Joint Committee on the Draft Online Safety Bill

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

Thursday 28 October 2021

6.20 pm

Watch the meeting: https://parliamentlive.tv/event/index/cf4e4a95-fbc3-4779-8e78-12f46acbadfd

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

Evidence Session No. 15   Heard in Public   Questions 233 - 249

Witnesses

[I](): Nick Pickles, Senior Director, Global Public Policy Strategy, Development and Partnerships, Twitter; Dr Theo Bertram, Director of Government Relations, Europe, TikTok.

USE OF THE TRANSCRIPT

1. This is an uncorrected transcript of evidence taken in public and webcast on www.parliamentlive.tv.

2. Any public use of, or reference to, the contents should make clear that neither Members nor witnesses have had the opportunity to correct the record. If in doubt as to the propriety of using the transcript, please contact the Clerk of the Committee.

3. Members and witnesses are asked to send corrections to the Clerk of the Committee within 14 days of receipt.

Examination of witnesses

Nick Pickles and Dr Theo Bertram.

Q233 **The Chair:** Welcome to the final evidence session today of the Joint Committee on the draft Online Safety Bill. Welcome to Theo Bertram from TikTok, who is giving evidence in the room, and Nick Pickles from Twitter, who joins us remotely. Welcome to you both.

Nick, could I start by asking you about a story that was very prevalent in the summer relating to Twitter and other social media companies? It was around racially motivated abuse directed at England footballers. It was not the first time that things like that had happened, and there has been criticism of it before. Indeed, in April, there was a social media strike by Premier League clubs and players. When did Twitter first start discussing in detail with the football authorities how you would deal with anticipated incidents of racially motivated abuse on Twitter during the European Championships?

Nick Pickles: Thank you for raising this. What we saw on the platform was something that we felt was unacceptable, and we took action on tens of thousands of tweets that violated our rules. The collaboration between the FA, the football supporters' groups, the Ministry of Justice and social media companies began back in 2016. It is long-standing collaboration. The information sharing and collaboration is something that we would do ahead of every major tournament. I think, having been at Twitter for coming on seven and a half years, what we saw this summer was unprecedented in the type of abuse and volume of abuse directed at English players compared to any previous tournament that has happened in my time at the company.

The Chair: It was not an unprecedented incident though. There have been lots of high-profile events, particularly of black footballers being targeted for racial abuse after a perceived poor performance or a missed penalty. That has happened in plenty of Premier League games. Indeed, that is why the Premier League took the issue so seriously and spoke out about it at times during the year. You could not really say that what happened in the summer was unprecedented or unpredictable.

Nick Pickles: That is why we have the collaboration in place. It is why we work with the footballing authorities. It is why we work with law enforcement to help them hold people accountable for committing crimes that are directed towards footballers. We have also identified gaps in our policies through that process. Later this year, we are planning on rolling out a stronger policy specifically addressing people who say that someone cannot be from a certain country because of the colour of their skin. That policy gap is one that we are strengthening. That was identified during the Euros as a particular issue where we were not taking action, but certain types of abuse were targeted using that language.

The Chair: Thinking of the England versus Italy match and what was seen on Twitter and other social media channels after that match, if you had systems in place to deal with it, and if you had effective collaboration

with the football authorities, why was it that terms of racial abuse were found among trending topics on Twitter that night?

Nick Pickles: I have heard that report. First, there are certain phrases that are simply not allowed to trend, so it is not a question of if they trend; we have systems in place to stop them trending. We saw some of those reports, and I have to say we were not clear where they were coming from because, technically, it is not possible for certain phrases that were discussed to have trended, so we can categorically say that they did not. We have not seen, for example, screenshots of things. That is an issue. On the type of content we took down, more than 1,900 tweets were identified proactively by our systems following the final. Just over 100 were identified through user reports. The ratio of proactive technology to user reports in the aftermath of the final was heavily weighted towards action being taken because of proactive investment.

The Chair: I will not ask you to list in a public session what those phrases are. If you were able to write to us, we would be happy to receive that. I saw a screenshot of the N word being used as a trending topic. Would that be banned?

Nick Pickles: I am not sure if the screenshot was of someone saying that it was trending rather than the trend itself. We occasionally see people who, for a variety of reasons, photoshop and create images that they share online to garner a reaction. That word in particular is on that list. Worldwide, it cannot trend.

The Chair: Okay. You are saying that your policy would be that that word could not trend and could not be used at all as a trending topic on Twitter.

Nick Pickles: Both the policy and the technology we have in place to stop it.

The Chair: Okay. Various people have given us evidence on that, and they are more than welcome to send us examples they may have that we can cross-reference against your checklist to see if that is the case. What enforcement action was taken against accounts that engaged in that sort of abuse?

Nick Pickles: Generally speaking, the action we take is at tweet level. If the tweet violated our rules, we would take action. If it was one of multiple tweets, we would suspend the account. If it was a one-off, we would look at the severity and the previous behaviour of the account. For example, has it broken other rules before? It can be a combination of things. It can be a first offence, permanent suspension because it was a violent threat or it can be one so-called strike against the account that we take into consideration for future violations.

The Chair: For racially motivated hate speech, how many strikes do you need to lose your account?

Nick Pickles: It depends. If someone uses violent language, it is one strike. That is one of the challenges. Depending on the context of the speech itself, the policies have different thresholds. This is an area where we have started publishing the strike system because we think it is important for people to understand. In the case of civic integrity and Covid misinformation, we tend to look at five strikes. In other policies like that, we are more aggressive.

The whole concept of strikes is something that, as user behaviour has changed, we need to look at more holistically because people having lots of violations for different policies are clearly unhealthy users, so the strike system allows them to stay on the platform longer than we want. This is a good example of a system that we are looking at for how we can change it.

The Chair: Do you think the strike system is effective? There is a good chance that a lot of this abuse came from accounts, let us say, that have a very short life span, act in a co-ordinated fashion, are probably automated and are often based abroad. I have seen evidence that a lot of the racially motivated abuse directed that night at the England players did not come from the UK but came from networks based overseas. Is a strikes policy effective if it is being done at an industrial scale by large networks of accounts that may be very temporary in their nature?

Nick Pickles: I have to say, from the data that we have from the analysis we did, the overwhelming majority of people engaging in this abuse were real people and they were in the UK. I do not think that is easy to hear. Sometimes, if we say it is an automated problem, it changes the nature of it from being less social and more technical. These were real people who were in the UK.

You are right. In the strike system more generally, if someone is running multiple accounts, there is no point giving all their multiple accounts one strike; you want to suspend the network. It is a useful tool. It is a good example, as the Bill is going through, of how you design processes. One of the things that we have tried is to send people a message before they send the tweet to say, "This might be hurtful and might be reported. Are you sure you want to send it?" That intervention is not a strike, but it stops the tweet being sent. Protecting the ability to innovate with different safety technology is something that is really important in the Bill.

The Chair: But you take enforcement against known networks of accounts where multiple accounts engaging in abusive behaviour are being controlled by a common source.

Nick Pickles: We do that quite commonly. We also, unfortunately, now see this activity linked to states, not in the Euros case but more broadly. If someone is running multiple accounts to deliver abuse, we suspend the whole network. We look at the technical indicators. That is something we can do using behavioural signals rather than always having to wait for a tweet to be sent. Those actors often use technology to evade our system,

so it is a very adversarial space. If someone is running a network of fake accounts that engage in any activity, we take them down because they are fake accounts in a network; we do not wait for a specific policy violation.

The Chair: Say someone was using a common phone number to run multiple accounts. If one account was taken down for being abusive, would you take action against the other accounts as well?

Nick Pickles: Occasionally, there are circumstances. As Members of Parliament, you might have members of staff who have phone numbers on your account. There are sometimes circumstances where we would look at whether there are two different accounts and one was personal and one was someone they work for, for example. Phone numbers are a very good signal that we use—one of many—to identify, for example, people trying to return to the service. Phone numbers can be gamed. You can buy SIM cards. You can use online virtual services. They are not foolproof, but they are a good signal to use for some of these systems.

The Chair: I understand that it is very easy to get fake phone numbers on the internet to set up fake Twitter accounts, and that is part of the problem. One of the issues we have to grapple with as well is cases like this where, if someone is abusing someone else and they are doing it in their own name and they are traceable, it makes it easier for law enforcement to take action, but if they are not doing it in their own name, that obviously makes it harder. What are your policies in complying with requests from the police for data and information about abusive accounts?

