



Select Committee on Communications and Digital

Corrected oral evidence: Freedom of expression online

Tuesday 23 February 2021

3 pm

[Watch the meeting](#)

Members present: Lord Gilbert of Panteg (The Chair); Baroness Bull; Baroness Buscombe; Viscount Colville of Culross; Baroness Featherstone; Baroness Grender; Lord Griffiths of Burry Port; Lord McInnes of Kilwinning; Lord Stevenson of Balmacara; Lord Vaizey of Didcot; The Lord Bishop of Worcester.

Evidence Session No. 11

Heard in Public

Questions 97 - 105

Witnesses

I: Dr Lucas Graves, Research Associate, Reuters Institute for the Study of Journalism; Will Moy, Chief Executive Officer, Full Fact.

USE OF THE TRANSCRIPT

This is a corrected transcript of evidence taken in public and webcast on www.parliamentlive.tv.

Examination of witnesses

Dr Lucas Graves and Will Moy.

Q97 **The Chair:** Welcome to Lucas Graves and Will Moy, witnesses for our session today. Dr Lucas Graves is a research associate at the Reuters Institute for the Study of Journalism, and Will Moy is CEO of Full Fact, the fact-checkers. Thank you both very much indeed for joining us and giving up your time to come and give evidence to our inquiry into freedom of expression online. We appreciate it very much. The Committee has a lot of questions to put to you.

Can I ask you to give us an introduction to yourselves, a bit about your respective organisations, a summary of your corporate structure and who funds you, and your initial perspective on the issue of freedom of expression online, particularly from the perspective of the role of fact checking?

Dr Lucas Graves: Thanks very much for the invitation. I should say that I was, until recently, a senior research fellow at the Reuters Institute in the UK, where I spent two lovely years, but have now returned to my usual post as an associate professor at the University of Wisconsin in Madison, Wisconsin, in the US, where I have been since 2012. Before that, in another life, I was a technology journalist and then, for a while, a technology analyst at a research firm.

The bulk of my academic research has been about fact checking, fact-checking organisations and fact-checking practices, first as they emerged in the US in the 2000s and then increasingly overseas, as this has become really a global movement and fact-checkers have popped up around the world with impressive speed.

I would start with a simple but important observation. We need to conceptually separate fact checking as a practice, how it works and how well it does or does not work, from the quite distinct problem of what sorts of policies we might choose to layer on top of the work of fact-checkers in different contexts: how we treat claims that have been identified as false on social media platforms, in parliamentary debates or in any other setting.

In my experience, professional fact-checkers, the people who work at fact-checking organisations such as Full Fact and similar ones around the world, almost never think of themselves as wanting to police speech. That is a burden that they are very reluctant to take on, and nor should they. In the same way that we would not want a Moody's analyst to be overly concerned about the downstream effects that her ratings might have in the market as she is doing her research and assigning them, we do not want fact-checking groups to be thinking all the time about how their rulings will be implemented in different environments.

Fact-checkers are the first to say, in my experience, that their work is often not black and white, that reaching these verdicts often turns on subtle questions of judgment and that it involves nuance. While it is entirely appropriate that the work of independent fact-checkers be used

by others to identify, label and, in some cases, actively inhibit misinformation in certain settings, such as on Facebook, that should always be done according to very simple and clear policies that, ideally, are laid out well in advance and applied consistently. The strictest measures, deleting a post or sanctioning someone for a statement, should be reserved for those really unambiguous cases where the falsehood is clear and where the potential harm is also quite obvious—health-related misinformation or misinformation about elections for instance.

Will Moy: Thank you for this important inquiry. Full Fact, of which I am CEO, is a charity that exists to inform and improve public debate, as our charitable objects put it. We are a team of fact-checkers, technologists and policy specialists. We look at the full range of ways to tackle the causes and consequences of bad information, but we are best known for our fact-checking work, which was seen by tens of millions of users last year. We believe in the importance of citizenship and everybody's contribution in a democracy. We believe in the importance of high standards in public life.

Full Fact is funded by thousands of people who choose to donate to us every month via a range of charitable trusts. We also receive funding from Google for the artificial intelligence work that we do and from Facebook for fact checking specifically of content on the products it owns: Facebook, Instagram and, latterly, WhatsApp.

The internet has been a tremendous force for democratising public debate. It has made information more accessible, created a more level playing field and helped people get together to pursue causes they believe in. Full Fact is possible because of the internet, and democracy is possible, in different ways, because of the internet.

Freedom of expression, online as offline, includes the freedom to be wrong, and fact checking is the free speech response to misinformation. A few years ago, we published a report called *Tackling Misinformation in an Open Society: What to do when the Cure might be Worse than the Disease*. We said then, "There is a moral panic about 'fake news' which is prompting frightening overreactions" by Governments and potentially internet and media companies.

On the other hand, bad information ruins lives. It damages people's health, promotes hate and hurts democracy. The very real threats to health during the pandemic from misinformation have seen an acceleration of efforts to control information online, with political pressure firmly on the side of removing too much and not too little, and with weak, if any, democratic oversight. We are concerned by these precedents and, in the online safety Bill, Parliament has an opportunity to correct them.

