance or response to what our reaction is. As I said, we are looking at how we can scale those new policies we have across areas such as hate speech and racism on the platform. We are six months into this, and we are not quite there yet.

Chair: Let me raise the same issue with Twitter. I know that you have now suspended the accounts of Britain First, Paul Golding and Jayda Fransen. About a week ago, I clicked on Twitter on an *Evening Standard* columnist’s tweets. She linked to a column that she had written about Nigel Farage and Donald Trump. When I then come back out into Twitter—having clicked on that link—I then find that I have “people also read” recommendations from Twitter, and that the second one in that queue was Paul Golding. Your organisation was recommending that I should follow or look at the tweets from Paul Golding because I clicked on an article in the *Evening Standard*. Why is that happening?

Sinead McSweeney: This is the point at which there is a tension between how much you use technology to find or flag up bad content, and how much you use it to make the user experience different. That tool that you are talking about is for somebody who wants to read or find out about the results of the FA Cup final, and they read one article and somebody suggests, “People who read that also read the following three articles.”



HOUSE OF COMMONS

That is the way that the technology works. To take that a step further and tell the machine, “We only want you to do that when it is about football, the Oscars or the election, and not when it is about certain bad subjects,” takes longer for the machine to learn. On the other hand, an academic might say to us, “I searched Twitter for X sentiments. I am doing research and it is helpful to me if you point me to those other places.” These are the balances, risks and decisions that we have to take. As Nicklas said, we are increasingly looking at how we can label certain types of content so that it is never recommended, but the reality is that the vast majority of a user’s experience on Twitter is something that they control entirely themselves—they control it through who they follow and what they search for.

- Q30 **Chair:** This is my final question. I think you have the same issue with Facebook, or if you go on One Page and there are recommendations about other people who like that page, or links and so on. The problem you have is that all three of your organisations use algorithms to encourage people who are interested in one particular thing to then follow something else. The police have said very clearly to us that they are extremely worried about online radicalisation and online grooming. Isn’t the real truth that your algorithms, and the way that you want to attract people to look at other linked and connected things, are actually doing that grooming and radicalisation? Once people go on one slightly dodgy thing, you are linking them to an awful lot of other similar things, whether that be racist extremism, Islamist extremism—your technology is doing that job, and you are not stopping it from doing so.

Simon Milner: I am happy to take that, given what others have said. I disagree that that is what the technology is doing, but I recognise that we have a problem—this is a shared problem with the police, yourselves and civil society organisations—of how we address the person who may be going down a channel that can lead to them being radicalised, either on the left or the right, or whatever, so that they ultimately become an extremist. That is one reason why we have established our Online Civil Courage Initiative in the UK, to really understand that phenomenon and see what we should be doing, using our technology, our people and our expertise, and considering how we can support other organisations as they use our platforms. It is unfortunate when those things happen. There are definitely things we are doing now, but there is much more to be done certainly by our company, and I am sure that my colleagues would say the same thing for themselves. That does not mean that we are happy that those things happen. We are really trying to address them, but we also see lots of examples of the opposite happening, and people who come online and are interested in modern-day Islam, and who find lots of really positive and encouraging content—lots of it linked together—which will produce a very positive outcome for that individual and community.

Dr Lundblad: An additional point is that first, if content is radicalising we do not want it on our platform. Secondly, I think you are right: there is a funnel of vulnerability where vulnerable individuals are exposed to content that is recommended. One way to think about how to address that is to



figure out whether there are ways to break that funnel. Yesterday we launched—we have been talking about this so it is not news—something called Redirect within the UK, and we will launch it later in other European countries. It is a way to catch people who are searching for this kind of content, and then to break that funnel and bring them to a counter-speech that essentially shows them or debunks what they are looking for. That initiative is first being used for anti-radicalisation work, and the idea is to catch people who are in a funnel of vulnerability, break that, and take them to a counter-speech that will debunk the myths of the Caliphate, for example. There are things that we can do, and we certainly take this problem very seriously.

Sinead McSweeney: This is an issue that I think about a lot, as a parent as much as somebody in my role at Twitter. I worry sometimes—I said this at the European Internet Forum recently—that when we talk about terrorism or when we talk about radicalisation or extremism, racism, grooming or any of these things, we put a huge amount of emphasis on where the messages are coming from and on stopping, blocking or censoring those messages. I think we need to put as much attention on looking at the audience—their vulnerabilities, building their resistance. That cuts across self-esteem, sense of self and digital and media literacy. All those initiatives have as much importance, because one of the most effective ways of defending against these messages and bad thoughts, bad thinking and bad motives is to build strength in the audience as much as blocking those messages from coming.

Q31 **Tim Loughton:** This is all sophistry, isn't it? I understand, where you have got individual postings, that one can have a subjective argument as to whether they abuse your individual codes or not. As the chair said, we are still talking about pretty basic stuff that the most basic technology should proactively filter out.

Yesterday there was a big spat on Twitter about somebody who had posted about "kick a Tory". To give them their due, someone in the Labour Party said, "This is appalling—I will report it". This morning, I went online on Twitter and searched for "kick a Tory". There is quite a lot there. There is a quote from former Dr Who, Tom Baker: "I'd love to kick a Tory MP in the bollocks", which has given rise to lots of responses, including, "But Tom, you can't, all Tories are the C word"—all readily advisable.

I am prepared to believe that social media platforms are equal opportunity abuse hosters, so I searched for "kick a Socialist" and found that the content had been taken down. Worse still, if you search "#killatory" you find: "If everyone who tweeted about #Thatcher killed one Tory each, the country would be fixed in time for Corrie Unite Sports Fans #killatory." Also: "David Cameron has doomed the young and deserves nothing short than a bullet to the head, #killatory". And: "I really do think we should all kill a Tory. Just think of the benefits if each family in the UK were to kill just one Tory." That is on Twitter, Ms McSweeney.



HOUSE OF COMMONS

Your code says, you will “not tolerate violent threats, wishes for the physical harm, death or disease of individuals or groups”. How does that comply with your code? Simple hashtag, simple search, simple take-down, or even better, simply do not allow them there in the first place.

Sinead McSweeney: There are a couple of elements to this. At its most practical level, we have 500 million tweets a day, we have 330 million users. Twitter is used in multiple languages. We can sit here in London and talk about one hashtag relating to one political party, but if you are to have technology that is proactively searching for the kind of words that you are talking about across all political parties in the world in all languages, that is a much more significant task, particularly if you then layer on the amount of human review that would be required to ensure that it is actually abusive. Of course, it feels like something that is simple—

- Q32 **Tim Loughton:** That’s because it is. You are in every country in the world. It wouldn’t take many hours to design a filtering mechanism so that every political party in each country in which you operate, if somebody puts up a hashtag or a phrase which says “#killatory”, “#killasocialist”, “#killaLibDem” or whatever equivalent, that wouldn’t happen. Under your “Hateful Conduct”, it says: “You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability or disease”.

You commented earlier about political figures being in the frame, recently. Are you prepared to add political figures to that list and would that then help you to be more proactive in filtering out really violent abuse against people in public life?

Sinead McSweeney: I worked in politics for many years. I knocked doors for politicians, I stuffed envelopes for politicians. I have a great love of, a respect for, the political process. I do not know of any country in the world that has included political figures as a protected category in its incitement against hatred legislation. However, I find it deeply concerning that we are now in a world where there is such an atmosphere of hostility towards the body politic, a growth in binary thinking, and abuse rather than considered argument. I think what you are talking about online is actually symptomatic of a much, much bigger challenge.

- Q33 **Tim Loughton:** So don’t allow it on your platform.

Sinead McSweeney: It is not a simple problem in society and it is not a simple problem on the platform. As we have seen where people have used abusive hashtags against members of the gay community, were you just to say, “Take down everything that uses that hashtag,” you would remove a lot of tweets and content from people like yourself or your supporters who are saying, “This is wrong. This shouldn’t happen. This isn’t how we should disagree about political issues. This isn’t how the left should disagree with the right, or the right with the left.”

- Q34 **Tim Loughton:** Ms McSweeney, this is not about left or right. It is not



HOUSE OF COMMONS

about taking away somebody's right to criticise somebody whose politics they do not agree with. It is about not providing a platform, whatever ills of society you want to blame it on, for stuff that incites people to kill, harm or maim people—that incites violence against people—because of their political beliefs, in this case. We could come up with a whole load of other examples. Frankly, you say, "This is a problem with society," but you can do something about it. You are profiting from the fact that people use your platforms, and you are profiting, I am afraid, from the fact that people use your platforms to further the ills of society. You are allowing them to do it and doing very little, through simple things like that, proactively to prevent them. That is the nub of the matter, is it not?

Sinead McSweeney: I disagree. I don't think you can boil it down to that being the nub of the matter.

Tim Loughton: But it is.

Sinead McSweeney: There are a couple of elements here. You are focusing on one policy. We have multiple policies under which some of that content may be found to breach our terms of service—harassment, targeted abuse and so on—but you cannot just isolate those tweets from the wider challenges that we are now facing in how we disagree with each other.

Q35 **Tim Loughton:** This is not about disagreeing; it is about inciting a crime against somebody else. Mr Milner and Dr Lundblad, you both talked about being proactive. You have spoken about taking on extra reviewers, although according to a report in *The Times*, in the case of Google, "the reviewers will not proactively scour YouTube for inappropriate content as has been demanded". What proactively are your two companies doing to prevent harmful hate stuff on your platforms?

Simon Milner: Particularly focused on public figures like yourselves?

Tim Loughton: Well, anything. I have just quoted one example.

Simon Milner: Specifically on that, one of the things we did for the snap election in the summer was to offer all 5,000 candidates from every party, as part of the training on using our services, safety training about how to keep themselves safe on our platform. Unfortunately, because of the nature of the snap election—we understand why—lots of candidates were simply not able to avail themselves of that. That is a real shame, because we had recognised that this growing phenomenon, as highlighted by the Committee on Standards in Public Life report, was something where we needed to do more and to step up and offer more help. We also did have a dedicated channel for candidates to let us know about abuse so we could deal with it quickly.

Q36 **Tim Loughton:** That's not proactive, is it? That's reporting.

Simon Milner: I think it is proactive.



HOUSE OF COMMONS

- Q37 **Tim Loughton:** It's not proactive. If somebody has already put something on your social media platform, that is not proactive. I want to know what you are doing proactively so the minute somebody tries to put on some offensive stuff that breaches your code—I have given examples and the Chair has given others—it is just not allowed on to that platform.

Simon Milner: When it comes to speech about politicians, we don't do that. We don't do that for the very clear reason that we are very cautious about interfering in political speech. There are some politicians who actively want to take on the trolls and abusers, and they do not want this content taken down. They want to address it. If we started to say, "We are going to take this down automatically," that would be very difficult—judging whether content is abusive or is responding to abuse is often very difficult, as Ms McSweeney has explained—and, as I say, different politicians have different perspectives about the kinds of protections they want in place.

- Q38 **Tim Loughton:** We are all a bunch of masochists, basically, you are saying.

Simon Milner: I didn't say that.

- Q39 **Tim Loughton:** I want to get to the bottom of this proactive stuff. I read in the paper a fascinating article the other day saying that facial recognition cameras linked to Facebook profiles can be used to tell shop managers which products to promote to individual customers. You are now providing very advanced technology. When I walk into a shop, those cameras will scan my face and analyse what buying mood I am in, and pass that information to the very helpful sales assistant to make sure I go and buy various things. That is quite advanced technology, and yet you cannot even proactively come up with the technology for these 10,000 reviewers who are not going to do anything proactive to stop the sort of blatant, clear abusive stuff that we have all quoted.