Nick Pickles: There is a long-standing legal process to be followed. One of the first questions is whether the individual is in the UK. Sometimes you have situations where UK law enforcement asks for data for someone outside the UK. That is a legal challenge for us. The UK and the US have worked to resolve some of it with the CLOUD Act. If law enforcement comes to us and says, "This account targeted a footballer or a British citizen. We are investigating a crime. Here is the Investigatory Powers Act paperwork", we review that, and, as long as it has things like the account name right and the technical information, we produce on those requests. Generally speaking, we receive around 1,000 requests a year from UK law enforcement, and we work with them very frequently to improve the quality of those requests and the speed and timeliness of them.

The Chair: How long does it take to respond? Where the information will be produced because the correct application has been made, how long does it take?

Nick Pickles: We have an emergency process. Without giving you a simple number, if someone comes to us, for example, on a counterterrorism investigation, there is a dedicated process that takes, generally speaking, minutes. It is a very fast process that we have a team of people working on globally. If it is more routine, I would not

want to give you a specific time, but it is certainly not something that I would expect to take more than a few working days.

There is occasionally back and forth. One of the challenges we have is that sometimes different police forces around the world have different levels of training, so they may wait to come to us and ask for information. One of the challenges we sometimes find is that, if that delay is too long, we may have deleted the information because we did not know there was a legal request coming. We have a preservation system in place so that if a complex investigation is happening and law enforcement says, "We think we might need this but we do not have the request ready", we can hold on to the data for longer if they need it.

The Chair: Okay. From what you are saying, a valid request made by the UK authorities for accounts in UK jurisdictions should take only a matter of days to process.

Nick Pickles: If they submit 100 requests on the same day or one request, that will impact it, but it is certainly not something that should take months in that case.

The Chair: But an individual request should just take a matter of days.

Nick Pickles: Yes, it should be a relatively fast process.

The Chair: One of the issues that we are concerned about is not just the activity of individual accounts but the whole pile-on effect that we see on social media, particularly on Twitter where people receive a lot of abuse within a very short period of time. Does Twitter have policies in place to mitigate that, to dampen that effect down, in order to try to protect people from a high volume of abuse? We have heard a lot of evidence about that, and the personal impact it has had on people who have received it.

Nick Pickles: There are a couple of different ways we approach that. One is policy. If you call for other people to dog-pile somebody and say, "Go attack this person", that is a policy violation. We have products in place. We recently launched something called safety mode, which is using the signal that you are getting a lot of incoming noise to start autoblocking and taking some of the burden off the individual.

We have had to extend our policy framework to something that we have called co-ordinated harmful activity, which—to your previous point—is that they are real people, not fake accounts, who operate in a loose network. The most high-profile example of that has been QAnon, and one of the challenges is that the co-ordination happens off Twitter. We do not see a call to action; we just see accounts moving around, perhaps sharing personal information and attacking in a co-ordinated way. We suspended more than 70,000 accounts as part of the QAnon suspension. That is the kind of policy we need because bad actors will change their behaviour to try to find different ways of attacking people.

Q234 **The Chair:** Thank you very much. Theo Bertram, on a slightly different

topic, I saw a report in the US press that TikTok had changed its privacy policy and was collecting biometric data from users, including face screens and voice prints. Is that something that is being done in the UK as well?

Dr Theo Bertram: Thanks for having me. I am very grateful to be here today.

No, our privacy policy has not changed. Our UK/EU privacy policy is very clear, and you can read it online. It says we do not collect biometric data. The confusion around it arises from different definitions of what is considered biometric by the US and the EU.

The Chair: In the context of the US, are you able to say what TikTok is gathering the information for?

Dr Theo Bertram: You can think about TikTok like this. When you are creating videos, you might have an image of a mask that, when you move your face, moves with you. The idea of where a face is on the screen is something that might be counted under US law but not under EU law. That is not the same in terms of whether we are collecting facial identities or things like that. We are not doing that.

The Chair: In America, where TikTok is gathering facial data—

Dr Theo Bertram: No, just to be clear, we are doing the same across both sides. There is a different definition within the law of what constitutes biometric data. In both places, we are not collecting facial identity or anything like that.

The Chair: What are you collecting?

Dr Theo Bertram: I do not know if you have done this, but there are lots of effects on TikTok when you are videoing yourself. There is one I have done where you are a dragon, and you move around and the dragon moves around and you have a little dragon on your shoulder. To do that, the phone needs to know the shape of your face and where it is moving to be able to project the image as it moves on the screen. We are not collecting that data; it is just to facilitate the video to be able to do that. Under EU law, that is not considered biometric data because we are not collecting it. Under US law, as I understand it—I am not a privacy lawyer—there is a slightly different definition. That is where the difference lies.

The Chair: I understand what you have explained the usage for. Is that data being collected off the device and stored off the device by the company to help it with the development of its products and innovations of its products?

Dr Theo Bertram: No, we are not collecting facial identity or anything like that.

The Chair: Okay.

Dr Theo Bertram: We have a list of all the things that we collect under our privacy policy. I gave it previously to the DCMS Committee. I am happy to give it to you as well.

The Chair: You do not use data like that for the age profiling of users.

Dr Theo Bertram: No.

The Chair: What sort of age profiling work do you do?

Dr Theo Bertram: When a user first wants to download the app from the App Store, first of all, you have a prompt in the App Store asking whether you are over 12 or not, whichever store you use. Then, when you open the app, you have to give an age, and we do not prompt you as to what that age must be. Even if you pass that age gate, whenever any video is reported for any reason on the app, a moderator will look at the creator of the video to see whether the person is 13, and we use AI to try to detect whether the creator or viewer is under 13.

We are the first company to routinely publish the number of under-13s we remove from the platform. We removed 7 million in the first quarter of this year and 11 million in the second quarter of this year. We had a debate internally about publishing those numbers, and whether they might make us look bad because of the volume, but we felt that being transparent was the right thing to do.

The Chair: Those are the numbers you removed. What would your estimate of the total numbers be if the regulator came to you and said, "Can you tell us how many underage users you think you have?"

Dr Theo Bertram: That is all the underage users that we know, because we removed them. That is why we are publishing that number.

The Chair: Is there a larger universe? If there is some sort of review? Do you have a system of flagging accounts that are likely to be suspicious that you work through? Seven million is the number you have removed, but there could be a bigger number that is subject to review.

Dr Theo Bertram: There are certainly people we remove because we believe they are under 13, who subsequently turn out to be over 13, and there is an appeal process for them. I do not have the exact numbers here, but if that is important for you I can share that with you. That is our estimate of the number of under-13s, because as soon as we find them we remove them.

Q235 **The Chair:** Is the moderation of content for UK users done by UK moderators?

Dr Theo Bertram: It will be done by moderators everywhere. If a moderator can spot that someone is under 13, it does not matter where they are. We have UK moderators as well as global moderators.

The Chair: Would a moderator in China be effective at understanding whether a young person in the UK is 12 or 13?

Dr Theo Bertram: TikTok and the Chinese business are separate things. You are not going to have Chinese moderators; you are going to have TikTok moderators, which is the rest of world, because TikTok does not operate in China. Our head of trust and safety is based in Dublin. That is our headquarters in terms of where our moderation is led from globally. We have moderators all around the world because we need them in each of those languages.

The Chair: But you do not have moderators in China.

Dr Theo Bertram: Not for TikTok. TikTok is separate in that way.

The Chair: In terms of the regulator's jurisdiction here, would that be the same for engineers who work on TikTok? You do not have any engineers based in China.

Dr Theo Bertram: You have the TikTok product and the TikTok app, which is served by the UK business here for the UK. The data is stored in Singapore and the US. Our approach, generally, is that we want data to be stored by region. We do not have a data centre in Europe yet, but we will next year, and that will be in Ireland. UK data will be stored there. In terms of data transfer between regions, we try to minimise that as much as possible. There are restrictions on who can access data and what kind of data they can access, and that is very strictly controlled. Just to reassure you on the idea of whether there is some sort of access to data from China, that would be extremely strictly controlled, and only in the cases where it was done under strict control.

The Chair: If you were a project designer or engineer working in China on TikTok, you might be allowed access to data as part of the development of a new product under specially supervised conditions. Would that be the case?

Dr Theo Bertram: Data does not—

The Chair: I appreciate that data is not leaving the data centres. You said that there may be circumstances where perhaps someone in China had access to it.