When it comes to freedom of expression online more broadly, I find it striking how many of our traditional intuitions and arguments about freedom of expression are disrupted by changes in technology. I will give a few examples. We were used to a world with a clear distinction between individual speech, which is indivisible, and public platforms,

which are divisible. Now we live in a world where internet companies have fine-grained and nuanced control over how public content is, which can be exercised by item, person or topic.

We were used to a world where there was a clear distinction between individual freedom of expression and institutional freedoms of the press. Now we live in a world with a far wider range of people and organisations acting as gatekeepers and as sources of accountability. It is not so clear what rights and responsibilities they all should have. We were used to a world where a question of speech rights on private property, so-called quasi-public spaces, was an esoteric one, with rare cases about protest rights in shopping centres. Now the question about whether we should have speech rights on privately owned property of internet companies is the central question about whether there is any role for government here too.

We were used to a world where fringe behaviour came at a high cost and was rare: drowning out other people's speech in a physical space or shouting "fire" in a crowded theatre. The technological analogues of these are easier, more common and less accountable. But perhaps most important is the effect of scale. Laws used to be written for humans to enforce, with the benefit of robust processes for protecting people's rights through the courts. Expression online is on such a scale that human enforcement is impossible, and technological enforcement is and probably always will be prone to error.

How people think about freedom of expression has always been shaped by the opportunities we have to express ourselves. While debate on this issue is shaped by the US First Amendment above all and, to a lesser extent, the ECHR and human rights approaches, many of the assumptions their drafters make about how expression works in practice are no longer true. We need to do three things to respond to this. One is to insist on the positive responses for the opportunities to inform and improve public debate, not just the proportionate responses to clearly identified harms.

The second is to insist on proper real-time transparency from internet companies, so that we can understand how their choices shape public debate. The third is to recognise that we do not need a law to solve these problems, but a new body of law, incrementally and cautiously developed over time, with the benefit of widespread conversations that are not yet happening in the way they need to happen.

The Chair: Let us get stuck into some of those issues. Will, what is it exactly that you are checking when you check for facts? It seems to me that you are not really fact-checkers, in that you are trying to establish the truth in a story, rather than whether the facts used to make an argument are, in themselves, strictly correct. Is that right?

Will Moy: If I have understood you, it is, in that a lot of what fact-checkers end up doing is reinserting shades of grey into things that have been described as black and white by campaigners or journalists. Campaigners and journalists need to compress things and make their argument in their best way, so we often spend time saying, essentially,

that it is a bit more complicated than that, not just whether something is technically accurate. Ultimately, our job is to help people make up their own minds about important topics, and that is why we and anyone else I would describe as a decent fact-checker always link to the sources we use, so that people can look at those and make up their own minds about them.

The Chair: We may come back to that in some of the further questions.

Q98 **Baroness Featherstone:** That was a really interesting introduction. You touched on some of the issues that I wanted to raise with you. What is it that we should concern ourselves with when we fact check online? What is the objective? Is it that all content should be checked? If it is not all content, for right or wrong, which content should we be looking at? Where should our focus be?

Does it matter whose content it is? Does it matter more if it is a prominent individual, a corporation or the public, or should the focus be only on the content, for example, where actual harm is delivered, as with, I would argue, the antivaxx conspiracy theories? What differentiation do you make between information or content that is posted to deliberately deceive, as opposed to just misinterpreting; or exaggerating, as opposed to being objectively untrue?

Dr Lucas Graves: You raise a very urgent set of concerns that are all the more urgent right now, in part because of the success that this new field has had. This has brought a lot of interest and additional resources, but also new demands on fact-checkers' time and attention, and rising scrutiny from critics around the world. There is one set of concerns about keeping up with demand and prioritising their efforts. Historically, a lot of fact-checkers, in the United States in particular but also elsewhere, where they were attached to news organisations, had tended to rely on their journalistic news sense to identify the most important claims to check and which claims their audience will be most interested in. This is, after all, a form of journalism, in some cases, that tries to build an audience and to engage readers, so that has to be taken into account.

At the same time, especially since 2016, there has been rising concern over the potentially harmful effects of misinformation. The infodemic illustrates that perfectly, so news sense is not enough any more. Fact-checkers have been struggling with ways to prioritise their work and to identify the most harmful and widespread misinformation. A lot of them rely, for instance, on CrowdTangle to figure out which rumours have spread most widely on Facebook. As Will says, greater transparency from Facebook and other platforms would make it easier for fact-checkers to make those judgments and to figure out where they can best direct their efforts.

Relatedly, there is a difficulty in balancing between traditional political fact checking, which many of these organisations, including Full Fact, originally specialised in—checking claims from politicians and claims that are circulating in the established press—and debunking wild online rumours, antivaxx conspiracies and so on. There is evidence that,

because of the funding that is available, some fact-checkers have been pulled, perhaps more than they would wish, to focus on debunking misinformation and online rumours, largely because of the partnership with Facebook on that.

That is important work, but fact-checkers have to think about how much they want to pull away from their traditional efforts to keep political leaders and elites accountable, especially when we know what an important role political elites and people with large megaphones play in amplifying dangerous online rumours. It is hard to draw a clear line between those two.

There is, relatedly, a concern with maintaining diverse funding sources in this climate. It may be that, in some parts of the world, fact-checkers are overly reliant on a single funding source, which is their partnerships with Facebook. That money has been really important in invigorating fact checking around the world, but I hate to think what would happen to the amount of activity around the world if it were to disappear. That is an ongoing concern.