Simon Milner: To be clear, the 10,000 refers to YouTube. Our reviewers are involved in some proactive work. As we have explained extensively, there are certain types of content that we focus on initially—the most extreme content. Child abuse imagery has been addressed for many years in a proactive way. We have all talked about terrorism and extremism. Some of the other areas we are talking about are more difficult. That does not mean we are going to stop—of course we are going to try to do more work on our own initiative, but when it comes to political speech, it is much more difficult. I am sure that you, as politicians, would also be concerned about the prospect of us restricting political speech in an automated way, without anybody overseeing it. That is a big step to take.

- Q40 **Tim Loughton:** For "Kill a Tory," substitute "Kill a black person" or "Kill a Muslim". It is all the same; don't try to make it about politics. Dr Lundblad, particularly to deter serial abusers online who hide behind the anonymity they can have on Twitter, YouTube or Facebook, isn't the answer a requirement, which at least one country in the world has, for people to reveal their identity to the social platform host and be approved before they can post any such material? Then, if they do abuse it, they



HOUSE OF COMMONS

can be struck down and cannot re-emerge five minutes later with another completely anonymous identity. Why are you not doing that?

Dr Lundblad: I would like to address two things. The first is the proactivity point. It is factually incorrect to say that they will not be engaged in proactive work—

Tim Loughton: They won't be.

Dr Lundblad: As I described, they will be identifying risks. Once those risk patterns emerge, with new things happening in hate speech or whatever area it is, they will find examples, and they will find a lot of them. That is where the proactivity point comes in. The misunderstanding we are sometimes fighting with is that if all these people just search the platform, that is not how you achieve scale. You achieve scale through finding examples and then training the machines, which will do the equivalent of 180,000 people's work, working a 40-hour week. There is definitely that sequence, which I described at the beginning to the Chair, where there is a strong element of finding examples and training machines. I want to make that clear first.

On the issue of verification and identification, we take down abuse on our platforms whether or not those are verified accounts or identified accounts. There are technologies that make sure people cannot immediately sign up for a new account, and there are ways for us to make sure that if someone's account is stricken, they cannot immediately sign up for a new anonymous one. We are somewhat hesitant to disclose those in a public forum, because they will then be circumvented and abused, but we are happy to give you a more thorough information security briefing on how we work to make sure that even anonymous accounts are not open to people abusing and then signing up and abusing again—the pattern that you describe as problematic.

The question of identification is much broader. I would submit that on YouTube, there are plenty of platforms that benefit from anonymity. There are, for example, communities of people who have suffered domestic abuse and who want to have anonymity, because that is part of their ability to share their experiences.

- Q41 **Tim Loughton:** But that's nonsense, isn't it? I came to your office with others from the Committee last month and spoke to your senior legal counsel. I was presented with a young Syrian woman who said that she absolutely needed anonymity because she was working for Syrian refugees at risk from the Government. I perfectly understand that, but who is to say she is not working for the Syrian Government?

Dr Lundblad: I am not sure I understand the question.

- Q42 **Tim Loughton:** Okay. You have just said that you require anonymity, because people could be at risk if their anonymity were compromised. You do not check on whether that is bona fide or not, and it could equally be used for malign purposes as benign purposes.



HOUSE OF COMMONS

Dr Lundblad: This is where the question of abuse, whether it is anonymous or identified, comes in, because we take down abuse no matter what. There are communities that tell us that they do benefit from the anonymity that is afforded. But, again, it is a very broad question. It is ultimately one that goes to the heart of a whole series of questions, if you want to identify everyone who is online in different forums. We have made the choice not to do so.

Q43 **Sarah Jones:** I am aware of time, so if you could keep your answers really short, that would be good. I want to ask a couple of questions about potential policy developments and a couple about money.

German legislation requires social media companies to remove illegal hate speech and other criminal material within 24 hours of a complaint. Would you be prepared to voluntarily have that target; say, to remove all criminal material within 24 hours and terrorist material within two hours? Would you do that voluntarily, all three companies?

Dr Lundblad: We have committed in the G7 to work, where technically feasible and within the protection of human rights, to get to a point within one or two hours. Today, when it comes to terrorism material, 50% of the material is removed within two hours, 70% within eight hours, and we are closing the gap. So we have already committed to the terrorist part of that. When it comes to the hate speech part, we are signatories of a code of conduct that Commissioner Věra Jourová has set up for the European Union, which is currently undergoing a test that has a 24-hour turnaround time requirement. The idea here is that this test will be published before Christmas and I think many of our organisations are a part of it.

Q44 **Sarah Jones:** What is your target for reaching the two-hour terrorist target?

Dr Lundblad: I do not know what the target is but we have said that, obviously, we aim to do this as fast as possible.

Sarah Jones: As fast as possible, okay, thank you.

Sinead McSweeney: As Nicklas has said, we are also part of the exercise that is being conducted across the EU and our systems are currently being tested to see whether we are meeting the commitments around the 24 hours.

Q45 **Sarah Jones:** And on the two-hour terror target, is that—

Sinead McSweeney: Right now, 75% of the accounts that we are removing have not even tweeted, so the question of the one hour is becoming increasingly irrelevant, through a combination of the fact that 95% of the accounts are being identified by our own technology and 75% of the ones we are removing are removed before they even tweet.

Q46 **Sarah Jones:** But you do not know how many are taken down within two hours? You do not have that?

Sinead McSweeney: Do you mean of the tweeting or creation?



HOUSE OF COMMONS

Q47 **Sarah Jones:** Any kind of terrorist information that is on there.

Sinead McSweeney: Seventy-five per cent come down before they ever tweet.

Simon Milner: For us, we recently reported that 83% of the terrorist content we take down is taken down within one hour of upload. So we think we are doing quite well against that target but we want to try and do better.

Q48 **Sarah Jones:** Do you want to try and get to 100% within one hour?

Simon Milner: We have committed to 100% within two hours, but as I say, we have managed to achieve 83% within one hour.

Q49 **Sarah Jones:** Okay, and 100% within two hours by—

Simon Milner: As soon as possible. It is really a case of homing in on the 17% that we are not managing to get within the hour. Why not? Let's try and get better at that.

Dr Lundblad: To be clear, we are not counting what we catch on upload so it doesn't ever even hit the platform. There is a lot of content that comes in. It is just the content that hits the platform, on Sinead's point, which is important.

Q50 **Sarah Jones:** The Government published a Green Paper on internet safety strategy. Is there anything in there that you disagree with?

Simon Milner: We have responded to that consultation. There is much of it we support. The one part that we are not sure is the right approach is the social media levy, in part because the parallels are drawn from industries such as alcohol and gambling, where there is lots of evidence that those things are never good for you. You might argue, especially at this time of year, that a bit of alcohol is okay. Gambling is never good for you, whereas in social media, there are lots of aspects of social media.

The Committee on Standards in Public Life last week said the vast majority of messages that MPs and candidates saw during the recent election were positive. So there is lots of good on social media and therefore to apply what is almost like a sin tax to us—we are not convinced. But that does not mean we do not think we should be called to account for how much work we are doing on educational awareness, which is what the levy is designed to pay for.

Q51 **Sarah Jones:** The Committee on Standards in Public Life said that potentially we should introduce legislation to shift responsibility for content—yes or no, do you agree or disagree that we should introduce legislation for that? It would shift responsibility and liability for content to the platforms that host it. Yes or no?

Simon Milner: I went to the launch of the report last week, and I noted that the Chair, Lord Bew, said that that will not be needed if the companies get better at taking this content down and are more transparent. I very much hope that that is where we will get to, because

we certainly have concerns about shifting the liability boundary in a fundamental way.

Sarah Jones: So at the moment you all disagree with that.

Simon Milner: We have concerns about it.

Q52 **Sarah Jones:** Fine, okay. I have one more policy question. You are essentially giant advertising platforms, aren't you? Do you think that Ofcom should have more control over what you do?

Dr Lundblad: I think it is not a question of binary application. We are heavily regulated and there are already 10 pieces of legislation in the UK alone, and European directives and so on, that apply to different aspects of the services that we offer. I think there are aspects within what YouTube does that could count as a kind of broadcasting and we register accordingly, with the audiovisual media services directive as our legislative basis in the European Union. For a certain part, I think there are legitimate questions about the regulatory situation. Otherwise, I think we are regulated differently for all the different services that we provide.

Q53 **Sarah Jones:** So there is some room—you think there could be changes.

Dr Lundblad: I think that if you look at the audiovisual media services directive and at how you register under that for content that you produce yourself on your platform—that is the case in a very small set of channels that YouTube produces—we would certainly consider being under that regulation.

Simon Milner: There are two types of regulation that apply to our advertising. One is around the data and how that is used, and data protection regulation applies to our platforms. The other is advertising, and the Advertising Standards Authority has a remit that covers online advertising in the UK, and therefore the advertising on Facebook that appears in the UK is subject to its regulation.

Q54 **Sarah Jones:** I have some very quick questions on money. Do you spend more or less money on technology to tackle ad-blocking than on identifying extremist content?

Simon Milner: That type of technology is not relevant to our service.

Sarah Jones: I mean Google.

Dr Lundblad: I don't know. I very clearly think that that would be highly unlikely, to be honest, but I am happy to look at that.

Q55 **Sarah Jones:** Could you go away and come back to us with the answer?

Dr Lundblad: I can certainly do that. I would submit in the interim that I believe we spend a lot more on the removal of extremist content.

Sarah Jones: Than on tackling ad-blocking.

Dr Lundblad: Yes, definitely.



HOUSE OF COMMONS

- Q56 **Sarah Jones:** Okay. Can you tell us how much revenue you lost as a result of companies pulling ads in response to being featured next to extremist content? How much did you lose from that?

Dr Lundblad: I don't know. I know that we are in the trust business and it is crucial for us not to have any abuse on our services. We hear fast from our creators, users and advertisers when there is any kind of abuse of content. It is certainly not good for us, and we take it very seriously because we want to maintain the trust of those three different groups. As to a number in pounds—I don't think that is something that we have disclosed.

- Q57 **Sarah Jones:** That would be good to find out. This the last question from me. How much does it cost to employ one person to monitor content in general—how much do the many thousands of people that you employ cost each?

Dr Lundblad: It will be market salary in the countries where we employ them. We will also move people around and they will do different things. Some of them will be expert in their area and will be looking at new patterns, for example. Some of them will be working just to make the human judgment at the end of the machine process.

- Q58 **Sarah Jones:** Is there an “-ish” salary? Is it quite low skilled? I am trying to get a sense of the nature of this work.

Dr Lundblad: No, it goes from high skilled to low skilled; it is really the entire spectrum. If you think about people who are expert in identifying terrorist content, that is a high-skilled thing. For somebody who is working more with the nuts and bolts of reviewing content, that is—well, it is not low skilled; I submit that it is still quite high skilled, and it is very intense work. We have made sure that working conditions for the people who do this are good—it is gruelling work—and that they have access to onsite and offsite counselling, and very limited working times. Things like that are really important to us because this work is not pleasant.

- Q59 **Sarah Jones:** Do either of you have a ballpark figure for how much it costs to employ one person?

Simon Milner: No.

Sinead McSweeney: No. It would vary from country to country; there are different levels.

- Q60 **Rehman Chishti:** Mr Milner, thank you for your encouragement earlier. It was very pleasing to attend the seminar and event in Parliament in relation to dealing with hatred towards faith communities, and in particular the Muslim community.

Just to clarify, you spoke earlier about the proactive nature of taking down material. What percentage did you say you take down?