Dr Theo Bertram: To make the product function and to make the product safe, we have people all around the world. There will be some engineers in China as well, potentially. The rules on who can access data from anywhere in the world across regions are strictly controlled. You can access data only if you have the right permissions to access that data, and only for a limited time and only for the correct data that you need to access. All of that is strictly supervised and audited. We know that there are concerns for all of these processes in general. We are probably the app that has been the most scrutinised in 2020. These issues have all been very thoroughly investigated by external experts and national security teams, and when they have had a look at how we work and operate and what those processes are, they have come to the conclusion that there is no national security risk at all.

The Chair: Okay. I was not talking about national security—

Dr Theo Bertram: No, I was anticipating your line of questioning.

The Chair: I am reassured to know that is all right. In the scenario you describe, which could in theory exist, where an engineer in China has access to UK citizens' data that is being held in Singapore in order to complete a time-limited piece of work, who signs off on that and says that the systems are adequate?

Dr Theo Bertram: The global head of security—

The Chair: For TikTok?

Dr Theo Bertram: Yes. It is Roland Cloutier. He is based in the US. He is a former US state employee so is an extremely experienced head of security. He designs those operations and is in charge of making sure that access is always restricted.

The Chair: So it is an internal process.

Dr Theo Bertram: We have external validators as well. We have made the commitment that we are happy for there to be external validation, both by regulators and by third parties that a regulator might want to nominate to come and examine what we are doing and make sure that we are doing it in a way that satisfies their concerns.

The Chair: Are you able to give an example of the sorts of reasons that could be given that would allow a Chinese engineer to access that data?

Dr Theo Bertram: Keeping the platform safe in trying to design the product in a way that is going to make it safer for users. I can give—

The Chair: It sounds like any product design feature. Presumably, you only have people working on things to make it safer and to make the user experience better, but any product-related reason could be a reason for accessing the data.

Dr Theo Bertram: The process would be that they would need to have a clear reason why they needed that access. That would need to be cleared, and it is strictly audited. It would be restricted. It is not the case at all that there is some sort of way of accessing data that is unsupervised, unverified or unaudited. It is controlled.

The Chair: This is principally not a data Bill, although there are overlaps between data security and the concerns of the Bill. If the regulator had concerns about the safety of a particular product, it could well be that the people who designed that product are based in China rather than in the UK.

Dr Theo Bertram: In terms of the product design, they are based all around the world. In terms of this Bill and the regulator, of course, they can have access to us. We are more transparent than any company when it comes to how our platform functions. Some of the members of this

committee have already attended our transparency and accountability centre virtually where you can meet our moderators and see how TikTok works. When we have the physical centre, we will also enable you to come and see the algorithm for yourselves. Regulators and academic experts can come and inspect that so that there can be certainty about the safety of our data and all the steps that we are taking to keep user data secure.

The Chair: Following up on that, you said that TikTok has engineers all around the world. What proportion of the engineering team is based in China versus the rest of the world?

Dr Theo Bertram: I am not sure of that off the top of my head.

The Chair: Ballpark?

Dr Theo Bertram: We will follow up with you on that. I do not want to give you a guess. I do not want to give you an estimate.

The Chair: Would it be reasonable to assume that it is more than half?

Dr Theo Bertram: Probably, but I will give you the updated figures. We are certainly growing the engineering teams outside China as well.

The Chair: Okay, thank you.

Q236 **John Nicolson:** Perhaps I could start with you, Dr Bertram. You will remember that you appeared before the Digital, Culture, Media and Sport Select Committee last December. You and I had a bit of a chat. We talked about Covid disinformation, and you told us how great TikTok was at taking it down. We managed to find within seconds somebody who had 600,000 likes for Covid disinformation. Do you remember her?

Dr Theo Bertram: Yes, I remember.

John Nicolson: She was not hard to find because her hashtag was "anti-vax", so it should not have taken anybody very long to find it. You managed to find it. Only when we told you about her did you take her down, and you did that during the course of our conversation. You are not doing much better though, are you, in the intervening year on taking down anti-vax stuff?

Dr Theo Bertram: Actually, we have improved the rate at which we are taking down Covid misinformation. I think the numbers are in the sheet I handed round: 83% of Covid misinformation is removed before it is reported, 87% within 24 hours and 70% before a single view.

John Nicolson: So how come children as young as nine are getting targeted by anti-vax stuff on TikTok?

Dr Theo Bertram: I saw that news article, and I think there are two parts to that. First of all, children as young as nine are not allowed on our platform. I have explained the steps we take to keep those under 13 off our platform. None the less, no one should see Covid misinformation on

our platform. We had our first Covid misinformation policy in place two weeks before the UK went into lockdown, so I do not think you can say we were unprepared. We have been developing our Covid misinformation policies around anti-vax and so on ever since then.

John Nicolson: You mentioned the article. It was an investigation by NewsGuard. Let me quote exactly what it said: "The barrage of toxic content—including videos asserting that the vaccines are deadly and that COVID-19 is a genocide conspiracy—came even though some of the children", some of whom were as young as nine, "did not follow a single account or search for specific kinds of information." They were being fed the content without even prompting.

Dr Theo Bertram: The videos that NewsGuard shared with us, most of those we have removed—

John Nicolson: After they shared them.

Dr Theo Bertram: Some of them we have not. One of them, for example, was a video of a young US comedian and was projecting to the future, "Blade Runner"-style, with the idea of a zombie coming towards him. It was clearly humour.

John Nicolson: The zombie coming towards him was somebody who had had a vaccination—

Dr Theo Bertram: Yes.

John Nicolson: Because he was proudly boasting that he had not had a vaccination. He was a lone survivor—

Dr Theo Bertram: No, that was not the case.

John Nicolson: All the folk who had had vaccinations had all become zombies.

Dr Theo Bertram: The good thing about this Bill is that it will take that out of the hands of companies like us, which I know you do not trust any more—I understand that—and put it into the hands of Ofcom and the Secretary of State. That is a good thing. If Ofcom says, "Okay, no jokes about Covid", that is what we will do, but at the moment we have to make those judgments ourselves. We allow satire around Covid, but if you feel, as the Commons and as Ofcom, that we need to go further than that and shut down freedom of speech because of the risk, I would understand that.

John Nicolson: Except it was not just satire; I think a lot of people thought that it was anti-vax. Setting that aside for one moment, that was just one. Lots of the other stuff that you took down, you took down only because it had been highlighted to you by this investigation. What I am saying to you is that whether it was me last year at another committee pointing out to you somebody with hundreds of thousands of views, or whether it is the investigation by NewsGuard, you always seem to be on

the back foot. You are always responding to what other people are telling you, rather than you and your organisation looking for this stuff effectively, because it is not hard to find. If I can find it as an individual, you, with all your company's resources, should be able to find it before me, and it is not even my full-time job; it is your full-time job.

Dr Theo Bertram: That is the nub of the whole debate and what the Bill is trying to get at. The big, fundamental challenge we have is that, across all these platforms, there is a huge amount of users at scale, and when it comes to harm, whether it is Covid or racism, getting rid of most of the stuff is straightforward; it is the bit where nuance and context are required, and how you do that at massive scale, which is difficult.

John Nicolson: Tell me again how it is nuanced to post a video that says that injecting yourself with a Covid vaccination means you are injecting yourself with dead babies? How is that nuanced when the hashtag is "anti-vax"? Where is the nuance? I do not see the nuance.

Dr Theo Bertram: No, I agree.

John Nicolson: Then find it.

Dr Theo Bertram: That would be straightforward to remove.

John Nicolson: But you do not.

Dr Theo Bertram: As I said, we removed 83% before it was reported, 87% within 24 hours and 70% before a single view. It is not 100%.

Q237 **John Nicolson:** It is certainly not 100%. Let us move on. Do you know what money muling is?

Dr Theo Bertram: I saw an article about it this week.

John Nicolson: Can you explain what it is, or shall I explain what it is?

Dr Theo Bertram: Why don't you tell me?

John Nicolson: I will explain what it is. Money mules are recruited, sometimes unwittingly—if they are children, it is unwittingly—by criminals to transfer funds obtained illegally, stolen funds, into their accounts. They say to these kids, "We're going to put £10,000 into your account. If you transfer £9,000 to a third-party account abroad, you can keep £1,000". An investigation by *City A.M.* found: "Social networks like TikTok are now full to the brim of posts promising young people quick ways to make cash". Do you know the kind of effect it has on a young person's life if they do that?