Will Moy: I agree with everything that Lucas has said. As I said, fact checking exists to help people make up their own minds about topics that are important to them, and Full Fact, as Lucas rightly said, began as a democratic service, primarily about people's role as citizens.

We see ourselves doing three roles with our fact checking at the moment. The first is informing the public. Millions of people choose to come to Full Fact to find out whether something is true. Over the last year, that has been overwhelmingly about the pandemic. It has overwhelmingly been about health, which we would not have predicted, like so many things, a few years ago.

The second, which is a response to the new environment online, is that we are first responders to the emergence of misinformation. Because we fact check a very wide range of things, we see emerging patterns of misinformation often before others do, and we can identify emerging harms. For example, a year before the Government responded to misinformation about 5G, we had publicly warned that this misinformation was spreading and growing.

The third aspect is about power and accountability. Some people do have responsibility to the rest of us to get their facts straight, to give evidence for what they say and to correct the record. That includes politicians in a democratic society; that includes the media. Fact checking is partly about helping people do their job as citizens by understanding government and the society we are in.

All of that means that we cover a very wide range of topics. There is an exploratory element. There is a news judgment-driven element. The keystone for us is thinking about harm: what harm would be done if this thing were wrong?

Q99 **Baroness Buscombe:** What you do is critical, but, of course, trust in what you do is uber-critical. What I am going to ask, you have touched

on, but let us push it a bit further. You talk about the push and the pull, the unambiguous falsehoods and so on, but are fact-checkers neutral in what they decide to fact check? Are the methodologies and processes they use, which are really important, and the judgments they reach, consistent? After all, to establish the truth, you have to ask the right questions, so how do you know what the right questions are? The truth, taken out of context, is not necessarily the truth. Will, you are shaking your head, I hope in agreement with me.

Will Moy: I am nodding in agreement. The best way to make up your own mind about Full Fact's impartiality and independence is to look at our work and judge it for yourselves. In a sense, Full Fact, over the last 10 years, has been an experiment and a stress test of whether it is possible to create an organisation that is trusted as genuinely trying to be impartial in the febrile environment of British politics over the last 10 years: general elections, referendums and, frankly, hostility throughout public debate.

I am pleased that we have a cross-party board of trustees and find our work being used on all sides of political debate, both for people correcting themselves and for people correcting their opponents. We have a lot of institutional safeguards that protect our attempt to be impartial and independent. We have clarity about our funding, which is available on our website. We have a conflicts of interest policy. We have restrictions on staff political activity. Everything we fact check is fact checked at least twice and usually seen by three people before it is published. We link to all our sources.

Impartiality and independence are the product of a process, not something that an individual can manage. They are destinations and things to aspire to, not things that you can perfectly do at all times. We have a corrections page on our website that lists all the corrections we make. When we make mistakes, we correct the record.

Part of that institutionalisation of how fact checking can be reliable, which Lucas can say more about, is the creation of the International Fact-Checking Network. We helped create the network and its code of principles, which establish that credible fact-checkers are committed to non-partisanship and fairness, transparency of sources, funding and methodology, and an open and honest corrections policy. That is backed up by inspections every year or two by an independent third party appointed by the IFCN itself, and those assessments are published on the IFCN website. You can go there and read the assessment of Full Fact or any other fact-checker that is a signatory to the code of practice.

All of that is what is in our control, but let me throw the question back to you in this context. If our fact checks are mainly seen, as they are, via internet companies and other mediators, what is our guarantee that those mediators are not changing the effects of the fact checking that we are doing? We do not know. We do not know for sure whether they select a subset of fact checks and apply them. We do not know whether they apply fact checks differently. We do not know what other factors lead them to amplify or downplay different bits of content.

If you want to answer that question satisfactorily, not only do you need to scrutinise us, but we need real-time information on suspected misinformation from the internet companies, not, as the Government are proposing in the online safety Bill, in annual reports. We need Ofcom to have the powers that the Financial Conduct Authority has to demand information at such times, in such form and verified in such manner as it may direct.

We need independent scrutiny of the use of artificial intelligence by those companies and its unintended consequences: not just what they think it is doing but what it is actually doing. We need real-time information on the content moderation actions those companies take and their effects, because these internet companies can silently and secretly, as in trade secrets, shape public debate. These transparency requirements, therefore, need to be set in the online safety Bill, and only then will we truly be able to answer your question.

Baroness Buscombe: That is very helpful and quite illuminating for all of us. Thank you very much.

Dr Lucas Graves: I would eagerly associate myself with everything that Will just said. By and large, in my research I have found that fact-checkers are neutral or at least, because, as Will said, neutrality is a process and a destination, these are organisations that make a good-faith effort to be neutral and to work in a fair-minded way. That has been embodied in the development and now the very widespread endorsement of the code of principles from the International Fact-Checking Network, which Will mentioned. This is required, for instance, to be a partner in Facebook's fact-checking effort and commits fact-checking organisations to the principles that Will mentioned of non-partisanship, transparency about funding and methods, and having a corrections policy.

At the same time, there are real differences in the quality of the fact checking that these organisations can engage in. I have found that the full-time established organisations tend to do the best work, if only because, over the course of their experience, they develop more rigorous methodologies that have been field tested. Other news organisations engage in ad hoc or occasional fact checking, perhaps during election season. Maybe they do not have dedicated staff involved in fact checking. While they may be excellent news outlets, they have not worked through a methodology in the level of detail that a well-established organisation like Full Fact has. That can lead, in some instances, to lower-quality fact checking.