Simon Milner: Of content related to the two global terrorist organisations that I mentioned, ISIS and al-Qaeda, 99% of the content we take down is on our own initiative.



HOUSE OF COMMONS

- Q61 **Rehman Chishti:** A week or so after the excellent seminar you had in Parliament dealing with hatred and tackling hatred towards the Muslim community, if one goes on the internet one will find material like this, which I will send to you, appearing on there: "Islam or die. I say, fuck it, let's. England, no surrender. So convert to Islam or die," with a crusader sign like that.

So you have this brilliant seminar in Parliament saying with regards to the Muslim community, we are doing everything we can—like every other faith community—in terms of hatred. Then you find material like that, and then you take into account what you say about a proactive nature. Do you not think that undermines your proactiveness and the building of confidence and trust in those communities?

Simon Milner: I can understand why people will feel frustrated. To suggest that that event that we held was, as it were, "That's it." Just to be clear, the proactive work we are doing, as we explained at the beginning, is focused particularly this year on terrorist organisations. My understanding of the content that I can see from here, is that this is not involving those terrorist organisations.

- Q62 **Rehman Chishti:** On that material, if I may just clarify, with regards to your comments earlier to the Chairman, people have the right to criticise religion but what you see there is not criticising religion. That is completely unacceptable. Do you not accept that?

Simon Milner: I would have to ask our expert teams to look at it and the context.

- Q63 **Rehman Chishti:** But, from the point of view of an ordinary person looking at that, would you not say that that material out there is completely unacceptable?

Simon Milner: I'm not sure that a picture of a knight saying, "England, no surrender", would be infringing our terms.

- Q64 **Rehman Chishti:** No, let us clarify. The top bit comes in conjunction with it—one image after another, after another. What does the top image say then? You read the middle one. What is the top image?

Simon Milner: I would rather not use one particular word, but I will say, "Islam or die. I say, F-it, let's".

- Q65 **Rehman Chishti:** And the last one?

Simon Milner: Is, "So convert to Islam or die".

- Q66 **Rehman Chishti:** Yes, with those images. You mentioned the middle one, but all three come together. I am asking you about proportionality and acceptability. People out there come to your seminars in Parliament and you say, "We want to work and engage with faith communities," and then they see this kind of material on your system. At this point in time, are you not prepared to say whether that is acceptable or not?



HOUSE OF COMMONS

Simon Milner: No, because that is not my role in the company. We have experts who are expert in reviewing content. I am very happy to ask them to look at this.

Q67 **Rehman Chishti:** Would you do that?

Simon Milner: Of course.

Q68 **Rehman Chishti:** Secondly, a point was made earlier with regards to the social media levy. In Germany, the Parliament has put forward legislation to say that it must be done within 24 hours or there are sanctions for the social media companies. You have come back to the Government on the Green Paper saying that you don't agree with that. The point I make is that if you are asked repeatedly to take it down and you give assurances, but we still see material like that, surely you should take some responsibility and be penalised. If you have nothing to fear, you should be okay with a financial penalty at the end of it.

Simon Milner: We take our responsibilities in this area very seriously. When it comes to the German legislation, that has now been completed but it is not yet in action. One of the concerns we have about it is that you are asking our companies, rather than the courts, to decide what is illegal. We think that it is problematic for national jurisdictions. Should we be held to account on whether we are applying our standards in a proper way? Should we be held to account on how transparent we are being? Absolutely. It is one of the reasons why we very much welcome—it may not feel like it sometimes—coming to this Committee to give an account of what we are doing. But we are not resting on our laurels: we know we need to get better.

Q69 **Rehman Chishti:** I take that point. I am running out of time, but I have two questions. On Twitter, you talked earlier about dealing with the different types of extremism. The Government do not have a working definition of extremism at the moment. The extremism commissioner will be appointed shortly, or is about to be in position to look at the working definition of extremism. When you talk about different types of extremism, one sees the mindset of how your policy is set. How do you define extremism?

Sinead McSweeney: I think what I was talking about there was probably in a slightly different context, which was the growth of extremist and binary thought, and this kind of opposition; you know, there is no grey area or complexity allowed in debate anymore.

Clearly, in the area you were talking about, there are terrorist groups—there are legal definitions of terrorist groups. We announced a policy a month ago, which we started implementing yesterday, about groups and organisations that advocate violence against other groups, so there is that.

When you get into the area of extremist thought, that becomes more challenging. I look back over the last five and a half years, and I would have given very different answers to those questions five years ago because Twitter was in a place where it believed that the most effective



HOUSE OF COMMONS

antidote to bad speech was good speech. It was very much a John Stuart Mill-like philosophy. We have realised that the world in which we live has changed, and we have had to go on a journey with it and realise that it is no longer possible to stand up for all speech in the hope that ultimately society will become a better place because racism, homophobia or extremism will be challenged. We do have to take steps—as we do now—to limit the visibility of hateful symbols, and to ban people from our platform who are affiliated with violent groups. That is the journey that we are on.

- Q70 Rehman Chishti:** A final question to YouTube. You talked earlier about counter-narrative online, in terms of debunking an idea and to defeat the “caliphate”. If I am right in terms of what you said earlier, you put a counter-narrative online for what is going on there, so that people see that, in order to—as you said—defeat the “caliphate”. As you and I know, words have significant meanings. A poison ideology is put forward by the death cult in Syria and Iraq, Daesh. It deliberately chooses to call itself ISIL, ISIS or Islamic State to gain legitimacy. This Government has taken a decision to use the correct terminology to defeat its ideology and the appeal of it. People see YouTube representatives use the word “caliphate”; “caliphate” was an old Islamic form of government that no longer exists, and a death cult has chosen to call itself that to gain legitimacy. Do you not think that it is important to use the correct terminology and wording to help defeat this ideology’s appeal, and that it is important for you to work with the Government to help you use the right terminology?

Dr Lundblad: Absolutely. I unreservedly apologise for using the wrong word. I am sorry, and I agree that words matter a lot. My apologies.

Rehman Chishti: Thank you.

- Q71 Naz Shah:** My question is to Google first, and it relates to algorithms. Recent reports have highlighted that inappropriate content is appearing alongside children’s videos on YouTube. How and why is this happening?

Dr Lundblad: It is interesting because this is a story where some of our expert groups with whom we worked flagged a phenomenon to us that has been emerging recently. What has happened is that you have innocently uploaded pictures that are then used in a deeply harmful way. This is something described by Mr Milner earlier. Comments can be of a completely unacceptable nature, for example. One of the things that we have seen is that as we start clearing the platform, you then see displacement effects. People go elsewhere. They use comment fields, they go to other services and so on. It is one of those things, and we quickly caught it.

There are three steps in our response: the first is that anything that is even vaguely criminal, in terms of soliciting or hinting at criminal abuse of children in other ways, is reported to the National Centre for Missing and Exploited Children—NCMEC in the United States—which has a network of law enforcement contacts across the world. We proactively look for that,



HOUSE OF COMMONS

make sure that we report it, and make sure that it is brought to the right law enforcement attention. That is the first step, and it is absolutely essential. There is no place for child abuse on our services.

The second step is that we have looked at these videos that are sometimes innocently posted but are subverted to these horrible uses. We have taken down 150,000 of those in the last month and we have removed comments from 625,000 videos in order to remove anything that could be within the realm of being close to the kind of practice that you describe. We come back to scale: we have taken swift action, and we have also trained our machines to catch and remove this.

- Q72 **Naz Shah:** But I am talking about videos appearing alongside children's videos. That is a bit different from what you are suggesting. My understanding is that if my six-year-old decides to go on YouTube, which she often does, to watch something, there might be something highly inappropriate coming up alongside that video that would be very upsetting for my child. What are you doing to tackle that bit?

Dr Lundblad: There are also examples of kids' cartoon figures being used in a way that is deeply inappropriate. This goes to the same policy shift that I just described. We have started to detect those videos, too. Where they are in breach of our guidelines, we remove them. Where they are adult humour put into a cartoon, which sometimes happens, we put them behind an interstitial and we make sure that you have to log in in order to see them and they do not appear next to children's material. We are rolling this out. It is not perfect yet, but we are taking care of this, because we take this very seriously and want to make sure that children are not harmed in any way, shape or form online. The other aspect of this is that at the end of the day we want to make sure that we have safe environments online for children.

- Q73 **Naz Shah:** So are you accepting that your algorithms were wrong, given how this happened?

Dr Lundblad: Yes.

- Q74 **Naz Shah:** And are you fixing that?

Dr Lundblad: Yes.

- Q75 **Naz Shah:** My next question is for all three of you, really. Every time you go on to a YouTube video and so on, you are recommended something similar, as the Chair suggested earlier. What are you actively doing not to promote the spread of similar hatred?

Simon Milner: I can give an example of this from the recent snap election. We deployed a relatively new product called Perspectives, so that when anybody interacted with news content around the election, they would see a service—or a product or a feature—when they came back to Facebook whereby they could see where all the parties stood on election issues. They would go into that and then content provided not by Facebook but by the parties would explain to them, on defence, the environment and so on, "Here are where the main parties all stand." We



have also tried showing people stories about the same issue from very different media outlets. One of the things that is quite hard is getting people to read it. You may well serve up alternative content, rather like the redirect feature that Nicklas explained, but you can't always get people to read it. There are things we are trying to do to address this issue of the so-called filter bubble.

Sinead McSweeney: Our suggested additional reading feature is actually relatively new. It is something that we are always looking at and evolving, and we are taking on board the kinds of challenges that you mention. The important thing to remember about Twitter is that, because of the open nature of the platform, if you access content—if you look at a tweet—you may get “People also read”, but immediately beneath that you will generally also get 10, 20 or 30 people vehemently disagreeing with the point in the original tweet. There is always that scope for alternative views and perspectives almost within the size of your phone screen.

- Q76 **Naz Shah:** Ms McSweeney, I am pleased with some of the evidence we have heard this morning about work that has been done—I am particularly pleased with the employment of so many people by Facebook—but Twitter has fared pretty badly. You have not even got it right where we gave you things that clearly violated your standards, not just ours, that you needed to remove months and months ago. What other assurances can you give us that you are actually actively challenging what we are talking about?

Sinead McSweeney: As I said at the outset, over the course of the last 12 months we have introduced 25 separate new measures and changes to our policies, our processes and the product itself. We have increased immensely the tools that we give users to limit the amount of abusive content that they see, whether that is through blocking or through muting content. You can choose never to see content from somebody who has not verified their phone number or has not changed their profile picture. You can mute certain words.

Separately, we have improved the way in which content is reported to us. We have improved the tools available to the teams reviewing those reports. We have, as I said, taken into account in a much more significant way reports from bystanders, and we have made significant changes to our policies. As I said, there are things that are against our terms of service today that were not against our terms of service not only when I joined five years ago, but last month. Who is to know what further changes will come and what further iterations of those policies there will be?

- Q77 **Naz Shah:** So am I right in saying, from what you just told me, that you rely on users to police your platform, as opposed to Twitter taking responsibility?

Sinead McSweeney: No.

Naz Shah: That is what I just heard. If I report it as a user report, you will look into it, but if not, it is not something you are looking for actively.

Sinead McSweeney: No, that is not correct across the board. There are certain areas of content—terrorism, child protection and so on—where we will adopt much more proactive measures than we will do at the other end. It is a combination of all those things. It is a combination of using our own technology and using reports from people who are the subject of abuse and people who are witnessing abuse.

Q78 **Naz Shah:** How many more people are you employing, or looking to employ, to address this?