Dr Theo Bertram: I read the article. I think it was referring to money muling but also to a wider set of financial concerns and the idea that users were seeing videos from creators that were about crypto or get-rich-quick schemes. We have strict rules, certainly for our advertising but also for branded creators, and we have rules around community guidelines on what you can post. If you think of our advertising, for

example, we do not allow ads for crypto. Anything we would see as being get rich quick, we would file—

John Nicolson: So why are you failing to detect these money mules?

Dr Theo Bertram: I have not seen the evidence that *City A.M.* has for that claim.

John Nicolson: Did you phone them up and ask for it straightaway?

Dr Theo Bertram: I did not; I was preparing for this committee.

John Nicolson: If I had been preparing for this committee, I would have thought, "Chances are these guys are going to ask me about this. I think I'll pick up the phone. I'll talk to *City A.M.* and say, 'That is deeply disturbing. Tell me about it'".

Dr Theo Bertram: We are working with the FCA. We work with the other agencies as well to make sure that we can take financial harm off our platform. We will look in more detail at that.

Q238 **John Nicolson:** There is a recurring theme here, I have to say: being passive and waiting for other people to do the investigations and provide the evidence.

I would like to move on briefly, if I may, to Mr Pickles at Twitter. My heart sinks as I ask you this question. I think everybody agrees that your policies about abuse are fine. The problem is that your enforcement is absolutely hopeless. Do you know what a greasy bender is?

Nick Pickles: I think this relates to the question you asked me at a previous committee hearing—

John Nicolson: Yes, it does. It is still there.

Nick Pickles: It relates to an account that we suspended. There were some issues in that case, but the account in question was suspended. I know we discussed following up afterwards, and I am not quite sure what happened there. I am fully aware of the context of that phrase, and the account was suspended.

John Nicolson: For people who have not followed me droning on about this, it is something that I was called on Twitter. I wrote to Twitter. I said, "This is offensive and homophobic". Twitter wrote back saying, "No, it's not. It's quite lovely". I wrote back saying, "No, it is really offensive, and here are your guidelines that specifically say you cannot target people with homophobic abuse", and Twitter wrote back to me once again saying, "No, that is not homophobic abuse". Is it a problem of language? We have heard about this at Facebook. Facebook is unable to determine the difference between American English and UK English. Is that the problem, really? You have folk monitoring there who just do not know what greasy benders are. If they knew, which I told them, they would take it down, so that is not really an explanation. Is language a problem for you and your monitoring?

Nick Pickles: Context and local phraseology that is used in different countries is something we train our moderators on. This was a miss, but we suspended the account. Actually, we had suspended the account before you and I had the exchange in the previous committee. The system is not perfect—

Q239 **John Nicolson:** You might be right, but somebody said to me that it is still up. I will happily stand corrected if that is not the case.

Are you allowed to pretend to be a real person, and post a photo of someone else and pretend to be a professional in a profession that you do not actually work in?

Nick Pickles: There are probably a couple of different issues there. You are not allowed to pretend to be someone else in a way that is deceptive. We allow things like parody accounts and accounts for fan clubs to run. You can be the—insert pop star here—fan club. You can pretend to be a public figure. There are many high-profile ones. Peter Mannion, the MP from "The Thick of It", is on Twitter. He is not a real person.

John Nicolson: I know all of that. I follow him. I keep seeing an account in Scotland where someone is pretending to be a teacher. They acknowledge that the photo they are using is not actually them, and they acknowledge that the name they are using is not actually their real name. Is that acceptable under the Twitter rules?

Nick Pickles: The key question would be whether they are being deceptive. It sounds like that person is being open about the fact that the name they are using is not their name. It is unclear from what you are saying whether they are pretending to be someone else or they have created a fictitious identity. You do not have to use your real name on Twitter. In significant parts of the world, that is something that protects people's safety and enables them to participate in a conversation when the issue is that they might not be able to do so. We allow you to use a pseudonym, but we do not allow you to deceive others that you are somebody else.

John Nicolson: You can use a pseudonym, you can use a false photo and you can pretend to be a professional in a profession that you do not actually work in. You are allowed to do all of those under the current rules. Do you think that is tough enough? Do you think that protects people enough, or do you think that we, as a committee, should do what we have heard Twitter would like us to do, which is to set down a clear set of rules and enforce them on your behalf?

Nick Pickles: From a quick look through some of the evidence, clarity is a word that comes up from a range of stakeholders, not just Twitter. This is a good example. There are big policy choices to be made about people using the internet in a way that is deceptive. Our policies protect against that.

Do you want people to use only their real identity online? That is a big public policy question. We think there is a lot of value in having

pseudonymous accounts and in allowing parody and satire. Those are the kinds of decisions on questions of journalistic content and democratic importance content, and the kinds of questions where we might see a piece of content that we would otherwise remove but was, for example, posted by a candidate for office or someone who self-described as a journalist, that it is unclear to us in the Bill how you would resolve. We have taken down content posted by journalists of the Christchurch attacker's manifesto. We think that was the right thing to do, but it is unclear to us how we resolve some of those tensions without much more clarity on the definitions.

John Nicolson: Okay, thank you.

The Chair: Thank you. Debbie Abrahams.

Q240 **Debbie Abrahams:** Thank you very much. My first questions are to Mr Pickles, if that is okay. I want to talk with you about the minimum age of use. I go to primary schools regularly. I went to one last week. One of the questions I have been asking recently is how they use social media and what they use it for. I would say that probably about a third of the kids I speak with—it is usually year 5 and 6 kids—say that they do. I know that your minimum age of use is 13. Given that I just have been able to search for porn on Twitter while I have been listening to you, how is that compatible with a minimum age of use of 13?

Nick Pickles: Thank you for raising this. How you verify age online is a really big question that this Bill can look at. If you look at the Ofcom data, for example, Twitter does not even feature in the top 10 platforms used by young people. One of the questions we have to look at is about the people using our service and how we balance the protections in place.

We have tools like safe search, which apply if you are not logged in. If you are just using Twitter, that protects the content that is coming through search. It is an issue where you have free expression on one side and important protection issues on the other. We, as a service, look at the people using our service. You have two very different companies before you today: a company that is overwhelmingly used by older people and a service that is overwhelmingly used by younger people. The challenge for the Bill is how you strike a regulatory framework that allows different services to have different rules reflecting who is using their service.

Q241 **Debbie Abrahams:** That did not really answer my question, but I will leave it there for others to follow up. I will come to Dr Bertram in a moment.

My next questions are specifically around corporate governance. Who is currently responsible for risk assessment within your organisation?

Nick Pickles: For Twitter, we have a risk committee of the board, which does that from a corporate governance point of view. We also have a chief legal officer who has a team of people working underneath her to do a range of risk assessments. We also do risk assessments as part of our

product teams and research teams. There is a range of people. Ultimately, the chief executive officer, Jack Dorsey, is responsible for running the company and making the decisions that affect how our business operates.

Debbie Abrahams: How regularly do you produce a risk assessment, and how will that change with the implementation of the Bill?

Nick Pickles: We do not produce a single risk assessment; we produce a range of different ones. In the case of every single feature that has, for example, a personal data issue, we produce data protection impact assessments where required. In the product space, we do a broad risk assessment identifying potential misuse and harms that we should mitigate. That work is ongoing. The question is how much of that should be in a single document that we provide to a regulator, versus the expectation that every single time you launch a product, an individualised risk assessment is produced. That could be clearer in the Bill. Then we have the highest level of corporate governance where several times a year a more umbrella document is produced for our board to help them understand the bigger corporate governance risks.

Debbie Abrahams: I presume that, if it is going to the board and so on, you will have a document with details of the different risk areas and your actions to mitigate those risks. Do you do that annually? Will there be significant change with the Bill?

Nick Pickles: The important thing for the Bill is to do risk assessments more commonly on more specific things, rather than focus on one very large risk assessment that tries to encompass all the risks our business faces. Certainly, with the UK regulator's issues in mind, some of the risks in the UK for us are very different from risks that we might face in, for example, authoritarian countries.

It is about being specific about what the regulator will find informative. Is it related to product and policy? Is it related to research, or is it related to our overall approach to compliance with legal requests, because there is recognition of illegal content as part of the Bill? We would certainly expect a lot of that process to be coming from UK courts, UK law enforcement and the regulator.