There are also legitimate differences in the approach that organisations can take, depending on the part of the world they work in, and the political climate and context they are embedded in. For instance, there are rules that seem second nature and quite obvious in the US or the UK around being completely transparent about the sources you have used to reach your conclusion, but there are some fact-checkers working under political threats who, in some cases, have to protect their sources, even academic or authoritative sources that they are using to disprove a claim.

Some fact-checkers are attached to partisan media outlets or political operations. That has been a challenge for the International Fact-Checking Network and for the code of principles, in deciding how exactly to establish whether an organisation is being fair-minded in checking claims from across the political spectrum. How much do you take into account the funding or orientation of the parent organisation when you are judging just the fact-checking outlet?

There are difficult questions that remain, but, in general, these are, or strive to be, fair-minded organisations. They are increasingly developing institutional mechanisms to put that into practice more consistently.

The Chair: Lucas, can I press you on the role of news media organisations in providing fact checking? It is quite an important point. Our last inquiry was on the future of journalism and it became quite clear to us that many agenda-driven news organisations that have an agenda and a perspective, perfectly legitimately so, none the less employ the highest standards of journalism. The fact that they have a partisan viewpoint does not, in itself, matter, as long as they maintain high journalistic standards. If they have a viewpoint and an agenda, and are not politically impartial in the way that Will describes his people, are they well placed to be fact checking other news organisations or is it something that would be best left to impartial, independent organisations of the nature that Will describes?

Dr Lucas Graves: That is a really difficult question and a very good one. As you say, sometimes we miss the point that news outlets that are dedicated to a cause or have a particular history of allegiance to a party can nevertheless do excellent investigative work and excellent journalism, with regard to the accuracy of the information they are presenting to the public, based on their internal fact checking. You can build a political argument more or less convincingly on the basis of sound research and reporting. We have many examples of excellent news organisations that do take a point of view.

When it comes to fact-checking outlets and external fact checking, in other words establishing the veracity of claims that are already circulating in the public sphere, it matters whether an organisation is willing to assess claims from across the political spectrum. There is an outfit in the United States called Media Matters that does excellent research and is dedicated entirely to establishing how much lying takes place on Fox News. It does excellent work. At the same time, and for good reason, it is not seen as being an independent voice, because it never takes up the question of misrepresentations that come from the left or that circulate in left-leaning news outlets.

While it is fine for partisan news organisations to check one another, that is not as useful in contributing to a solid bed of facts on which we can try to build policy and carry out a public conversation.

The Chair: Thank you. That is interesting.

Q100 **Lord Griffiths of Burry Port:** Fact against fake news seems such an easy binary pair to sew up in a couple of minutes, but you are already

proving that it is anything but that. Facts may help people to make their mind up, but, in an argument, say, between the State of Israel and the Palestinian state, you will have two sides putting what they think are facts, which they have chosen to make an argument, and the arguments will be moving in different directions.

If I may give a different example, when the then Archbishop of Canterbury was facing the Lambeth Conference, which was going to be riven down the middle by two different views about sexual identity, he went away for four months and came back with a 320-page biography of Dostoevsky, written to persuade him that ambiguity was the best outcome in many of the arguments that we put forward anyway.

All this grey area that has been alluded to is where my interest lies. If something does harm, let us call it. You cannot drink disinfectant to get rid of Covid-19. Poetry is my background, and poetry is about metaphor, imagery, suggestion, hinting and all that stuff. How do you get a group of people to make judgments, which are going to be subjective, and then make policy out of those conclusions, given the welter of human material produced by our interactivity, which defies the simple question, "Is this a fact"?

Who looks at the work of the fact-finders? Quis custodiet ipsos custodes? Who controls those who keep the controls? Is there an appeals process when judgments have been made? If so, how effective is it? I looked at Byzantine history, and this all feels a bit byzantine to me, so please clear my mind for me.

Will Moy: Could I take a stab at the poetry and let Lucas talk about the appeals processes, and see where we get to?

The Chair: Give it a go.

Will Moy: Politics is what happens when two people look at the same set of facts and reach different conclusions because they have different priorities, principles and appetites for risk. It is characteristic of the world that looking at the same set of facts does not lead you to the same conclusions. Any fact-checker who tried to get everybody to reach the same set of conclusions about what to do with the facts would be way overreaching their job. That is not our role. The more we can clear out factual conclusion and let people debate the shared world we all live in, the better.

Often, we cannot do that, because many important facts are unknown or uncertain, and the job of fact-checkers becomes expressing the uncertainty or talking about things that are not known, despite the fact that some people may confidently claim that we can be sure of them. There is a need to distinguish the modest role of fact checking, which is to establish which facts can be established and to be clear about which cannot, from the role of politics, which is debating what to do about them, including the uncertainty, principles and value judgments that go into politics.

As long as you keep fact checking in its box, recognising that you can build different things on the foundations of a shared reality, it has a

useful role to play. An easy way to slip up in fact checking is to try to check a claim about the future—as we always say, you cannot fact check the future—or a value judgment. You can sometimes inform people’s choices about those judgments but you can never tell them what the answer is, because it is not a question of fact. I will let Lucas talk about the appeals and I can answer any specific questions about how Full Fact handles that in more detail afterwards.