Sinead McSweeney: It's interesting: in terms of scale, compared with the other companies, Twitter as a company is only 3,500 people globally. That is the entire workforce. Over the last two years, we have gone through two rounds of redundancies. Despite that, our user safety teams and user services teams have consistently increased, and they continue to increase. I speak at our new hire orientation in Dublin every week, and every week there are new people joining those teams. They will continue to increase, and particularly now, with the new policies we have introduced in recent weeks, we will require more people to enforce those policies.

Q79 **Naz Shah:** What is YouTube doing to actively redirect people?

Dr Lundblad: The redirect project that I described earlier is one example of how we use advertising technology to catch people and then bring them to counter-speech of different versions. The other thing we do is that as we review videos and find that they are unsuitable, even though our experts may find that they are not to be taken down, we limit their features, so that they do not appear in the recommendation stream that you see. The idea is to roll that out broadly—we have started to apply it in the last six months—so that you will not get things in your recommendation stream that are of a disturbing nature.

Q80 **Naz Shah:** Dr Nicklas, how ineffective are your algorithms if you are not able to protect children and have got that one wrong, and you have not been able to redirect hate speech, despite being before us so many times? What is it going to take to get this right?

Dr Lundblad: I think we are getting it right, to be honest. I think we are making huge progress. What I mentioned in terms of take-downs, the amount of content we are addressing, and the speed with which we are addressing it—all of those are quickly improving. To be fair, we have realised and come to the conclusion that this is a year of sea change for us. We are putting enormous resources into this, and we are very serious about getting it right.

Q81 **Naz Shah:** I went to a YouTube event that was very good. From some of the things that your YouTubers were talking about, I know that you have the talent there on your platform. Should you not be using that talent actively—especially in your Muslim community, who were at that event—to counteract some of the narratives? I don't see any of your users being directed to an alternative message on YouTube.



HOUSE OF COMMONS

Dr Lundblad: Thank you for the question. We have a programme called Creators for Change that does exactly that, where we fund new creators. It is also important to say that the vast majority of the content on YouTube is good content. It is educational. It is entertaining. It is valuable content. As we roll the Creators for Change programme out in the UK—it is already under way—it will allow us to find all those alternative voices and bring them to the surface. You're right; there is a limited amount that a corporation speaking can do, but when the creators on the platform—the people who use and shape the community of the platform—speak out, that is really important. We have done this in other countries as well. We have a programme called #NichtEgal in Germany, for example, where young people speak out against any kind of online hate or online harassment, in order to get the cultural shift that we need. That is important.

Q82 **Douglas Ross:** Ms McSweeney, what have Paul Golding and Jayda Fransen done in the last 24 hours to get their accounts suspended?

Sinead McSweeney: We announced a policy—

Douglas Ross: On 15 November.

Sinead McSweeney: In the first week of November. We began implementing that policy yesterday. It says that you cannot be on our platform if you are affiliated with a group that encourages violence against protected groups, and that we would take into account both online and offline behaviour.

Q83 **Douglas Ross:** We are aware of the statement you made on 15 November. What have you done in the more than a month between making that statement and removing those accounts 24 hours before this Committee meeting?

Sinead McSweeney: As we indicated at the time, there are a number of things. We published a blog, I think in early November, which explains what happens when we change a policy. One of the things is we very often sense-proof that policy with our Trust and Safety Council and with other groups. We then set about training the various agents who have to enforce that policy. There is actually a requirement under EU law that when you are making a substantial change to your terms of service, you have to give your users 30 days' notice of that change. That is what we have been fulfilling—all those requirements—since we announced the policy change.

Q84 **Douglas Ross:** So along with those two accounts, how many others have been taken down in the last x number of days?

Sinead McSweeney: I don't have a number. When we started enforcing the policy yesterday, it wasn't one day of action; it was day one of enforcing the policy, and the policy will continue to be enforced on the platform until such time as it is updated.

Q85 **Douglas Ross:** So, it's not just these two?



HOUSE OF COMMONS

Sinead McSweeney: No, there were many, many accounts across the world which were suspended, but we will suspend accounts today, tomorrow and right through as individuals or groups with those affiliations come to our attention.

- Q86 **Douglas Ross:** You said you were conducting a review of verified accounts, including verification from accounts whose behaviour does not fall within these new guidelines. How long will that process take?

Sinead McSweeney: I don't know. It became clear recently that our verification process was broken. People had been verified who should never have been verified. People's understanding of what it meant to be verified on the platform varied. It was originally introduced during a health crisis. I can't remember if it was bird flu or swine flu, but it was to verify that the public health information coming from the World Health Organisation was valid, because there were a number of impersonation accounts. That is the genesis of verification. It was to authenticate that public health information was coming from a recognised source. Over the years, it evolved, and people began to start looking at it as a status symbol or endorsement. We can't have it being any of those things, so the entire system of verification is now suspended and it is being fully reviewed.

- Q87 **Douglas Ross:** What was the public interest in verifying those two accounts?

Sinead McSweeney: As I said, it became clear to us that there are a number of accounts that should never have been verified.

- Q88 **Douglas Ross:** Yes, but your statement prior to 15 November said that accounts must be in the public interest before they are verified, so I am interested to know: what was the public interest at the time you verified those accounts?

Sinead McSweeney: That is the point I am making to you.

Douglas Ross: No, sorry. Just so I can get the right question, in case that is causing you trouble. The point is you have changed your policy now in early November but at the time you verified these accounts it was very clear—you have it stated in your policy at the time—that you would only verify accounts when it is in the public interest to verify them. So there must have been a point when someone at Twitter got a request to verify these two accounts and someone, a group, an organisation such Twitter, said, "Yes, that does meet our criteria to be in the public interest." I would like to know what the public interest was in verifying those two accounts at that time.

Sinead McSweeney: Okay, firstly, and I do understand the question, there are two separate issues. There is a policy that we have about whether or not people can have accounts on Twitter or whether they can post certain content on Twitter and then there is a verification policy. I have said that our verification policy was broken from top to bottom. Not only did people externally not understand what it meant to be verified, we



HOUSE OF COMMONS

were increasingly running into challenges that the very people that you are talking about who had to make decisions on who should be verified—the account had become so layered and complicated that it wasn't being applied consistently. But the reality is that at the time those accounts were verified Britain First was still a registered political entity in this jurisdiction.

Q89 **Douglas Ross:** Yes, but the Britain First Twitter page was not verified?

Sinead McSweeney: Yes, but the two individuals were members, affiliated with it.

Q90 **Douglas Ross:** Sorry, you are using your argument to verify these two accounts by saying a political party or Britain First was registered as a political party, yet Twitter themselves would not verify the Britain First account, but you would verify the account of these two individuals. I am just quoting from your own verification procedures—you tried to suggest that there was a difference between setting up an account and setting up a verification, and I want the Committee to be clear that I was using the terminology for the verification element. Twitter says that “an account may be verified if it is determined to be an account of public interest”. I am not sure—what was the public interest of these two individuals, when you have accepted yourself that the party's Twitter account presumably was not in the public interest because it was not verified?

Sinead McSweeney: As I have said to you, and as we have said publicly, our system of verification was broken. Not only did people outside not understand it, but we increasingly came to realise that the individuals who we had tasked with enforcing it didn't understand it. It is entirely in suspension and subject to review.

Q91 **Douglas Ross:** Google and Facebook, you have both spoken about the increased number of people who are looking at quoted claims, and images and so on that should be taken down. What targets are set for those members of staff to deal with these inquiries? We have heard about 100% in two hours, but I mean more generally. I am asking about a general member of the public who reports something to you but, as the Committee has heard, they do not hear anything back. What target are you setting your staff to say, “When this comes in this is how we respond; this is the timescale in which we respond”? How do you monitor adherence to those targets?

Simon Milner: We have different what we call turnaround times for different types of content and different teams, depending on the issue and, to some extent, on events. For instance, when there are terrible events around the world, including terrorist attacks for instance, we will put something like a SWAT team on to that, to look out for content that might be praising the attackers. We want people who are ready so that as soon as the names of victims are published we can memorialise accounts and work with families around that. It does depend on the issues, and on events.

Q92 **Douglas Ross:** If we take out exceptional events, what is your standard



HOUSE OF COMMONS

target for saying, "Here are some complaints that we have received from a member of the public. I know that my department is working to target, because it is dealing with 98% of them within 48 hours"?

Simon Milner: We aim to get to every single report within 48 hours, but many, many of those teams will have much shorter turnaround times, particularly on the kind of content that we have been talking about. There are also lots of reports about very low-level abuse such as spam. As I said to the Committee when I appeared previously—I am happy to repeat it now—you are very welcome to come to our Dublin office to meet the teams who are responsible for this work, and the leaders of those teams, and to hear more about their work.

Q93 **Douglas Ross:** So every report you get from someone complaining about something on Facebook will be responded to within 48 hours?

Simon Milner: That is absolutely our aim, and we aim to do it much more quickly for more serious issues, particularly when people are feeling at risk.

Q94 **Douglas Ross:** So if I as a Member of Parliament ask for people to contact me if they have reported something to Facebook that has not been dealt with and they have not had a response within 48 hours, I will get zero responses from any member of the public in the United Kingdom.

Simon Milner: I doubt that very much because we are far from perfect. We are dealing with millions of reports every single week.

Q95 **Douglas Ross:** So of those millions of reports, what percentage are being dealt with within 48 hours? You are saying that they all are, but you also say that you cannot respond to so many. I am not sure what the answer is.

Simon Milner: We've also said that we recognise that we need to be more transparent about this. At the moment we do not have a public number of the type you are describing, but we are certainly looking to be more transparent about the effectiveness of our reporting and review processes over time, and there will be more to come on this for us next year.

Q96 **Douglas Ross:** But you have a number. You know internally if your departments are working well or not.

Simon Milner: We certainly have lots of metrics around how effective our teams are at handling reports, but I am not aware of a headline number in the way you describe it.

Q97 **Douglas Ross:** How do you decide if your teams are working efficiently?

Simon Milner: As I say, we have lots of different metrics. There is an awful lot of focus on this area. Just recently, our CEO, Mark Zuckerberg, said on our earnings call that we are going to prioritise keeping people safe on Facebook and improving our resourcing and effectiveness on that ahead of profits. We have made that very clear, but it is also an indication



HOUSE OF COMMONS

that we are far from there yet, and there is an awful lot more to be achieved.

- Q98 **Douglas Ross:** But you are saying the headline point that we should all take away and people should be reassured by, is that once you've reported something, it will be dealt with in 48 hours.

Simon Milner: No, I'm not saying that. You were asking me specifically about it. I did not say, "That's the most important statistic." I shared some numbers earlier about the particular issues raised by this Committee last time I was here, in March, around what we are doing about terrorist content, how much of it we are addressing on our own initiative and how quickly we are doing it. I have given you two very significant numbers: we take down 99% of content of our own volition, and 83% of that is taken down within one hour of upload. Those are more important numbers than an aggregate number of what proportion of reports are dealt with in 48 hours.

- Q99 **Douglas Ross:** You give us those very impressive figures for that section, but what about the figures for racial abuse of politicians, racial abuse in general, homophobic abuse, religious abuse and so on?

Simon Milner: I recognise, Mr Ross, that these are really good and important questions. We do not yet have answers for them, but there is more to come from our company when it comes to transparency.

- Q100 **Douglas Ross:** Can you understand my suspicion, then, that you can give us great figures, like 98%, which sounds really good, yet you cannot give us figures for many other areas? Is that because the response rate isn't so good for them?

Simon Milner: No, it's just that we do not yet have our system set up for us to be able to produce those numbers and to feel confident in them. Believe me, we have dedicated teams working on this very hard, including some of our engineering teams here in London. There is more to come from our company on this. I would be very happy to come back to the Committee in due course. In fact, you are welcome to come and visit our office in London—that is much nearer than the one in Dublin.