It is probably not a very good answer, but the challenge is that we could provide a huge number of individualised risk assessments, or we could provide something more comprehensive that might be less detailed. What still has to be fleshed out is which way the regulator would find most useful.

Debbie Abrahams: Okay. Going back to the questions the Chair asked at the beginning, did you have in your current risk assessment the potential for the online pile-on of racist abuse against English footballers in the Euros?

Nick Pickles: Looking at big picture risk assessment, when we launch features, do we assess the risk of things like dog-piling? Yes. That is why we have a policy against calling for dog-piling. It is why we have a policy on things like co-ordinated harmful activity. As for a specific risk assessment on, for example, a single game of football, I do not think we are doing risk assessments on that granular a basis.

Some of these issues were areas where we identified gaps in policy and are going to address them. In other areas, specifically the use of emojis, I am not sure that a risk assessment would have got down to the granular questions about which emojis in an emoji library may be used in an abusive manner, but that was certainly one of the things we had to react to very quickly. If you are thinking about future risk assessments of new product features, as well as looking at text-based content and image-based content, we now have to include emojis in risk assessment. The interesting question there is, if we see a problem with a particular emoji being used in a racist way, how does that then flow into the risk assessments done by the phone manufacturers and OS manufacturers building emoji libraries that are part of the device? How a risk assessment from one business flows into a risk assessment for the whole sector and whether that should go through the regulator or through companies is a good question to ask.

Q242 **Debbie Abrahams:** Okay. Thank you very much, Mr Pickles. Dr Bertram, would you care to answer the questions on age verification?

Dr Theo Bertram: Sorry, which question?

Debbie Abrahams: Specifically around age verification and what TikTok's policy is on that. TikTok, I have to say, is the one that comes up all the time when I visit primary schools.

Dr Theo Bertram: With primary schools, I would definitely be keen to take up Baroness Kidron's offer of there being some sort of government-led information for teachers. We are doing a project with the South West Grid for Learning to try to train teachers. It would be helpful from our perspective if there was a clear message in primary schools that TikTok is not for anyone under 13. It would be good for that to be part of a wider group of companies as well. Do you want me to go back through the explanation I gave at the start of how we keep kids off the platform?

Debbie Abrahams: No, that is absolutely fine. Given that I had asked Mr Pickles, I did not want you to feel left out, Dr Bertram.

Dr Theo Bertram: Okay.

Debbie Abrahams: I am more than happy to move on to corporate governance issues.

Dr Theo Bertram: Sure.

Q243 **Debbie Abrahams:** Specifically, who is currently responsible for risk assessments at TikTok?

Dr Theo Bertram: Let me answer that question specifically, then, if I can, I will give a broader view on how we build safety at TikTok. As Nick said, the way I read the Bill, the model is GDPR and data protection impact assessments. You should be doing a risk assessment when you are designing the product or when you are making a significant change. That will happen with product teams doing that, and then it is accountable to someone who, in this case, would be a named person for Ofcom, based on the way the Bill is designed.

A question from one of your colleagues in the committee to the others asked about the process for looking at this Bill in general. I am part of the European leadership team. We have a risk assessment. Both the OSB and the DSA, the European equivalent, are on the list of things we need to be taking into account. Both at the macro and at the micro level, that risk assessment is built in.

We have disadvantages from being a much younger company than the other platforms you have talked to today. I heard Will Perrin when he came and talked to you. He described this era when tech companies were seen as "magical beings". We have never enjoyed that period of time. It was long past by the time TikTok was created. When TikTok was created three years ago, we knew that we would be in an environment of much greater scrutiny and much higher regulation. We have built trust and safety right at the centre of everything we do, whereas for older companies it has been a more gradual add-on to what they are doing. For us, trust and safety is right at the heart of what we do. The trust and safety lead is based in Dublin and reports to our general council, which reports to the CEO.

Debbie Abrahams: This is a similar question to the one I just asked Mr Pickles, although not event specific. Are issues of underage use included in your current risk assessment?

Dr Theo Bertram: We are absolutely committed to getting younger users off the platform. Part of the risk assessments in general would be whether it was increasing or decreasing the likelihood of an underage user.

Debbie Abrahams: What are your top three?

Dr Theo Bertram: If we are trying to anticipate what will be in the risk assessment designated by Ofcom, we probably need to wait and see what you decide the priorities will be. Certainly, we want to look as broadly as possible at safety and think about it holistically.

One of the things that this Bill does that nowhere else has done before, and is ground-breaking, is the concept of the developmental stages of childhood. We are talking about 12 year-olds and how we need to keep them off the platform. Then there is the sense of whether we just look at this in a binary way, as children and adults. One of the things that Baroness Kidron has led on, and the Bill is ground-breaking in doing, is

seeing that we need to think of different stages of development, and if we treat teens like primary school kids, that is not good for the teens.

We are doing that in our product design. You can see it in the way that we have made privacy by default for those under 16. When you create a video on TikTok, if you are under 16, the people you are going to share it with will only be the people you have chosen; it is not going to be widely shared. Things like direct messaging are not available for those under 16. There are things you can do in product design that take in the principles of this Bill, which are really ground-breaking and important, and set a great precedent for other countries.

Debbie Abrahams: Thank you. I am trying to understand where you are now and where you need to be. You have said that safety is really important for the company. I am trying to understand where you are now and what organisational change you will have to develop to get to where you need to be under the Bill. What are the top three risks that you have currently identified?

Dr Theo Bertram: You can look at the risks that we have identified on our platform in our enforcement reports, which we publish quarterly. At the moment, we are looking across a broad range. I can send you the list of how we categorise each of those harms so that you can see how we are thinking about it. Certainly, child safety is right at the top.

Debbie Abrahams: Okay, that is very helpful. Thank you so much, Dr Bertram.

Q244 **Dean Russell:** Can I come to you first, Mr Pickles? I am hopefully going to ask you a very easy question to answer. I hope you will answer it in the way I would hope. If you or Jack Dorsey were in the room together now and there was a young child—perhaps a nine year-old boy with cerebral palsy—under attack, would you step in to help them?

Nick Pickles: I am assuming you mean physically in the room.

Dean Russell: Yes. If you saw a young child, perhaps someone with cerebral palsy, with their life potentially under threat, would you or Jack Dorsey step in to save them from the people who were attacking them?

Nick Pickles: I would certainly want to do my best to intervene.

Dean Russell: You would intervene, yes?

Nick Pickles: Without knowing any of the context—

Dean Russell: If there is a young child in the room and they are under attack, and you have the ability to step in and stop them being injured or harmed, you would step in, I would assume.

Nick Pickles: Yes, I would like to think so.

Dean Russell: I would imagine Jack Dorsey would do the same.

Nick Pickles: I would hope so.

Dean Russell: Solve this dilemma for me, then. There is a young boy called Zach who has epilepsy and cerebral palsy. He is one of the people with epilepsy to whom people on Twitter send flashing images, which could cause them to fall over, have a fit—if that is the right phrasing to use—lose their job, lose their life or definitely come to serious harm. All Twitter needs to do is stop those images playing as soon as the person receives them, or stop them showing at all, and potentially save their lives. Why has Twitter not done that? Why has it taken the Epilepsy Society to try to create a Zach's law to stop flashing images being played on Twitter to cause harm when you could just stop that right now? You would intervene if it was in real life in front of you. Why is Twitter not intervening now to stop those images being shown or displayed?

Nick Pickles: There are two or three different parts to that. One is how the image is being produced. I know for a fact that previously we had an issue with gifs that come through a third-party service being used in that way. In that case, we worked with the third-party provider to ensure that they did not have images in their library.

If people post them direct to Twitter, that is a violation of our abuse rules. There is also a need for a deterrent; I think this was called out in Young Epilepsy's evidence to the committee. To your point, in the offline space, an assault like that would be punishable by criminal law. As Young Epilepsy called for in its evidence, this is a good example of where there is a need for strengthening the criminal law because there is currently no criminal offence that would cover that issue, which is my understanding, as it suggested in its evidence. I absolutely see Twitter having a role there.

I hope that, through working with our third-party partners and our own technology and rules, we are protecting people, but there also has to be accountability in criminal law. That comes through a number of different areas. If we are asked to police illegal content but the law does not criminalise the activity, that is challenging. The Law Commission works generally in this space, making sure that the law is fit for purpose. We fully support it being a criminal offence. If there is more we can do, we will certainly work with those partners to understand that.