Dr Lucas Graves: I would agree with Will that some of the most important work that well-established fact-checking organisations do is in deciding what cannot be checked, and which claims or texts are beyond the purview of their fact-checking work. They take this very seriously, in the best cases, and have well-articulated policies about trying to distinguish between checkable claims and statements of opinion. At the same time, that also can be a hard line to draw. There are disagreements among fact-checking outlets about what are and are not appropriate things to check.

I will give you one example. A lot of American fact-checkers, in the heat of the 2012 race in the United States, agreed that the Republican claim that Obamacare, President Obama’s healthcare reform proposal, constituted a government takeover of healthcare was provably false. It was a ludicrous claim. Many of them called it their lie of the year. There was a counterargument that, while one person’s government takeover of healthcare may not be another person’s government takeover of healthcare, it is inherently a matter of opinion. The fact-checkers responded by saying, “We have models for what a government takeover of healthcare looks like. We have reasonable benchmarks. Private insurance remains in place and there are no government-paid doctors. It is nothing like the Canadian system or the NHS. All that evidence suggests that you cannot reasonably describe this as a government takeover of healthcare”.

However, people could disagree. As Will says, mostly what fact-checkers do is add nuance to questions. They bring to light the ambiguity. There is a tendency to focus on their ratings, how many Pinocchios they hand out or whether they say that something is officially false or true. Some do not use ratings, such as Full Fact. Even for organisations that do stamp labels on things, that always comes after a lengthy article, sometimes thousands of words about a relatively simple question, which brings to light the nuance and the ambiguity.

I realise I have not spoken about the formal appeals processes. They need work in general. Most fact-checkers are quite transparent in publishing corrections. They will often even publish critical mail that they have had that did not lead them to make a correction. They will explain why they did not make the correction, but they will present the argument that a complainer made. That needs to be institutionalised more clearly. Not all fact-checkers do that, so that process could be formalised a bit more.

Will Moy: Specifically at Full Fact, there was an interesting case of self-restraint in fact checking. Lots of people would have had us fact check

the economic predictions around the EU referendum, and we did not. You cannot fact check the future. We did publish an article explaining how economic models work and the kinds of questions you might want to ask about them, so that choice of self-restraint in fact checking is a crucial part.

We have a complaints process on our website and it has multiple layers, if you are the subject of a piece. There is a hugely important balance, as an independent organisation, between having an effective complaints process and having a process that can be used as a means of lobbying us. I know that other media organisations have a challenge in maintaining independence and handling complaints, which can be very time-consuming, in an effective way.

It becomes really interesting where fact checks are used by internet companies. The only company we know of that provides public criteria for what fact-checkers it works with, and an appeal process on fact checks, is Facebook. The appeal process just says, "Email the fact-checker". It is so opaque that, by and large, people find the wrong fact-checker to email about the thing they want to complain about, so we often get emails about other people's fact checks to our email address, which is confusing.

Where fact checks are used privately by companies, not for them to show publicly but for their human moderation processes or to train their machine learning models that purport to identify false information automatically, there is no accountability at all. If a mistake is somewhere in that pile, it will not be visible. That is why it is so important that this Committee recommend a strong information requirement power for Ofcom.

Q101 Viscount Colville of Culross: You have just been talking about the increasing power of machine learning in fact checking. Many are concerned that AI is too blunt a tool for fact checking. YouTube said that, when it reduced the number of human moderators, the number of videos removed doubled and there was a huge increase in appeals being upheld. Is there a problem in that AI often does not get the cultural context for a piece of information and that, as a result, it poses a threat to freedom of information? Will, in your introduction, you said that technological enforcement is prone to error. If you could address that question, I would be grateful.

Will Moy: There is that problem. There are many other problems with technological enforcement. Automated processes do not provide an effective means of fact checking at internet scale. We would advise that any claims for how technology can do this reliably should be treated as unreliable, unless they are independently verified. That includes claims from the major internet companies.

Although internet companies have very fine-grained control over how content spreads, the technology that tries to support identifying false or misleading information is both limited and error-prone. They are like a driver who has great speed control but poor hazard awareness. In a

sense, that is not surprising, because there is no single source of truth for them to turn to. There is no source of facts that you can look everything up against. One of the questions we need to talk about, if we want people to really enjoy freedom of expression and the freedom to impart and receive information, is the paucity of high-quality information provided in our public life, partly by our public authorities. If we do not correct that, the rest of the debate falls.

We have worked on artificial intelligence in this field for many years. We have automated part of our process using AI, but always with a view to helping a member of staff have better and faster access to helpful information, not to replacing them. We know that internet companies employ people to manually review algorithmic decisions. We do not know anything about how the review was triggered, what actions they can take, how often it happens or anything else. Again, there is no transparency.

I can take you through two roles that algorithms play here. One is what we call claim matching. If you have identified a claim that is false, where else does it appear that you can then take action against? This is seriously limited in its accuracy. A headline example from the last few weeks is Facebook banning posts containing references to the place Plymouth Hoe, mistaking the word "hoe" for something pejorative. More seriously, we had a case of a police force having a post warning about Covid scams labelled by an internet company, because the technology so crudely believed that "Covid" and "scam" being in the same post—at least, as much as we can guess—was enough to label a post by a police force warning people against being defrauded.