- Q101 **Douglas Ross:** I've already been there, actually. When can the Committee expect those figures?

Simon Milner: You will definitely be getting more from us next year.

- Q102 **Douglas Ross:** Early next year?

Simon Milner: I don't know. I'm afraid I can't give you definitive timing on that.

- Q103 **Douglas Ross:** And Google?

Dr Lundblad: There are different turnaround times for different subjects, just like Mr Milner said. We have the code of conduct with Věra Jourová in the European Union, which is 24 hours for hate speech and one to two hours for terrorism. But in order to follow on from the Committee's



HOUSE OF COMMONS

recommendations to hold us accountable, we are going to publish a transparency report in the spring on our turnaround times and the kinds of categories and flags we get. I hope that we can turn that into some kind of standard so it is possible to look at exactly what those are and how they improve over time.

Q104 **Douglas Ross:** Time is very short, so let me move on to a different area. How many requests to take down terrorist content has each organisation received from the Counter Terrorism Internet Referral Unit?

Simon Milner: I can't give you a definitive number on that, but I am happy to look into it and to write to the Committee.

Q105 **Douglas Ross:** You are all very high-profile people within your companies. Is it single figures? Is it hundreds? Is it thousands? Do we have a rough ballpark figure?

Sinead McSweeney: It is more.

Q106 **Douglas Ross:** More than single figures, hundreds or thousands?

Sinead McSweeney: It is double figures per month.

Q107 **Douglas Ross:** Is it the same for everyone?

Simon Milner: I suspect it may well be three figures over the course of the year.

Dr Lundblad: Yes.

Simon Milner: But rather than guess I would much rather write to the Committee to let you know that number.

Sinead McSweeney: There is an important point here, and it is a point I actually made at the European Internet Forum. Given how much better companies are getting at proactively identifying content, we need to be very careful not to judge the effectiveness of internet referral units based on the number of referrals they make, because there is less content there for them to refer. But they continue to play a really important role, because they are at the coalface of counter-terrorism investigation in identifying and communicating new trends to us. With a background in law enforcement, I think we need to be careful not to judge the success and work of internet referral units based on the number of referrals they make to us.

Dr Lundblad: This is a good point. We have seen those numbers decrease over time, but that is because of better collaboration.

Chair: Some more written information on that would be helpful.

Q108 **Stephen Doughty:** Can I just be clear with Google and YouTube and with Twitter? Information from proscribed terrorist organisations in the UK should not be on your platforms. Is that correct?

Sinead McSweeney: Yes.

Dr Lundblad: Yes.

Q109 **Stephen Doughty:** Okay. Can you both explain, then, why there is so much material on both of your platforms from dissident terrorist organisations in Northern Ireland, both republican and loyalist? I have just done a very quick search. There are videos on YouTube entitled “Up the IRA” and “Fuck the British Army”, and pictures of bombings of buses. We are not talking about historical content here; we are talking about promoting current activities. There are also videos from loyalist terrorist organisations, including one entitled “KILL ALL TAIGS”. There are similar things on Twitter. There is even a Twitter account called “@uptheira”. Can you explain why that content is on your platforms?

Dr Lundblad: If there is a proscribed organisation, it should not be on our platforms. If it is there, it is there because we have not caught it and taken it down. We are happy to look at those particular videos and make sure that we take action on them and get back to you with a full review.

Q110 **Stephen Doughty:** And Twitter?

Sinead McSweeney: Similarly, if they are terrorist organisations, they should not be on the platform. Some of those organisations will also come within the violent extremist groupings, particularly some of the loyalist groups you mentioned. Again, following the commencement and enforcement of that policy yesterday, I would see those accounts coming down.

Q111 **Stephen Doughty:** Okay, but this stuff has been on there for a significant amount of time. Why has it not been removed proactively? This is very clear-cut. We are not talking about historical stuff, which I understand there is a legitimate interest in, particularly with regard to some of the organisations; we are talking about contemporary accounts and information. Why is it not being proactively removed?

Dr Lundblad: One of the changes that has been made after speaking to this Committee and others is that we are now turning to identifying risks, actively looking at content and then training machines to get it. That means that we still have a lot of work to do in order to make sure that we take care of things like this. We are taking that very seriously.

Q112 **Stephen Doughty:** Do you proactively search for content from the official list of proscribed organisations provided by the UK Government?

Dr Lundblad: Yes, we—

Stephen Doughty: All of those organisations?

Dr Lundblad: What we do is look for those organisations and then train the machines.

Q113 **Stephen Doughty:** So how is this content still on your platform?

Dr Lundblad: I think the answer is that we have not caught it yet and that we are going to review it. I do not know what the content is, but I am

more than happy to look at it. If it is from proscribed organisations, I agree with you that it is unacceptable.

Q114 Stephen Doughty: Just to be clear, you search extensively around the names and related terminology used by all the organisations on the UK Government's official list of proscribed organisations?

Dr Lundblad: As I said—

Stephen Doughty: Yes or no: do you search all of them?

Dr Lundblad: In the last six months, we have started implementing all of those controls. We should do better.

Q115 Stephen Doughty: So that's a no. Twitter, do you search proactively for all of the proscribed organisations, as defined by the UK Government?

Sinead McSweeney: No.

Stephen Doughty: No?

Sinead McSweeney: The majority of our proactive technology is currently being leveraged towards ISIS in particular, because they are currently posing the biggest threat.

Q116 Stephen Doughty: Just for information, does Facebook search for all the proscribed organisations?

Simon Milner: If you don't mind, I would prefer to come back to you in writing on that, because I want to be really clear on the answer. I know that there are certain things that we do proactively, but I want to be clear whether it is the full list.

Q117 Stephen Doughty: As Mr Loughton next to me is saying, this is pretty basic stuff. We are not talking about stuff that can be judged one way or the other. These are the organisations officially proscribed by the UK Government, and I find it quite astonishing, particularly from Twitter, that you are not searching for them. These are known terrorist organisations. They have names. They have terminology associated with them, and you are not searching for it.

Sinead McSweeney: Like Simon, I would prefer to check directly with the teams, but I do know that a huge amount of the emphasis has been on the activity of ISIS, as I said, because that is what the attention from people like yourselves and from Government has been on.

Q118 Stephen Doughty: Do you think it is acceptable for there to be a Twitter account called @uptheira?

Sinead McSweeney: I spent four years of my life working for the Police Service of Northern Ireland. I worked with people who had colleagues and family members who were murdered by the IRA, so I have very strong personal feelings about that. I would like to see the actual accounts and the content on them before I answer specifically on their presence on our platform.



HOUSE OF COMMONS

Q119 Stephen Doughty: There's plenty, and they are extensive. We don't even need to go into the information that the Clerks have provided for us; I had a quick look this morning for anti-Semitic content on Twitter, for example, and straight away I found an account called Creator Skinhead that said "My race is my religion" and had far-right extremist imagery in the background. I will not repeat the words here, but it was stuff that is easily searchable and is not questionable in any way. You are just not proactively getting this, are you?

Sinead McSweeney: Again, we have made the changes on hateful symbols and use of hateful terms in profiles, account names and bios in literally the last 10 to 20 days, and we are working our way through the implementation of those policies.

Q120 Stephen Doughty: Okay, but clearly you cannot cope proactively. You have not given us the number of staff you are using to look at these things. It is absurd. On that far-right account, which I found literally just at the start of this Committee, one of the tweets says that the "account is temporarily unavailable because it violates the Twitter Media Policy", but right above it is a tweet that says—I will not use the exact words—"No k—! No n— or gays...Hey hey! Ho ho! Commie scum has got to go!" Beneath, you have something that appears to have been removed, and above is something that has not been. That further suggests that the content above is okay. The bottom line is that you do not seem to be taking seriously enough the proactive removal of very obvious content. We have heard this from everybody on the Committee.

Sinead McSweeney: I can understand how difficult it is for you to just take my word or for me to convince you of how seriously we take this. I have just come back from a week in San Francisco, where there are three gatherings every day of people from right across our policy teams, our product teams and our engineering teams, with our CEO on occasions and other senior members. At those gatherings, we talk about these issues and measure and level-set on where we are with them every single day.

Q121 Stephen Doughty: So how many staff in Twitter are dealing with these issues?

Sinead McSweeney: I have already said that our whole population is 3,500, so in terms of the figures you have heard from the other companies, no figure I give you will sound like it is enough, because—

Q122 Stephen Doughty: And what were your global revenues in the last quarter?

Sinead McSweeney: I don't have that figure off the top of my head, but I can tell you that—

Q123 Stephen Doughty: It was \$574 million in the last quarter. You say you only have 3,500 staff. Do you not think that you ought to be spending a little bit more money on proactively searching for some of this content?

Sinead McSweeney: I can tell you that there is a significant percentage of staff—not just the agents who review content, but the people who



HOUSE OF COMMONS

devise and implement our policies and huge swathes of our product and engineering teams—who are working on safety issues right now and are off other projects.

Q124 Stephen Doughty: You ask us to trust you. Issues have emerged about Britain First and Donald Trump and others in the last few weeks. We heard from Mr Ross about the German policy of having to remove stuff within 24 hours. It took me nearly 24 hours to get a response from one of your senior staff when I raised concerns about the verified nature of the accounts and what they were tweeting, and it then took you nearly three weeks to take any action. Do you think that is acceptable?

Sinead McSweeney: I have explained to you why the implementation of that particular policy had to occur a period of time after we announced that we were making that change. I have also explained the challenges that we have experienced with our verification process and the fact that we have very openly put our hands up and said, "This is broken." We are therefore suspending the process of verification and will fully review it and announce a new policy in the new year.

Q125 Stephen Doughty: Okay. I am sorry to say that the impression—this is specific to Twitter—is that either you cannot cope or your organisation does not care about these issues. The lack of action taken so far and the lack of response, despite these things being repeatedly raised with you as an organisation, is simply not good enough. Quite frankly, I don't trust where you are at at the moment.

Sinead McSweeney: I am sorry to hear you say that, and nothing that I say sitting here in front of you is going to persuade you otherwise. All I can say is that I have worked for this company for almost six years. Prior to that, I worked in policing and in politics. It is not in my nature to work for or with a group of people who I cannot respect or who I do not believe in. I believe in the people. I believe in my colleagues. I believe in their intentions, and I believe that our trust and safety teams are passionate about making our platform a better place for everybody. That is all I can do here today. In terms of how transparent we can be on figures, that is also constantly changing. I am in a position to give answers to questions that I would not have been in a position to give two or three years ago. We are definitely becoming more open, and I hope that we will—

Stephen Doughty: With respect, I think that that reflects the cultural problem in the organisation. We will leave it there.

Sinead McSweeney: I think we will continue to become open.

Q126 Stephen Doughty: The facts suggest otherwise, Ms McSweeney. I am not suggesting that it is you personally, but the facts suggest otherwise. The reality is that we have had to repeatedly raise these issues, action is not taken even when the Chair of the Home Affairs Committee or senior Members of Parliament raise these concerns and you have to be dragged here to answer these questions.



Sinead McSweeney: I definitely wasn't dragged here. I was quite happy to have a trip to London in Christmas week. I can totally understand where you are coming from. I had this problem when I worked in police communications; you would talk about what our response times were for call-outs to burglaries, or what our detection rates were. There is always going to be one case where the police officer didn't say the right thing, or turned somebody away from the information desk. There will always be those cases, but believe me, the sea change I have seen in how this company addresses abuse and safety during the five and a half years that I have worked here is absolutely fundamental. Personally, I would not still be working for them if it hadn't happened.