Dean Russell: Can you commit, before we try to do any legislation around this—perhaps even before Christmas—to working with the Epilepsy Society to stop those flashing images being displayed on your site and putting at risk the lives of young kids like Zach?

Nick Pickles: Absolutely. I will also make sure it is a priority for our product teams to investigate whether there are any technical ways we can automate detection of those images. This is a good example where, if other companies have also developed those methods, there is an opportunity.

It is not really covered in the Bill, but one of the things we are quite concerned about is safety technology becoming the purview of only the largest companies. If other companies have developed technology to identify those kinds of triggering images, there is a question of whether the Bill can facilitate the sharing of that technology because, even if Twitter builds it, I am sure there are a number of smaller companies that might not be able to build it themselves as well.

Dean Russell: Thank you. I will take that as a very positive commitment that you will get it sorted before Christmas on your side, and we will work on the legislative side for the criminal activity. Hopefully, no one will have to be sentenced for it because it will have been stopped by your platform. I appreciate that confirmation; it is very welcome. If I may move to you now, Dr Bertram, would you allow a child you knew to eat a Tide pod?

Dr Theo Bertram: No.

Dean Russell: Why is it then that TikTok has had challenges to encourage children primarily—teenagers in particular—to do challenges that have detergent? There is one in the US at the moment, I believe, to slap their teacher. These challenges are very prolific on these platforms and definitely do harm, not just to the individual but to others. Why is TikTok not stopping those?

Dr Theo Bertram: Let me answer that. If I may, I will explain what we do on photosensitivity.

Dean Russell: Yes, please.

Dr Theo Bertram: When a creator creates a video, we have detection to see whether it is likely to trigger photosensitivity. If it is, we say to the creator, "Are you sure you want to post this, or do you want to repost it without that?" If they post it with photosensitivity, when it appears in your feed, the user will get a message saying, "This video contains photosensitivity. Do you want to skip?" Users usually skip that. Therefore, a creator will tend not to post a video with photosensitivity in order to avoid being skipped. We designed it in that way.

Dean Russell: Could I commit you to speak to Twitter and work out how that technology works?

Dr Theo Bertram: Different platforms work in very different ways. I do not necessarily think that one design solution for one works for another. That is how we do it.

Dean Russell: Okay.

Dr Theo Bertram: The Tide pod challenge predates TikTok, but I think your general—

Dean Russell: I believe TikTok was the main place where it happened, was it not?

Dr Theo Bertram: Not the Tide pod one. I think that one predates us.

Dean Russell: Okay, but there was a similar challenge.

Dr Theo Bertram: Your general point about dangerous challenges is something that we take very seriously. We have seen trends and challenges on our platform. Obviously, this area is against our policies; I know you have heard the tech companies say that a lot. We have removed that content, and I can share with you the removal rates for that specific type of content.

We also know that we need to do more. I know that you have heard that a lot as well. We are doing some work at the moment with academics and external third parties to inform us on what we can do proactively to try to deter this and how we best understand children's attitudes to risk—some of the driving factors are children wanting to take risks—so that it helps us design the product in a way that keeps that off the platform. I acknowledge your concerns. The case of Tide pod is not one that is specific to TikTok.

Dean Russell: Okay, but the principle is that there are many dangerous challenges on TikTok of which you are aware. I could list them if you like.

Dr Theo Bertram: One of the things that I would flag on this is the concept of hoaxes as well as challenges. Sometimes, the virality of the outrageousness of a challenge significantly exceeds actual views of that challenge. For example, in the middle of Covid last year, there was a post on Twitter of a challenge allegedly on TikTok, which was the "Lick the seat" challenge—

Dean Russell: Yes.

Dr Theo Bertram: It got lots of attention in the media that there were people daring each other to lick a toilet seat to prove that Covid was not a risk, which seems insane. It is obviously against our rules. We found one video, and the total number of views in the UK was zero. None the less, that challenge was then widely reported and—

Dean Russell: I appreciate that there will be examples of that, but there are many other examples, to be fair, where they have gone viral and have not been stopped. I was interested to know what your process is for that.

Dr Theo Bertram: I agree. Certainly, we have content moderation, and it is against our policies to act in that dangerous way.

Dean Russell: One of the concerns—we have heard this all afternoon—is that every platform we speak to says, "We've got these brilliant policies. Everything's safe. Oh, gosh, it's fantastic. We've got the best policies in the world", but they are not being implemented, or at least they are not being caught. It should not take legislation—it should not take an Online Safety Bill—to stop platforms doing it. I am trying to understand why at the moment you are not catching those "Slap a teacher" challenges—I

know it is in the US and it is different from the UK—which are sickening.

Dr Theo Bertram: Whenever it arises, we take it down. It is hard to anticipate what the future challenges will be. Our trust and safety team is looking at how we see what is starting to trend. The way our content moderation works is the stronger the virality, the more scrutiny there is.

Dean Russell: It comes back to comments we heard earlier about the algorithms. We have heard from witnesses about platforms tending to focus on hateful content because that works better for the algorithms because more people are going to look at it. Does that ring true to you?

Dr Theo Bertram: That came from the Frances Haugen testimony, and I thought all of her deposition was really interesting. In some ways, it has been helpful in moving us to, "What are you going to do about this?", rather than tech companies just being shouted at for failing and for the lack of trust that there is. Instead, I think she was really focused on what we do.

From my understanding of it—I do not want to misrepresent her—she was talking about the way Facebook found that, no matter where it drew the line on its community guidelines, it was the content that bumped up against it but did not breach the lines that would often get the most virality. I know Facebook has taken steps to address that. One of the drivers behind that is that it is a social graph rather than a content graph. What I mean by that is that the thing that amplifies content into your view of Facebook is what other people are doing and how that amplifies to you, whereas TikTok is a more personal, singular experience. That may mean there are the challenges, but the way it works is that what you see in your feed is based on what you are interacting with, not what other people are pushing in. The virality of groups, which is what she was describing, is a different challenge from the one we have.

Dean Russell: I would like to move on to a connected point, which is around anxiety, especially in teenage girls. What we have been hearing is that—this is good promotion for you—TikTok is Instagram on steroids. That is one of the phrases we heard. What research have you done, as TikTok, into anxiety in teenage girls? It seems to be a massive issue for the platform. What work have you done on addictive content, and how do your algorithms feed into that? What actual research has been done to make sure that children and young people using your platform are being protected?

Dr Theo Bertram: I would not describe us at all as Instagram on steroids. We are quite different in how users interact with our platform. The most successful creators on our platform are successful because they are authentic, not because they are perfect. If you asked Simon Cowell to pick who would be the most successful music act in the UK on TikTok, I do not think he would have ever picked—

The Chair: Sorry to interrupt, but I am rather conscious of time. If we could answer the questions that are asked rather than giving extended

background, that would be helpful.

Dr Theo Bertram: I was just trying to illustrate that what we serve is authenticity.

Dean Russell: What research have you done on anxiety for teenage girls, and what research have you done into the potential psychological harms of using TikTok? If you have not done any, it is fine—just say.

Dr Theo Bertram: We have been working with Beat and other organisations associated with eating disorders to make sure that we have the right policies in place. For example, when you search for particular hashtags like pro-ana, you see a message to call the Beat helpline immediately. We are working with them to understand it. Is there a smoking gun document for us like the one that Frances Haugen had? I have not seen that. I do not think we have that research.

Dean Russell: Okay, thank you.

The Chair: Has that research been done, though? Is that the sort of research you do—analysing user experience, particularly among more vulnerable age groups?

Dr Theo Bertram: We have just done a huge piece of work with focus groups, led by Internet Matters, where we brought parents and teens together. It led that and then told us what the analysis from that showed. It was across a whole range of issues to do with safety.

Q245 **The Chair:** You quite rightly said that TikTok is a content graph business rather than a social one. What people see in their "For you" feed are data profiling recommendations based on content—not people you follow but content that matches your data profile.

We have heard evidence in this inquiry and received written evidence, particularly relating to teenage girls, of very real concerns around images of self-harm, including girls cutting themselves in TikToks and that being promoted through "For you" to vulnerable young girls. The challenge that Frances Haugen threw down in that context to Facebook and Instagram—it applies equally to TikTok—is that, for people in that position, if that content is not being picked up, the more vulnerable you are, the more you will see content that may make you self-harm. What is TikTok doing to understand that problem and address it?