This claim-matching technology is very limited and we should recognise that it can be wrong in two ways. Either it can overreach and hit too much, or it can underreach. Right now, all the political incentives are for overreach, because of the health threat. It is essential that we correct course on that, and the online safety Bill is an opportunity to get proper parliamentary oversight on this.

You have to understand the political process behind this. Last year, the Secretary of State for Digital, Culture, Media and Sport issued a press release saying he had summoned bosses of internet companies to his office, taken them to task over 5G misinformation and ordered them, essentially, to take down as much of it as possible. I do not know whether to be grateful that a press release was done about that, so we know that this was happening in the open, or concerned that a Minister believes it is their job to order internet companies to try to remove certain content from the internet, with no parliamentary process for overseeing that in a democratic way. That is a dangerous oversight that is the responsibility of Parliament to correct when the online safety Bill comes before you.

The second role of algorithms is what we jokingly, in house, call robo-checking: automatically trying to determine whether something is true with a computer. This is possible only when you have authoritative reference data for the computer to respond to. For example, we have a

piece of software that can automatically look at a claim such as “the economy is growing”, go off to the Office for National Statistics, get the statistics, work out whether that is true, look at the context and so on.

You can think just how quickly that kind of software falls apart if you say, “the economy is growing in Wales”, “the economy is growing in Northamptonshire”, or “the economy is growing in my experience”. This kind of technology is very sensitive to small changes, so even where good reference is available, which it often is not, the technology is limited in what it can do. Just to emphasise, I have talked only about a simple case of text in English. So much important misinformation spreads in images, videos and other languages.

Most importantly, though, algorithms do crude things. We are used to very nuanced legislation to protect freedom of expression when we inhibit speech. We are used to multiple criteria: whether people have a reasonable excuse, what was in their mind, whether they could foresee the effects and all that kind of thing. Algorithms cannot answer any of those questions.

For that reason, the idea that what is illegal offline should be illegal online is attractive in principle but fails in practice, because it is impossible to enforce technologically. Current laws are impossible to enforce at internet scale. I do not foresee that changing, so you have a hard choice in front of you. Would you prefer to write cruder laws or to have error-prone enforcement of laws, which the law as written never envisaged? Our technological experts would be happy to provide more evidence about any of that.

Viscount Colville of Culross: That would be very useful. It is a bit of a grim choice for us.

Dr Lucas Graves: I will add one point to Will’s excellent summary. Artificial intelligence can enhance the reach and effectiveness of human fact-checking work in a couple of ways. It is useful in helping to identify and prioritise claims to check. That has not been a problem lately, but it can be a difficult part of fact-checkers’ work, needing to turn up interesting claims from the daily political grist. Algorithms can help there.

As Will said, AI can be useful in spotting when a claim that has already been checked has been repeated somewhere. That is important, in order to get the correction out quickly into the new outlet or context where it has been repeated. As Full Fact can attest, it is also important because, in the long term, it helps organisations to figure out whether they are effective, where they might direct their efforts and where claims that they have checked are continuing to circulate.

Ironically, one of the least transparent environments in this respect is Facebook and other social media platforms. The area where fact-checkers get less feedback about the effects of their work is, perversely, on the social media platforms, where that data is abundantly available but just not presented to them. Of course, the holy grail of completely automated fact checking is conceivable only in cases where there is a narrow statistical question being checked, something that can be looked

up quickly. Even then, it often fails, so you would always want to have a human in the loop.

Viscount Colville of Culross: Thank you very much indeed. That was fascinating.

Q102 **Baroness Greender:** Thank you so much for the compelling evidence so far. It has been brilliant. You have talked a lot about nuance in the way you do your work, but we have a perfect case study that is the opposite of nuance, which, of course, is Donald Trump. It would be great to hear from you about what the platforms should do when a fact is incontrovertibly proved to be wholly wrong.

Dr Lucas Graves: That is perhaps the most difficult question. Laying out the range of possible interventions is useful. At one end, you have just adding context to a claim, by making other information, such as relevant articles, available to people who see the false claim that is circulating. One degree up, you have labelling it as false, giving people a warning message that says, "There is reason to believe that this claim is not true. Here is the work of independent fact-checkers showing that it is not true".

Then there is more actively inhibiting the spread, either by a pop-up warning presented to people when they are about to share it on their feeds or across their network, or, under the surface, algorithmically, by simply making a given post less likely to spread or to be shared between users. We know very little about how that works. That is a sort of black-box process. Finally, the most extreme intervention is removing the content entirely or perhaps banning that account permanently. We have seen some examples of that recently.

We need to make these decisions collectively and as democratically as possible, but, in general, platforms are right to remove misinformation that poses a clear potential harm, for instance election-related misinformation on the eve of a vote. A post informing people that the vote will take place the following day, or that they need to go somewhere other than their usual polling station, is clearly harmful and qualifies for removal, as do bleach cures, false remedies or antivaxx information.

Those instances are relatively rare, and I agree with Will's earlier point that part of the value of defending free speech is that people have the right to make mistakes and to be wrong in public. Ideally, that is an opportunity for counterarguments to be presented, so we should apply those harsh measures very reluctantly. More importantly than anything, the platforms need to be consistent in applying labels and warning messages. A politician should be as subject to having their posts stamped as false or misleading as any other member of a social media network. Platforms should be much more transparent about which rules lead to which actions, in which parts of the world.