Q127 **Will Quince:** We have had a lot of questions about the proactive and preventive steps that you could take. I want to touch on reactive steps that you can take, and to some extent are taking. If I can pose a hypothetical scenario, let us say that tomorrow you walk into a shop with your young children, and all of a sudden somebody starts hurling racist or homophobic foul language and abuse, perhaps even body shaming—the list is potentially limitless. Do you think that the shop owner or management has some responsibility for what their customers do within the shop? That is the first question.

The second question is how you would feel if you approached the management of said shop and they said, "Look, we don't really screen our customers for what they say, and we don't really remove potentially offensive customers." How would you feel if they said, "You can deal with this yourself. Consider the context of what they said, and perhaps just consider ignoring them"? How would you feel if that were the case?

A third and final question: if you then did persist and reported it to the manager, and the manager said, "Put it in writing, and perhaps we'll respond in a few days. Maybe we'll ban them from the shop, but if they come in in different clothing, then we'll just let them back in again to commit exactly the same acts," how would you feel?

Simon Milner: Who is that for?

Will Quince: All of you, if that's okay.

Simon Milner: Of course I understand what you are driving at. I wouldn't use that shop. One of the—

Q128 **Will Quince:** Are you saying therefore that we shouldn't use your platforms?

Simon Milner: No, what I am saying is—

Will Quince: That is kind of what you just said.

Simon Milner: Just to be clear, what I am saying is that if we don't get this right and continue to improve with it, we will absolutely find that not only will shoppers not come to our shop, but the organisations that advertise with us will also not want to do that. We have all had experiences of that over the years. When we have been found wanting,



HOUSE OF COMMONS

advertisers have said, "We're really not sure about whether we want to advertise with you." We get this. We absolutely get it. It is one of the reasons we take it very seriously. We are held to account by people on Facebook—you might think of them as customers—by politicians, by regulators and by advertisers all the time.

There is intense focus on us in this area. That is why all of us are devoting more resources and energy to this. It is right from the top of our organisations. Sinead talked about her CEO; as I mentioned, my CEO, Mark Zuckerberg, has been extremely public about this, and has said to our investors, "We are prioritising this above maximising profits, and you need to know that as we go into 2018."

If I was in that shop and that store owner said the things I have just said to you, in a different vernacular for that environment, I would think, "I want to make sure they do it, so I am going to expect to see, over the course of the next few months, that they do what they say they will do." I hope you will recognise—Ms Shah has left now—that, as she said, we have got better since last time we were here, but we know we are not done. We've got to continue to improve.

Q129 Will Quince: If I could drill into that, you mentioned the point about prioritising safety over profit. That is good news, because £20 billion buys an awful lot of safety. You mentioned what Mark Zuckerberg says publicly, but your company consistently says on the record that you don't want to be a moderator and take an editorial stance. But if you are going to tackle this and take it seriously, you have to do that. You have to be a moderator, and you have to take an editorial stance.

Simon Milner: If you can show me the quote, I would be interested in it. It could be around the issue of whether we are a publisher or not. We absolutely take responsibility for what happens on our platform, and that is one of the reasons why we are investing so much in adding so many people and developing increasingly clever technology. Some of our best engineers are working on these issues, because we know we have enormous amounts of responsibility. Does that mean we are responsible for everything that happens on the platform before we even know about it? Clearly not. What we are trying to do is to get better at improving our knowledge of what is happening. In the past, that was all about reporting. Now it is increasingly about technology and the sophistication and expertise of the people who we employ to achieve knowledge, so that we can act in a responsible way.

Sinead McSweeney: I'm smiling because I love the analogy. It is similar to one I used when I appeared at this Committee a number of years ago, when I talked about the levels of enforcement of good behaviour in a society. In the context of a public park, the extent to which, if you have antisocial behaviour, there is a certain response from the people in the park, you go to the park warden and then ultimately you have types of behaviour that require a police interaction.



HOUSE OF COMMONS

As online platforms, I think we should be aiming for something which is the opposite of what you have portrayed, but there are a couple of other elements. If I, as a mother, had taken my child into that shop and that happened, as well as complaining about what had happened, I would also get down to his level and look him in the eye and tell him that there are bad people in the world who hold hateful views and that he needs to stand up against them. I go back to the point that we have to focus on that element of this problem. Removing these statements and content from our platforms won't stop people thinking them or acting on them. We need to be committed in that area as well.

In terms of the journey that we have come on, when I think of my son, I visited London with him recently and for the first time, I was handed change in a diner and I got the Jane Austen £10 note and I got incredibly excited about it, and he couldn't understand. "Why are you so excited by this £10 note?" I told him the story of what had happened to Caroline Criado-Perez, and I actually sent her a message after and I said, "In terms of a learning moment, we covered everything from people being hateful, misogyny, people standing up for what they believe in, people looking for change." That was a huge learning moment for Twitter. We didn't learn enough, quite frankly, at the time, and it is only in the last year that we have really faced up to this issue. But I think it is useful to use offline analogies, like you did, to make us really think in a deeper and more impactful way about what our platforms should feel like.

Q130 Will Quince: As a follow-up for you, how can it be right that I can report someone for abuse and within an hour, with a new email address, they have set up a new account and they are back at it?

Sinead McSweeney: We are moving increasingly to prevent that happening. If people are suspended for a breach of our terms of service, firstly we look for various signals to prevent them coming back on the platform if it is a temporary ban, because we do have a lever in trying to get behaviour changed. We try to maybe give people time out, as it were. They have limited functionality until they delete certain tweets, but they will only come back on if they verify their phone number or email address. We have seen some success with that. Of people who have been put into that limited functionality state, 25% are only ever in it once, and we have seen 65% are only in that state once, so they are learning from those experiences.

Q131 Will Quince: In terms of your response to this, I had a couple of accounts reported, do you respond quicker and more proactively to blue-tick accounts or verified accounts above ordinary Twitter members?

Sinead McSweeney: No, it is based on what the report is about. Reports about violent threats are going to be actioned much more quickly than reports about spammy behaviour. The triaging is based on the severity of what has been reported.

Dr Lundblad: I agree with Mr Milner that we absolutely see that we are responsible ourselves for our platform and need to make sure that it is a



platform that is acceptable to everyone, not only through our own actions, but also through building an ecosystem around the platform that works well. One thing we are doing is building a network of trusted flaggers, which are expert organisations that will help us to not only flag content but to identify new trends and risks and so on, so that we can actively act on those. One thing that came out of the Committee's last hearing is that those trusted flaggers should not be expected to work for free, so we are funding the Internet Watch Foundation and the Institute for Strategic Dialogue, both in order to make sure that they can do that on a reasonable basis and because it is fair.

Another part of the civil society analogy that you have—I concur; I think it is a good thing for us to think in offline analogies—is that we also need education. We are investing heavily in different educational programmes. The two in the UK—Internet Legends, for younger kids, and Internet Citizens, for older kids—teach kids that they should not behave differently online from how they do offline. Some 93% of the kids who go through the programme actually say that they will change their behaviour.

Is that permanent? Is it true? Will it really happen or not? I think we can grow those educational programmes, together with other organisations, NGOs and the Government, in order to make sure that we preserve civil society at the heart of our online presence. That translation of civil society to the online presence is something that I think we need to take responsibility for and bring educational resources to.

Q132 Will Quince: Okay. I will read a quote from a former vice-president of Facebook: "I think we have created tools that are ripping apart the social fabric of how society works". That is pretty hard hitting. If you are serious about tackling this, not only the Committee but the public need to see that you are serious.

My eldest is getting a tablet for Christmas—I hope she is not listening to this—and she will have access to the internet for the first time. That terrifies me, quite frankly. If parents and regular users of your platforms are to have confidence, we need to know how responsive you are being to all these different reports that are being made to you.

Will you publish those reports—on a quarterly basis, as a minimum—categorised by what the reports are, in terms of different categories, and how responsive you were to each of those individual reports and the action that you took? That is the only way that we will clearly see, without hauling you before the Committee twice a year or however often it needs to be, that you are taking this as seriously as we believe you should be.

Dr Lundblad: Yes.

Sinead McSweeney: I think we have to be careful and make sure that what we publish is meaningful. That is the piece of work that is going on right now. We publish transparency reports twice a year, and to date those have focused on Government requests. As we move on, we need to look at



HOUSE OF COMMONS

how we talk about terms of service breaches, but we have to be sure that it is meaningful.

For example, after we announced the implementation of our policy yesterday, various alt right message boards were full of calls to action for people to try to game our reporting systems. Our stats yesterday on the number of reports we received will therefore not give you a measure of how many people felt abused on the platform, and our enforcement rates for yesterday will not be a measure of how effective we are. That is the bit we have to work on to ensure that the figures we share with you actually tell you something.

Q133 Will Quince: Is that a yes or a no, though?

Sinead McSweeney: That's what we're working on.

Simon Milner: I can't give you a simple yes or no. I can tell you that we are going to produce more meaningful data, to help you and many others who are asking the same questions to understand this. I hope that your daughter is in one of the schools where our programme—either through the Diana Award or with Childnet—is rolled out.

If she is, she will have somebody in that school who we have empowered to help young people like her with their new tablets to understand how to keep safe and how to get the best out of the internet, because there are lots of wonderful things. She is very welcome, provided she is 13.

Chair: This is an unusually long Committee sitting, and I appreciate everybody's patience and understanding of that. Given the seriousness of the issues, I think it is appropriate for us to take this time. However, I ask the witnesses not to repeat answers they have already given in response to questions.

Q134 Preet Kaur Gill: My first question is to Mr Milner. In May, *The Guardian* reported on leaked documents that outlined Facebook's internal guidelines for content removal. Things such as "I hope someone kills you", live stream attempts to self-harm and some photos of non-sexual physical abuse and bullying of children would not be deleted. Can you tell me what legal and safeguarding organisations you consulted before coming to some of these conclusions?

Simon Milner: There are a number of organisations that are part of our global safety network; I mentioned one of them earlier, which is Childnet. We, like the other companies, are strong supporters of the Internet Watch Foundation. They are a very important part of that global safety network. And there are many others. In the interests of time, I am very happy to write to you and set out those partnerships.

Q135 Preet Kaur Gill: Okay. But are you saying that it is okay for someone to say, "I hope someone kills you"?

Simon Milner: No, I am not saying that.

Q136 Preet Kaur Gill: Are you saying that these organisations have given you



HOUSE OF COMMONS

legal advice to say that? This is in your guidelines.

Simon Milner: I apologise. Again, I think I may have misunderstood your question. There are a number of different organisations that provide advice to us on our community standards and the application of them. As to the specifics of the individual things you are raising, as often with leaked documents, you shouldn't necessarily believe everything you read. Often those things can be extremely out of date and are not necessarily being applied at the moment. I am pretty certain that that is the case with the story you are referring to. But I am happy to write to you to explain our response to that story.

Q137 **Preet Kaur Gill:** I have read in lots of articles, though, that a lot of your reasoning is that there are benefits to raising awareness of difficult issues. Do you think that that justifies not removing content that depicts things like child abuse, self-harm or violent death?

Simon Milner: Let me take the issue of self-harm. One of the tricky new phenomena we have come across is this business—with a product called Facebook Live, which enables you to go live on Facebook—of people indicating that they wish to harm, even kill themselves, and they are live on Facebook.

When we consulted suicide prevention agencies, including the Samaritans in the UK, about what we should do in those situations, their advice, uniformly across the board, was, "Do not take that content down because when that person is live and they are still alive you can help them. Instead, look to work with law enforcement"—which we do—"and other organisations, and friends, to figure out how to get a message to that person, to say to them, 'Please stop. We're here to help you'".