Dr Theo Bertram: First of all, our community guidelines do not allow the promotion of—

The Chair: I appreciate you have policies in place, and you take it very seriously and it is a big priority for you. But it is happening anyway, so what are you doing about it beyond that?

Dr Theo Bertram: We work with outside organisations.

The Chair: They can tell you it is happening. It is happening. You know it is happening. Inside the organisation, what are you doing to prevent

those recommendations?

Dr Theo Bertram: The way the algorithm works is that we are always trying to diversify what we send to a user. We are trying to look at how we avoid creating filter bubbles. That is in our commercial interest as well as for the safety of our users. We need to try to make sure that we do not get them into those pockets. It is unusual on the platform that you get that kind of filter bubble. Even in those instances, we are trying to break that filter bubble. We are trying to analyse where we see—

The Chair: But it breaks a little by serving more diverse content. The issue is stopping it happening when self-harm as a category is being recommended to teenage girls through "For you". Your policy does not allow it and they are not searching for it, but it is there anyway. How do you filter out those TikToks? That is what it boils down to.

Dr Theo Bertram: One thing is users searching and finding content.

The Chair: But we are not talking about searching.

Dr Theo Bertram: Another thing is the "For you" feed. The more virality the video gets, the more it is moderated to try to avoid that risk. The problem area is content with a low number of views that is not yet caught by our content moderation.

The Chair: I feel this is something you should be doing more research on because it is a bigger problem than you are admitting. It is a very big problem for the people who are targeted by it.

Dr Theo Bertram: I hear that. I am happy to commit to it.

The Chair: Do you do separate analysis of the different sides of TikTok, particularly Alt TikTok, and the content that people who engage with Alt TikTok may see? There are slightly different genres of content and a slightly different audience, but from what I have heard, the problems there could be worse.

Dr Theo Bertram: I am not entirely sure what you mean by Alt TikTok. Certainly, there are hashtags on TikTok that you can follow.

The Chair: Are you familiar with the expression Alt TikTok?

Dr Theo Bertram: What you are describing is the idea that there are—

The Chair: Sorry, are you familiar with that expression, Alt TikTok?

Dr Theo Bertram: I am not familiar with Alt TikTok.

The Chair: You are not? Okay. It is quite widely spoken about, particularly among teenage girls. It is presented as a more Gen Z version of TikTok, if that is possible. I have heard it reported that the incidence of harmful content particularly targeted at teenage girls is more prevalent on that side of TikTok.

Dr Theo Bertram: When you say that side of TikTok, that is where I am not quite understanding what you mean, as if there is a separate—

The Chair: I believe that "sides" is the expression used to determine different categories of TikTok.

Dr Theo Bertram: You have your "For you" feed. When people talk about Alt TikTok, what I assume they mean is, "My 'For you' feed is personalised to me, and that is not like my dad's. It is mine and what I see in mine is different. What I see in mine is stuff that I am interested in and stuff that I did not think I was interested in as well". What another user will see will be a different feed. I suspect that that concept of Alt TikTok is, "It's for me", but there is not a parallel Alt TikTok channel, as it were.

The Chair: Okay, we will leave that there.

Q246 **Baroness Kidron:** I just want to pick up on something. You said TikTok does not work on your contacts, et cetera. Would you mind saying what it does work on?

Dr Theo Bertram: I can send you the list. When you first open the app, you get a set of videos with a very high number of views that are very clearly moderated. You might flick past a video. The experience of TikTok is that, because they are shorter videos, you are going to watch more. We can take the first signal as, "Did you watch it through? Did you flick past it?" As you go through, those are the kinds of signals—"Did you like it? Did you comment on it?"

Baroness Kidron: Not where you are, who you know, or anything like that?

Dr Theo Bertram: Language is obviously key, so geo is necessary for that. The football clubs that you might like might be more local. Every independent researcher who has looked at what we are collecting generally feels that we are collecting less data than—

Baroness Kidron: I was more interested from a recommending point of view. You just mentioned diversity to the Chair, so you are popping some other things in. What is the basis of the recommendation? Are you just saying engagement?

Dr Theo Bertram: Yes.

Baroness Kidron: Profile?

Dr Theo Bertram: Not so much. It is the engagement. It is the content. What we then do is say, "You flicked past this video. You spent more time on that video. There is a cluster of videos here that we think you might like", and then we will try some of those and see how you engage with them. That is how you end up with things you know you like, which the algorithm has anticipated. Sometimes, there are things that you did not know you liked that you then find you are interested in.

Baroness Kidron: As the Chair was trying to get at, if you say, "They're liking this, they're liking that", on the other material that you think they might like, do you know what is in that material, or are you doing it on a mirror basis?

Dr Theo Bertram: The algorithm is quite content agnostic in that way.

Baroness Kidron: Absolutely. But could you unknowingly therefore recommend things that are mirroring perhaps unhealthy things that are being liked?

Dr Theo Bertram: You can see whether there is a clustering of behaviours and a filter bubble. We are already working to see if we can identify that and take steps to burst it. The thing that makes people enjoy TikTok is that it does not just show you the things you would choose if you ticked the list of things you want to engage in. It shows you stuff where you think, "Oh, I didn't know that".

Baroness Kidron: I understand that, and I recognise that a lot of young people love TikTok. I am not suggesting they do not. I am just saying that you are looking at that and you are trying to puncture it, but as it stands that is a problem—the mirror problem of someone getting into an unhealthy space and similar material being recommended. That has to be the case from what you just said. I am not saying you are not trying, but that is as it stands.

Dr Theo Bertram: The vast majority of usage will be a really diverse set of content because that is what we are trying to drive, and that is what makes people enjoy the product. Filter bubbles are a challenge for everyone in the industry, us included.

Baroness Kidron: Thank you for that. I was interested in your evidence that you wanted the complaints processes to be more defined and limited in scope. I thought that was quite interesting. Can you say to the committee what you mean by limited, and definitely always more defined? I am very interested to hear that.

Dr Theo Bertram: With the complaints process, can I follow up with you on that one? I am trying to remember where we were on the complaints process. I would prefer it if I can check and come back to you on that.

Q247 **Baroness Kidron:** That is fine. Nick, you said in your evidence that we should consider a wider range of platform design interventions to deliver online safety. Then it segued into user controls. I am hoping you are going to tell me something about safety by design and what we could look for from a regulatory point of view—that is, saying "Have you looked at these things?" This sort of goes back to the long conversation you had about risk assessment earlier. I am interested in user controls, but I do not think that the answer can always be, "Let's give more responsibility to the user".

Nick Pickles: Thank you for bringing that up. Safety by design is a foundation of the Bill. An example you have heard evidence on already is

that, if you are building a space, should it be public or private? That is a fundamental design question. Twitter is overwhelmingly open. We have certainly looked at other issues on the platform that Peers have seen. Keeping things public and not creating closed spaces is one thing. On the issue of echo chambers, for example, it is looking at whether there is a way of providing content recommendation that is based on a topic. We would preselect something like climate change, a human would vet the accounts that go into that topic, and then that topic becomes a recommendation rather than focusing on perhaps an individual person or an individual piece of content.

A good example of safety by design is that we tried something that some people did not like. We said, "Before you retweet this, do you want to read it?", which is a relatively simple intervention, but 40% of the people who got that read the article who perhaps previously would not have done. In our evidence, what we meant was that all those things do not involve content moderation in a strict sense. They are design decisions. They are what you might call nudges. Our hope is that in the regulation you protect the ability of those kinds of interventions as well as content moderation, rather than content moderation being the primary, or only, means of dealing with harm.

Baroness Kidron: Okay. It is very late now, so I am going to go quickly. Many of us on the committee are very interested in that. If you have anything more to offer on that subject, we will be very grateful to hear from you.

Briefly, I was interested in your conversation about risk assessments. Something in the Bill that has not had much of an outing is risk profiles for different companies. If there was a market risk assessment by Ofcom right across the piece, with lots of different issues, but it also said, "For these kinds of companies, we are narrowing it down because you have those kinds of problems", and then took a safety-by-design approach, does that seem to you a reasonable way of looking at it, rather than something of a fear that there will be a very long list, much of which may not be useful to you?

Nick Pickles: I think you can ask questions to inform you about the people using your service, where you are marketing your service and where growth is coming from—all those kinds of things. Mumsnet, in its evidence, made some really good points around how to identify potential risk to free expression on a sector basis and a service basis, and worrying that, if the safety-by-design process does not have proportionality in it, you may over-index on one issue to prevent harm that then unwittingly ends up silencing a whole group of people trying to engage in a conversation. Proportionality, but recognising that all services are different, is a really important part of that process. As ever, everything you have outlined sounds wholly sensible to me.