Lastly, I would love to see a central database of all the claims that have been fact checked. This would not be easy to spread or to share in the way that most posts are, but it would be a spot on Facebook, for

instance, where you could see a record of all the claims that have been checked in a given country or globally, gathered in one place. Then we could understand how well this process is working, and see what kinds of misinformation have been circulating and what has been inhibited. They have been very reluctant to be that transparent, but it is essential.

Baroness Greender: Will, what you said about your funding model at the beginning was really interesting. Anyone who has worked in a large charity that receives government funding knows that they have to get the right balance with individual funding, so that they can continue to be critical of government. I am just wondering where the balance fits for you. First, if you could answer the original question, I would really appreciate that.

Will Moy: Most internet companies take action on every piece of content on their system. They decide how many people they are seen by, how they are displayed and so on. Those choices collectively are probably more important than specific content moderation decisions on content that is known to be false. Those choices are treated as commercial secrets, but they can powerfully enhance our ability to impart or receive information, or they can infringe on our freedom of expression. That is why we need strong information powers in the online safety Bill, so we can start to understand not just the content but the effects of those decisions. Parliament also needs to consider what constraints they should operate under in making those choices.

When it comes to action on specific pieces of content, we should start with freedom of expression, which includes the freedom to be wrong. We do not believe that an internet company or anybody else should necessarily take action just because somebody says something that is not true. If you look at speech offences in UK law, rarely, if ever, is a false statement on its own sufficient to trigger a response or penalty. Offences take into account the topic, motive, intent and audience, whether somebody did know or ought to have known what they were saying was false, and how somebody else might understand their statement.

When action is taken on specific content, the starting point should be giving users information that helps them make up their own minds about whether to trust what they are seeing. Baroness O'Neill calls that "assessability". Fact checking is one way of doing that. Features that help show the original source of content and advertising transparency requirements are two other examples. Lucas has given you a summary of a range of things that internet companies can do. There is a larger one in the recent UNESCO *Balancing Act* report. The ones with most impact on freedom of expression are downranking, taking down content and sanctions on users.

Offline, those kinds of restrictions on content tend to be a response to behaviour: either behaviour leading to a real risk of harm, as Lucas was saying, such as falsely calling "fire", or someone breaching their responsibilities, such as advertising truthfulness or accuracy rules for broadcasters. A problem of behaviour cannot be solved by restrictions on

content. Restrictions on content that affect both sharers and readers will very often be a disproportionate response when we have the right to impart and receive information.

Let me give you two examples. The first is President Trump. His political statements were very frequently detached from reality. That is fundamentally a problem of political behaviour, which led, in turn, to a problem of content moderation, how the media should handle it and lots of other things. You cannot solve that upstream problem by trying to deal with a content moderation problem down the line. The same challenge is here, as was mentioned earlier. If people are not willing to challenge those on their own side of the political debate over accuracy, we will not have a culture of accuracy in political debate, and what we do online does not matter one whit.

The second example I want to give you is about 5G. As I said earlier, we identified, a year before it became a major issue, a spreading concern about 5G safety. It was completely predictable. The same happened about 4G, 3G and things before that. We can tell you the whole history. What we said at the time was that there is no good public health information. Public Health England should fill that void by putting out reliable, trustworthy, public information about the health effects of 5G. That is the free speech response to that problem. That did not happen and, a year later, we had telecoms technicians being harassed doing their jobs and buildings being attacked. We had a serious real-world problem.

This is the question: had we intervened sooner with better information instead of coercive content moderation, would we be in a better place than ending up with Ministers asking internet companies into their offices? In doing that, where does responsibility lie? Is it the public health organisations? Is it the telecoms industry that spent millions of pounds rolling this out but failed to explain it adequately to the public? Is it the media? Is it the internet companies? Is it a bit of everyone? Turning it into just a content moderation problem is not the right approach in either of those situations.

To your point on funding, I have recommended that this Committee disregard claims made by the internet companies about the accuracy of their AI, unless they are subject to independent verification. You will have to make up your own mind as to whether we speak independently of them, and you will have a wider range of evidence to rely on, but I have no problem criticising the internet companies when they deserve it.

The Chair: As you indeed have. Thank you.

Q103 Lord Griffiths of Burry Port: How successful are fact-checkers at reaching their intended audiences? Can they find that they have reinforced prejudice or got people to run away from the whole exercise? Do users value fact-checking services?

Dr Lucas Graves: On the one hand, fact-checking sites have been able to pull together impressively and surprisingly large audiences in many cases. Will mentioned the tens of millions of visitors that a site like Full

Fact can get over the course of a year. That is impressive for a policy-oriented, highly technical operation such as that. The same has been true in the United States with sites such as factcheck.org and PolitiFact. Fact-checking operations attached to broadcast newsrooms, such as the BBC's Reality Check, which goes out across various networks and programmes, or Channel 4's fact-checking operation, can also reach millions or perhaps tens of millions of people, in some cases.

However, we know almost nothing about how often a given fact check reaches the relevant audience, in the sense of reaching someone who happens to hold the particular misperception that the fact check is addressing. There is every reason to think that that does not happen very often, especially when it comes to fact checks distributed online. What little that we have suggests that it happens quite rarely.