There have been a number of cases since we launched Facebook Live, of us being able to get the police to that individual in order to safeguard them. Not every time, I'm afraid. If somebody does take their life, we will immediately take the content down. That is an example of the difficulties here, and why having expert partners to advise us on new phenomena and what we should do about them is incredibly important.

Q138 **Preet Kaur Gill:** Facebook Live will be very difficult for you to regulate and monitor. Are you looking at other mechanisms? There is loads of technology out there and you could maybe think about how you use Facebook Live differently.

Simon Milner: You are right. Facebook Live is a tricky one because it is happening right now. But particularly, if somebody reports something to us and the thing they are reporting has already happened in the live stream, how do we look at it?

One of the things we have done is to improve our technology to allow our agents to rewind the live stream to look for the particular thing that somebody has identified. So there has been a lot of engineering effort deployed around Facebook Live so that we get much, much better, both at



HOUSE OF COMMONS

looking really quickly at things that people alert us to and at spotting some things using our technology.

Q139 Preet Kaur Gill: I know lots of people have given you lots of examples today of concerns that have been relayed, but I would like to know what you think constitutes a credible threat in relation to material that may encourage terrorist activity. What resources or additional information are used before coming to that judgment?

Simon Milner: Do you want me to answer that as well?

Preet Kaur Gill: You can start.

Simon Milner: There is quite a lot in that, and what I can tell you is that we have now employed some of the foremost experts in the world on the behaviour of these kinds of organisation and on what we can do to spot them and address their activities. It is through their work that we have been able to be much, much more effective at addressing terrorist content and the supporters of terrorism.

On the specifics you ask for, again, it is quite a detailed thing so if you want to follow up on it I would rather write to you to explain how we do that.

Dr Lundblad: Just to add to that, it is important to state that we also work together with Europol and other organisations to help us to understand this, as well as recruiting our own experts, to have a good broad understanding. There is also within that question the very real question of how to give emergency legal assistance when something seems to be imminent. I am more than happy to share that with you as well.

Q140 Preet Kaur Gill: In terms of the position that Germany is taking, they would say that even with the work with Europol, it is still taking a long time to get the material down.

Dr Lundblad: Europol themselves are very happy with the development and have said so in the European Internet Forum. We agree that we can do better—both Europol and us—but we are improving, and we are getting to a point that they feel better about. We can always do better.

Q141 Preet Kaur Gill: According to Google and YouTube, what constitutes a credible threat in terms of encouraging terrorist activity?

Dr Lundblad: I would say the same thing. It is a complex concept. I would be happy to ask our experts to provide you with the exact assessment that they do, because that is the best way you can get a good, solid answer. We have teams of experts who work on this, and I would be very happy to have them come back to you.

Sinead McSweeney: Similarly, I don't think I would do justice to the complexity of how those decisions are made. As well as the information that we share among each other through the Global Internet Forum to Counter Terrorism, we work with various law enforcement agencies and



HOUSE OF COMMONS

leverage academic organisations like VOX-Pol. Maura Conway, for example, is one of the world's leading experts now on extremists' use of the internet. She is based in Dublin and has been in to talk with our teams.

Q142 Preet Kaur Gill: Twitter has a new policy—I appreciate that it is very new. However, there are accounts that appear to have fallen foul of the new rules. Those are people like Jared Taylor, who is the head of the American Renaissance website, which champions racial difference; the American Nazi party; the Traditionalist Worker Party, a white supremacist group; and an account called End Time Paradigm, which specialised in anti-Semitic content. They fall outside of the rules. How will you address some of those organisations?

Sinead McSweeney: I cannot categorically say that any account that is still on the platform today does not fall foul of the rules. Again, the implementation of the policy started yesterday. It was not one day of action. It is a policy that now exists and that we will continue to review accounts against.

Q143 Preet Kaur Gill: Twitter's GIFs for users are provided by a third party, and there has been some anti-Semitic stuff. I know you have been working with organisations to address that, but what minimum standards do you apply to contracted services?

Sinead McSweeney: I would have to come back to you on the detail of that. In all our relationships with third parties, there are contracts and agreements. I can come back to you with the detail of the standards that we expect.

Q144 Mr Chope: Most of the problems we have been describing today would be avoided if there was openness and transparency of the identity of the people who are using your platforms. Why do you not go down that road?

Simon Milner: Perhaps I could take this first, because our service at Facebook is somewhat different from the other services. We do require authentic identity on Facebook. However, that has not solved the problem, so I would disagree with you that—

Q145 Mr Chope: Hold on. You say you require authentic identity on Facebook, but if you are using Facebook, you don't know the identity of the person with whom you are dealing, necessarily. That is what I am talking about. If everybody knew when they were using these social media platforms with whom they were dealing, it would be the same as when one looks at a letter in a broadsheet newspaper, and at the bottom it will tell you the name of the person who wrote it. What is the problem about having that transparency, and with it the accountability?

Simon Milner: I understand the point you are driving at, but just to be clear, when there are letters in *The Times*, *The Times* cannot be certain that that person is who they say they are, or that they live at the particular address they have written down. Authentic identity is a very important part of our platform. It is not necessarily about abuse, but more about how people use our platform to connect with people they know. That



HOUSE OF COMMONS

is broadly what Facebook is for, and that is why authentic identity really matters.

We do not suggest that that is the solution that would solve the problem of abuse. I am afraid that there are lots of people who are quite happy to use their real name and to be very abusive towards MPs and other people, and there is no suggestion that they are hiding behind anonymity.

Q146 Mr Chope: You say it would not solve it, but it would certainly reduce the problem because it would deter a lot of people from engaging in this activity. Because they are able to hide behind their anonymity, they behave in a different way, rather like people do in gangs or when they put masks on. Dr Lundblad was saying earlier that people shouldn't behave any differently on social media than they do in their ordinary lives, but they do—largely, in my submission, because of the anonymity. Would you accept that if we removed that anonymity, a lot of these problems would be solved?

Simon Milner: Some of it would, but certainly not all of it. It is not a silver bullet.

Q147 Mr Chope: If some of them would be solved, why don't we go down that route?

Simon Milner: I think that is a matter for the other platforms.

Sinead McSweeney: I agree. Unfortunately, I don't believe it would solve all of the issues. Yes, with some of the behaviours, people engage in them because they feel in some way protected. What we have done on the product side is that you can choose not to see content from accounts or not receive notifications from accounts who don't change their profile picture or don't verify their phone number or their email.

For me, if you take the view that it won't entirely remove the problem, we believe the benefits of having anonymity to us as a platform—we have made the decision—outweigh the extent to which it is used for abuse.

Q148 Mr Chope: What are those benefits?

Sinead McSweeney: We have seen the platform and anonymity on the account used by journalists, by whistleblowers, by human rights defenders, by people working in areas of conflict and also among groups of young people who may have mental health issues.

Nicklas mentioned people who have experienced domestic violence. Because it is an open platform, there are ways in which people can find others who are experiencing the same problems, but they may not want to identify themselves publicly as somebody who is struggling with depression, or panic attacks. We believe all of those use cases are important and we are proud to have a platform that allows people to have those conversations and to flag issues, but not in a way that jeopardises their safety.

Q149 Mr Chope: But surely people could opt in to anonymity, rather than, as



HOUSE OF COMMONS

you are saying, having to opt out of it.

Sinead McSweeney: The fundamental decision we have made is that there is a value to anonymity, which we are proud of; we are proud of having it on the platform, and we adopt other measures to address the abuse and safety issues that arise.

Q150 **Mr Chope:** That is a policy decision. You think that is better.

Sinead McSweeney: Yes.

Q151 **Mr Chope:** As I understand it, in South Korea there is no anonymity and there has to be openness and transparency. What happens in South Korea? Do any of you operate in South Korea?

Sinead McSweeney: We do operate in South Korea. I actually don't know the answer to that question, but I am happy to ask my colleague who covers South Korea and I will come back to you.

Q152 **Mr Chope:** Thank you. What about your colleagues? Do you know about South Korea?

Dr Lundblad: I only know about the policy, and I think it was—I will not guess; I will get back to you with my South Korean colleague's assessment. I think that is better.

I won't repeat what my colleague said, but I would like to add one point. We need, as companies, to hold ourselves to a standard where, even if there is anonymity, we should be able to take care of abuse. That is something that we need to be able to do. It is a rightful claim from the users and creators on our platform to say that.

Again, for the vast majority of the content on YouTube, anonymity is not a challenge. The sliver of content that we are discussing that has real dangers and real problems that need to be addressed as the highest priority, that is the content where anonymity is a problem.

Q153 **Mr Chope:** But you would accept that prevention is better than cure? If you didn't have anonymity, we would prevent quite a lot of this activity, which takes place because of the anonymity that is associated with it.

Sinead McSweeney: Unfortunately, that is increasingly not the case. I think that there has been a change in the way in which we disagree with each other and debate issues, where people, with their full name and photograph and lots of identifying features, are engaging in abusive behaviour. The key point is that whether an account is anonymous, or whether somebody is using their real name, they cannot hide from the rules. They may hide their identity, but they are still subject to our rules; their accounts are still suspended.

We have seen instances in the UK and in other countries where people who work using anonymous accounts to abuse public figures were successfully prosecuted in court. Number 1, increasingly we are seeing people who don't need anonymity to be abusive towards others—but even



HOUSE OF COMMONS

where they are anonymous, they are subject to the rules and they are also subject to the law.

- Q154 **Mr Chope:** Can I give you an example? Somebody recently tried to create a parody account of me. In order to establish that it was a parody account and not genuine, I had to supply Twitter with a copy of my passport—my identity. Why did I have to do that in order for that parody account to be identified as such, when the person running that parody account did not have to disclose to me any ID whatsoever?

Sinead McSweeney: The reason why you had to verify your identity was so that we would be sure that it was you who was reporting the parody account. There are parody accounts on Twitter of some very high profile figures, which I am sure people in this room follow and which are very entertaining, and which the real-life figures have never reported to us—maybe because they also find them entertaining.

I do not want to get into too much detail about which country or region this occurred in, but we have had instances in which we have received competing claims from a political figure as to which account was their real account and which was an impersonation account. Those are the reasons why we would require you, in reporting the parody account, to verify that you are the person who should report it.

- Q155 **Mr Chope:** Okay, but why shouldn't I be able to know who is operating the parody account? Why shouldn't I be allowed to know that?

Sinead McSweeney: Because that user has a right to privacy under our terms of service, and we don't—

- Q156 **Mr Chope:** Because of your terms and conditions, you are giving them the cloak of anonymity. I suggest that, if you didn't give them the cloak of anonymity, there would be even openness and transparency, and a lot of this abuse would go away.

Sinead McSweeney: As I have said, we have made a policy decision that we are proud of; the benefits of anonymity are such that we allow people to be anonymous on the platform. Revealing somebody's identity to a third party is something that we can only do under legal process under data protection law.

- Q157 **Stuart C. McDonald:** I just have a couple of questions on the possibility or otherwise that new technology or innovations can help us out here. First, Dr Lundblad, you have mentioned a couple of times the new "limited state" category that Google is applying to certain videos. I think that basically means that certain features, such as adverts or comments, do not apply there, and that you can essentially only really access a limited state video if you have the specific URL.

How do you decide what goes into that category of video? What effect has that had in the three months it has been operating? Is its potential undermined by the fact that certain websites that gather all of these limited state videos together already appear to have sprung up? In fact, if you do a Google search of "limited state", one of the first results that

comes up is a collection of those videos. In fact, I think one of those websites actually also has a Twitter account.