Q248 **Baroness Kidron:** I would like to ask my last question of you both very briefly. It has been a bit of a theme of the day that everybody has fabulous policies, but they are not landing in the lives of the users. Some

of the Bill is really based on the idea that you are going to uphold your policies and that there will be regulatory action, or at least there could be regulatory action, for failure to uphold your policies. As there has been a lot of discussion around director liability, I am interested to know your response. If you put out your policies, you say what you are going to do and you do not do it, how far should the regulator be able to go to ensure that you do? I will start with you, Nick, because I am looking there, but I am coming your way, Theo.

Nick Pickles: Thank you for raising that. We published a paper a few weeks ago talking about how to protect the open internet globally. On criminal liability for staff, there is a reason why these laws are called hostage laws. They are being used around the world to apply pressure to individual staff, who often are not the decision-makers, to try to force companies to do things that they do not want to do. If there is a question of how the UK leads the world in protecting the open internet and leads the world in designing regulation that is proportionate and effective, we have to think very carefully before introducing criminal liability for staff because there are countries around the world that will do the same thing, and use it to have journalists' content and political content taken down by threatening staff with those sanctions, and then will say, "We are just doing what the UK is doing". We have to think long and hard about this because the consequences are global and significant.

Baroness Kidron: Do you accept that there is a big gap between what is being said and what is being done, and then trying to find some levers—I take your point—that are non-financial to ensure that that gap is narrowed?

Nick Pickles: We are all businesses, and financial levers are very significant. Something like GDPR—4% of turnover—has changed behaviour meaningfully, and compliance has been driven, not just in Europe but around the world. We should not underestimate the effectiveness of financial penalties in changing corporate behaviour. As we discussed earlier, that is exactly the kind of thing boards and corporate boards are asking about, and it drives improvements in behaviour. It will drive improvements in the effectiveness of systems that you have highlighted.

Dr Theo Bertram: I agree with Nick, to coin a phrase. There is a big stick already with which to beat people like me and the companies that we represent. If you look at what this Bill will do, it is setting a precedent for content moderation in the way that GDPR did for privacy. What we have seen over the last five to 10 years since GDPR came in is a massive industrial complex around privacy and safety. This is the start of what will be an ongoing trend. We are seeing the DSA. This Bill will probably be a little bit ahead of that. We are seeing it in Ireland with the OSMR. There is a Bill in France as well. We are going to see the spread of this globally.

Across all those places, we are going to see huge industrial growth around how we improve on safety, and I think that will help drive innovation, standards and competition. The Bill is good because it focuses

on systemic failure rather than just on individual failure. Sometimes, we end up focusing on moments of tragedy rather than on systems. What is really important is how we design safety. That is key. It is a competitive part of the industry now. In her statement, Frances Haugen said it was in the profit interest of companies now to do that. I think that is true. There is a competitive case now for us to be safe as we can be.

Baroness Kidron: We have not seen a lack of safety damaging profits yet, but hopefully we will. Thank you.

The Chair: We have a final question from Suzanne Webb.

Q249 **Suzanne Webb:** Thank you, Chair. The joy of coming in at the end is that most of the questions have been asked. First, I want to make the point that policies are just pieces of paper, and it is what you physically do about them that will really matter, and really matters now.

Secondly, following on from what Dean Russell mentioned about Zach and Zach's law, why do you not go and meet him as well? I am sure you will be even more persuaded about what needs to happen about those flashing images. Dean and I have had the pleasure of meeting Zach. It will be a fantastic opportunity for him to explain to you exactly why it is so important.

I think that user safety has got lost in the corporate world somewhere. I mentioned shutting the stable door after the horse has bolted. I think that is what has happened. Everyone is playing catch-up and trying to sort out the unwieldy mess that has happened in user safety and the harmful content that is online.

To come to the point of the question, Debbie Abrahams mentioned corporate governance. How seriously has that been taken by your corporate structures? How often are they talking to you about it? Have they been fully briefed on the Bill? How often are you talking to them about it? Has everyone read the Bill? Quickly, I would like to understand from both of you what your favourite parts of the Bill are. What are your key takeaways from the Bill? Can you let us know?

Dr Theo Bertram: Do you want me to go first, Nick? I have talked to our CEO. I have talked to my boss and our general council about the Bill. More broadly, our CEO is insistent that safety is his top priority. It is a business goal across all the teams, and everyone cares about it. It is not just one part of the team trying to push the others. It is a priority across all the teams.

What is my favourite part of the Bill? The principle of trying to get stuff out of the hands of tech companies into the hands of a democratically accountable public body is a good thing. The way this Bill does it allows flexibility and allows it to be done in a proportionate way with systems rather than just focusing on individual moments. It is really ground-breaking.

The bits of the Bill that I am more wary of are the democratic exception and the journalistic exception, and the possibility that those two things

are abused by bad actors. If this Bill works really well, it will help us manage difficult social challenges about how society thinks differently about harm over time. It is not that long ago that the Government of this country thought it was harmful to promote homosexuality in schools under Section 28. There are still a lot of countries that think that. That kind of change takes time. There is no party in this House that would advocate for that now. I hope this Bill can manage that kind of societal change in a healthy way led by the regulator. My worry would be if it became an area that was used and abused as a political flashpoint.

Nick Pickles: On the risk committee point, I and others have briefed our corporate risk committee of the board directly and been part of discussions on content regulation as a field. This is certainly one of the Bills that is on its radar, also noting countries like Australia, and the DSA, which has already been mentioned, as part of that overall landscape.

On favourite parts of the Bill, I am from Yorkshire, so my default reaction when it comes to things like favourites is to be slightly tongue in cheek. When I read through the definitions of illegal content and saw the phrase "Other illegal" as one of the categories, I raised an eyebrow. It has been a few years since I was at law school, but that might have been one that the student Nick would have been raising some moot points on to suggest that it might not be the clearest of drafting.

The thing that stands out to me in the Bill is the risk of policy objectives being in conflict—in the case of Mumsnet, very directly. People's perspectives of what is harmful are different. The Bill could be used to try to silence the perspectives of people who share, in their view, harmful opinion, potentially through mass complaints to regulators, and attempting to push the system in a direction. Secondly, to take the journalistic point, protecting journalism is incredibly important. Journalism is the lifeblood of Twitter. I do not know how this works without defining who is a journalist. There is a big public policy question about whether that is a good way of regulating.

Today, you have been talking about 6 January a lot. We have removed accounts that said that the election was stolen because we believe that contributes to harm, and it certainly contributed to 6 January. A widely read newspaper has given the former President space on its editorial pages to make that very argument today. Reading the Bill, I do not know if the expectation is that we would stop access to that article because it is spreading something that has not been accepted by Congress, is widely regarded as having been debunked and potentially could cause further societal harm. They are the really big questions that we need to get into to make sure that the Bill is understandable and clear for us and for the wider public.

Suzanne Webb: Thank you very much. Chair, my last comment is a kind of rhetorical question, which I have mentioned at most of the evidence sessions. Why do we need this Bill? Why do you not just get on with protecting people like Zach now and doing the right thing: taking that harmful content down and sorting the algorithms out? Why are you

waiting for the Bill? Time is cracking on. It is really a rhetorical question.

The Chair: And has been answered as such.

Dr Theo Bertram: Although we have already done the photosensitivity, so we have cracked on at least in that regard.

Nick Pickles: Some of the questions in the Bill have to be answered by democratically elected legislators like you. For some of these questions about what speech is legal and what speech is unlawful, it is absolutely right that Parliament sets those standards. It is important that Parliament sets those standards and not a regulator, or through secondary direction of a regulator. They are big democratic questions, and we need Parliament to be making those decisions.

Suzanne Webb: They are absolutely 100% democratic questions, and I understand that as a Member of the Parliament more than anything, but there is user safety and protecting people like Zach as just one example. It is about keeping them safe. It is the human cost of all that is going on that concerns me greatly, and the long delay and wait we will have until this Bill is fully legislated and in place.

The Chair: Thank you very much, Theo Bertram and Nick Pickles, for your evidence this afternoon. After nearly five hours of public evidence, I thank the members of the committee and the committee staff for their forbearance this afternoon. Thank you.