Will Moy: Full Fact's fact checks, like those of some other fact-checkers, are directly embedded in Google Search, Facebook, Instagram and so on. Roughly speaking, we have 10 times the number of people seeing our fact checks in Google Search pages than on our website, so there is a huge influence from the choices internet companies make on who fact checks reach. We do not know a lot about that. As Lucas says, there is a real evidence gap there.

Q104 **Baroness Bull:** I wanted to seek your views on digital and media literacy and their role in this relationship between information and the perception and use of information. Media literacy has emerged over the last few years as everyone's favourite solution to the issue of understanding what is online, but people are not very clear about what they mean by it or, indeed, what policies should be put in place in order to improve media literacy. In our own report, *Breaking News?*, we found worrying disconnects between levels of literacy, particular in older audiences. Do people have the skills necessary to determine what is true online? How could we improve media literacy particularly among older adults?

Dr Lucas Graves: Media literacy is vital. To give just one illustration of that, there is recent research suggesting that higher news literacy in particular, understanding how news organisations work, the conventions of layout in a newspaper, and the difference between news and opinion, correlates to greater scepticism about online media content and claims circulating on social media. At the same time, there are no easy fixes. Media literacy is not a panacea. It is unfair to ask schools to educate away patterns of belief that result from social or economic divides in many cases. It is not a burden that they can carry.

I would also echo the media scholar danah boyd and point to the paradox in the things that we demand of media literacy. On the one hand, we want people to be sceptical. We want them to question authority. In many cases, we have encouraged media literacy in the past because we wanted people to challenge the received wisdom of the *New York Times*, the BBC and so on. At the same time, in the midst of the moral panic that Will pointed to, we want people to believe and to trust. Too often, that simply means that we want them to have faith in the

same institutions that we have faith in. Again, that may be important but it is not something that one hour a week in ninth form is going to address.

Will Moy: I agree with everything Lucas has said. I will point to a real evidence gap here about what are effective media literacy techniques and, as Lucas alluded to, the potential for unintended and sometimes detrimental consequences from well-intentioned media literacy interventions. You might think about three different types of intervention. One is a response to specific information gaps or misinformation, like we suggested on 5G and public health. An information response to that might have been better than enforcement.

The second is media literacy interventions at the point of decision. Prompts integrated into internet companies' products that suggest fact checking or "read this before you share it" have been shown to make a very significant difference to user behaviour, because you do not have to remember much. It is right in front of you.

The third kind, which is what people mostly think about, is educational interventions to change general attitudes and skills, which are hard, expensive and require two-way commitment from those you are educating, as well as those who are trying to do it. Sitting down an entire adult population and teaching them how to use the internet is probably not a feasible goal. We are looking forward to the Government's media literacy strategy and seeing how they put these ingredients in the pot.

Q105 **The Chair:** It has been very interesting. There is one area we have not touched on, which is the role of fact checking in the current public health crisis. You have spoken about your role in dealing with conspiracy theories, such as antivaxx and 5G, and that seems reasonably clear cut, but there are some areas of the public health emergency where there is genuine dispute among experts. Platforms are inclined to and want to take down content that contradicts public health advice and the view of the World Health Organization or Public Health England.

Are you coming under some pressure to go beyond your expertise and determine which expert is right? We probably will not know until this all over whether masks really are that important and whether they work, and whether lockdowns are the most effective way of dealing with the public health crisis. The evidence at the moment is that, on balance, they probably are, and it is certainly the view that is being taken by the public health authorities, although they have changed their mind, yet there are some very eminent experts who disagree. Are you in a position to determine which of those eminent experts is right? Is this such an emergency that we should just follow the bodies that are there to make these decisions on our behalf?

Will Moy: It is not Full Fact's job to communicate on behalf of those bodies or anyone else. If they want to get their message across, that is their problem. We are not under that kind of pressure and, if we were, it would be water off a duck's back. That would not be really of any interest to us. The challenge is how we do justice to exactly this problem. Our job

is to help people make up their own minds. When you are trying to make your own mind up about evidence that requires years of training to even begin to appreciate, you are left in a really difficult position.

In practice, we see two tiers of discussion about this. On the one hand, we are dealing with the most amazingly flawed and dangerous misinformation, which has no plausible basis and is based on clear misunderstandings that have clear health consequences. A lot of what we do in our online fact checking sits in that space. It may purport to be about the safety of masks; it may, in fact, have citations on the bottom, but it just does not have a credible attachment to an attempt to understand the truth and is not hard to fact check.

On the other hand, you have these contested areas of expertise. There is a limited use of the old legal test of a responsible body of professional opinion: one person with a PhD saying something does not necessarily make it a thing, but is there a responsible body of people with serious credibility who are having this discussion? At that point, our role as fact-checkers is to do our best to give people the information and help them make up their minds.

For example, just as I told you about the questions that you could ask about economic models, we can do that kind of thing. We do not have to just say, "Here is a bunch of papers. Read them and weep". We can say, "Here is a summary and here are some of the key differences", but you eventually hit the limits of what fact checking can do. The limits are our ability to all make up our own minds about highly technical subjects. Sometimes, you just have to choose who you believe.

The Chair: Indeed, I guess that is the key point to end on. Thank you very much indeed, Dr Graves and Will Moy, for a very interesting and useful session. We may well be in touch with you for a little more information. You have offered to send us more in one or two areas, which we would appreciate. We have a few more witnesses to come, but, for the moment, that brings this session to an end.