Dr Lundblad: You address three important questions. Let me take the last one first. There is always the question of the evolving nature of the threat; as soon as we do something, the people who want to abuse our services do something else. We are following that very closely, in order to figure out what we can do.

However, if we then go backwards through the effects, there is a significant effect on these videos. When we limit the features on videos—including by putting them behind an interstitial, removing comments, removing recommended next videos, removing monetisation and all those things—the view count drops dramatically.

The numbers vary somewhat, so we are not sharing a number yet, but that is something that we may be able to do in our transparency report or do later as we see more of the effects. There are now thousands of videos with limited features for which we have done this. The way we decide is by having a set of experts—as we have for all our policy issues—who look at a number of different things.

One thing, which came up after a discussion with the Committee at the last hearing, is what we should do with all those actors who are really toeing the line of our guidelines. They are gaming them and shifting a little bit, to stay just inside being acceptable. We want to clearly signal that that is a behaviour that we want to deter, so we are now working with and experimenting with limiting the features of those videos to see what the effects are. It is early days, but it is very promising, in terms of the view count; in some cases, the view count has dropped so significantly as to make the videos unimportant.

Q158 **Stuart C. McDonald:** Is there any possibility of some sort of similar system being applicable on Facebook or Twitter? Indeed, would Twitter take some sort of action to take down an account that is drawing people's attention back to these limited state videos that Google is trying to potentially hide away a bit?

Sinead McSweeney: If you share the details with me I can take that back to our team. To double on the point, like all bad actors, whether online or offline, as quickly as we learn, they learn and try to go one step ahead. Whether that is criminals being forensically aware, people similarly become aware of the measures that we put in place to protect our users and try to circumvent them. That is why we are never done. Our aspiration is not just to keep up but to try to stay one step ahead.

Simon Milner: There are lots of different elements of the Facebook product. It is not just about videos. We do not call them limited states, but there are things like feature blocks that we might put on particular pages or groups, there are certain pages that will not be eligible to have any advertising associated with them and we also can and do put warnings, or interstitials as they are called, against some content. So there are a



number of things that we do. It sounds like YouTube has something that feels more “on or off”. We have not got something quite as straightforward as that, but we have got a number of different elements.

Q159 Stuart C. McDonald: More generally, we have heard and read a lot about various initiatives like photo DNA, video fingerprinting and hash databases, which have been mentioned already today. I get the impression that with photos and videos, and particularly when it comes to child sexual exploitation and terrorism, where you are clearly so far over any line of acceptability that it is categorically clear, technology can help us there; but the huge issue, particularly with text and the use of specific words, when you get closer to the grey area, is that you find that technology is of very limited assistance.

Simon Milner: It can certainly help, but this is also about behaviour. It is not just about what somebody is saying, but the way in which things are being said. The behaviour of accounts, groups or pages is where the technology can help us drive towards a much more efficient process for our reviewers.

I mentioned before that we do not want our reviewers wasting their time looking at content that is fine. Unfortunately, in the old school of just reporting there was so much noise in that system that the vast majority of reports were of perfectly innocent content that was not infringing our terms and not even close to doing so. What we want to do is use technology to surface to our reviewers content, including speech, that is much more likely to be infringing, and then we can get much more efficient at addressing that content quickly.

Q160 Stuart C. McDonald: The problem you have is that the technology still drives your reviewers to a mixed bag. Obviously there is stuff that has to be taken down, but it also still has them reviewing lots of stuff that is ultimately benign. Then you come to the question, and some folk will say, that you all seem much more concerned about protecting the benign tweets, Facebook posts or videos and leaving them up and less concerned about letting some malign tweets, posts and whatever else stay up. Is the balance right? Have we not got to a stage where you could say that technology has flagged up a particular post, so you are not going to let that appear on a page until it has been reviewed? Why not turn the process round?

Sinead McSweeney: We are fortunate to live in a democracy and we are fortunate to have certain rights. Obviously, with those rights come responsibilities, but in all walks of life it is like that. It is better that one guilty man go free than 10 innocent men be convicted, particularly when it comes to speech and, even more particularly, political speech. Personally, even though I work for a company, I would be uncomfortable with private industry having the dial on how much innocent speech might potentially get swept aside in a quest to remove abusive speech. I think that it is hugely difficult. I would not say that we prioritise keeping up the benign content over getting rid of the malign content, but I think we want to be so good at identifying the malign content that we are not sweeping away.



HOUSE OF COMMONS

It is easier for us to think that—we are in a democracy, we are English speaking—but you have to think of all the other countries and the uses to which rules and reporting mechanisms can be put to distort politics and to silence people. Those are the complexities that we have to navigate when we address this.

- Q161 **Stuart C. McDonald:** Are there ways in which you can use the technology that you have to focus on particular accounts that are more likely to or will have a higher number of posts and tweets and whatever else that cross a boundary? For example, is there any research that suggests that a higher percentage of takedown occurs for new accounts, or accounts that are under a couple of weeks old, or accounts with content that has been taken down previously? Is that a way that you can get more out of the technology rather than having to apply it uniformly?

Simon Milner: You should come and work for us. That is exactly the kind of stuff that our teams are doing, and more. You are very welcome to come and meet them and hear more about it.

Sinead McSweeney: That is the important thing. It is not just about the words. There are other signals, and conduct and behaviours around accounts that are taken into account.

- Q162 **Stuart C. McDonald:** Is it, for example, how old the account is or are new accounts more likely to contravene the rules?

Sinead McSweeney: If there is one thing I learnt when I worked in policing, it is that the last thing that you talk about publicly is how you investigate bad actors and how you detect. I am sorry to go back to “trust us”, because you clearly don’t, but there is something about being somewhat constrained in how we describe how we do this work.

- Q163 **Stuart C. McDonald:** That leads helpfully to my final question. I suspect you will be back here before, but say that in two or three years’ time you are back before this Committee, what will success look like? I do not think anyone expects that we will have a social media free of abuse or hate or any terrorist content at all, but what will success look like? Importantly, who should measure it? Given what you have said about a lack of trust, does it have to be someone independent and who is trusted who decides whether you are making sufficient progress?

Dr Lundblad: I think success looks like reducing this problem to a fraction of what it is today. There will always be challenging content and problems, because bad actors are evolving. This is the point we made at the beginning: the nature of the threat evolves. We need to take care of that and make sure that it is a fraction of where we are today. On the way to measure it, I think that at that point we will have—I am speaking for my company—a transparency report industry standard that can be validated where you can look and see the clear impact that the policies, the technology and the people that we have brought to this problem make in terms of improvements.



HOUSE OF COMMONS

Simon Milner: On a personal level I really enjoy these sessions. I am quite unusual in that respect in our companies. You might think success for me is to keep coming back because these are good and important parts of how we are accountable as organisations. We get to talk proudly about the work that our companies do. I kind of hope that success might be that you have some other companies here, because it is usually us three. There are lots of other online companies, many of whom are not applying the same kind of endeavours that we are. For me, that might be some success, if there are three different companies sitting here in a year's time because you think we are doing a good enough job and you do not need to address us in this way. I know that is optimism, but it is Christmas week.

Sinead McSweeney: I guess my measure of success would be that Mr Doughty would say, "I believe you."

Q164 **Stuart C. McDonald:** It is all very well your coming here to highlight a certain number of posts taken down at such and such a time, but there is still this, "You just have to trust what we are saying." Is that good enough? Shouldn't we have somebody else who is maybe able to look under the bonnet, as it were?

Simon Milner: Earlier on Dr Lundblad mentioned that the European Commission have been doing extensive work around hate speech involving lots of different civil society organisations and reporting content to us with an assessment made of how well we are doing against that. That is a really good example of work where it is not what we say; it is what the European Commission say. I expect there will be more of those kinds of initiatives in the coming years.

Dr Lundblad: Europol will do that same thing on terrorism. There are several third party organisations that validate them, and that is important. You are absolutely right.

Q165 **Chair:** Thank you for your time and patience. I have a couple of quick points that I want to clarify. I certainly welcome what appears to be a significant shift in attitudes towards this, compared with the last time that we took evidence from you in March, and towards the level of responsibility that your organisations need to take, and also the increase in staffing and resources being dedicated to this. I just want to clarify. Both YouTube and Facebook gave us numbers. Twitter, did you give us numbers of staff?

Sinead McSweeney: No.

Chair: Can you do so?

Sinead McSweeney: I will follow up.

Q166 **Chair:** Secondly, to clarify, YouTube has said you are going to do an annual transparency report. Twitter and Facebook, will you be doing an annual transparency report?



HOUSE OF COMMONS

Simon Milner: We have not committed to that yet but you will see more from us on this next year.

Sinead McSweeney: Similarly, we already do twice yearly around Government's requests and we are looking at how we can find a meaningful way of giving numbers around terms of service.

Q167 **Chair:** I urge you to do more. Twitter, you said when you talked about the suspension of Britain First and the related accounts that you looked at both offline and online evidence. Facebook, you have not suspended Britain First or the EDL. Is that because you have lower standards than Twitter or is that because you do not look at offline as well as online content?

Simon Milner: It is not a simple answer. With Britain First, up until recently they were a registered political party that had had people running in elections in the UK, so they were deemed by the authorities to be legitimate in that respect. However, there are clearly issues with their page on Facebook. A number of pieces of content have been taken down. We are obviously reviewing it but we are very cautious about political speech.

Q168 **Chair:** Can I ask a specific point? Do you look at offline behaviour of organisations as well as online?

Simon Milner: We do.

Q169 **Chair:** I have noted some of the things that have changed this year that we have welcomed, and also Mr Milner's point. As I said at the beginning of the March session, we obviously have you before us because you are the biggest and we do recognise that you do far more than other organisations that also need to be challenged. You will also have heard the immense frustration from all members of the Committee about the scale of this problem.

The areas where we would like you to look further, and to hear more from you, include the extent of your proactive searching for all terrorist and violent extremist content, not just ISIL and Daesh. You heard also about Northern Ireland and far-right extremism.

Secondly, those include your proactive activity around hate crimes and abuse of individuals, particularly where there is independent evidence of particular individuals being targeted—who might or might not be politicians—where it is clearly about contributing rather than just waiting for individual reports. The speed of your decision making has been raised time and again, particularly the speed in response to people across the country who are not trusted flaggers or institutions that raise things.

There is also the use of your technology and algorithms, potentially to promote extremist or inappropriate content. I would say for the record that although the examples I used earlier were around Twitter and YouTube, I also have examples for Facebook.

For example, when I do a quick search of the word Islam into Facebook pages, I immediately get very high up pages that have a relatively low



HOUSE OF COMMONS

number of likes being promoted when they are clearly Islamophobic pages, including, by the way, since we started this session, the page that I mentioned at the beginning of the session, which is now coming up as No. 2 on the search list.

Simon Milner: That's probably because you have searched it before.

Q170 **Chair:** That does seem to be a problem; it wasn't at the beginning. We would like to hear more evidence on that. You have also heard more concern about the issue of anonymity. We would also like to have more information from you in response to potential policy changes.

You were questioned about changing the law on whether you should become publishers and be treated as publishers, whether there should be an independent regulator and what other potential policy changes are being put forward. Your response to us is effectively: "We're working on it. We are doing some stuff. We have done a lot of stuff and we've got to do some more."

The reason we are pressing you so hard about this is because it is so important. In the end, this is about the kinds of extremism, whether that be Islamist or far-right extremism, that can lead to very violent incidents. It is about the kinds of hate crime that destroy lives. It is about the kind of harassment and abuse that can undermine political debate and democracy. You are some of the richest companies in the world and that is why we need you to accelerate and do more. Thank you very much for your time.