

ELECTORAL MATTERS COMMITTEE

Inquiry into the Impact of Social Media on Elections and Electoral Administration

Melbourne—Thursday, 19 November 2020

(via videoconference)

MEMBERS

Mr Lee Tarlamis—Chair

Mrs Bev McArthur—Deputy Chair

Ms Lizzie Blandthorn

Mr Matthew Guy

Ms Katie Hall

Ms Wendy Lovell

Mr Andy Meddick

Mr Cesar Melhem

Mr Tim Quilty

Dr Tim Read

WITNESSES

Professor Stephan Lewandowsky, School of Psychological Science, University of Bristol, and School of Psychological Science, University of Western Australia, and

Professor Ullrich Ecker, School of Psychological Science, University of Western Australia.

The CHAIR: I declare open this public hearing for the Electoral Matters Committee Inquiry into the Impact of Social Media on Elections and Electoral Administration. I would like to begin this hearing by respectfully acknowledging the Aboriginal peoples, the traditional custodians of the various lands that each of us is gathered on today, and pay my respects to their ancestors, elders and families. I particularly welcome any elders or community members who are here to impart their knowledge of this issue to the committee or who are watching this broadcast of these proceedings.

I welcome Professor Stephan Lewandowsky, University of Bristol and University of Western Australia, and Professor Ullrich Ecker, University of Western Australia. I am Lee Tarlamis, the Chair of the committee and a Member for South Eastern Metropolitan Region. The other members of the committee here today are Bev McArthur, Deputy Chair and a Member for Western Victoria; the Honourable Wendy Lovell, a Member for Northern Victoria; Andy Meddick, a Member for Western Victoria; and Dr Tim Read, Member for Brunswick.

All evidence taken by this committee is protected by parliamentary privilege. Therefore you are protected against any action in Australia for what you say here today. However, if you repeat the same things outside this hearing, including on social media, those comments may not be protected by this privilege. While you are covered in Australia under parliamentary privilege for any comments you make today, you should note that Australian law cannot give you the same protection with respect to the publication of your evidence in the jurisdiction you are giving evidence from. All evidence given today is being recorded by Hansard. You will be provided with a proof version of the transcript for you to check as soon as it is available. Verified transcripts, PowerPoints, presentations and handouts will be placed on the committee's website as soon as possible.

I now invite you to make an opening statement, which will be followed by questions from the committee.

Prof. LEWANDOWSKY: Thank you. I think Ulli and I have decided to take turns, just 2 minutes each, so maybe I can start out. For me the most important thing to recognise is the context in which information is disseminated online. One of the things a lot of people are saying is that at the moment the internet is not regulated. However, I do not think that is true. It is actually highly regulated but it is regulated or governed by algorithms that are designed by platforms to enhance user engagement in order to maximise their profits. Those algorithms determine at least in part the information diet of a lot of people online, so I think one of the things we have to look at is the context in which misinformation unfolds—things like algorithmically driven polarisation, which we know exists. We know that Facebook knew that their algorithms were polarising the American public, but they chose not to do anything about it because they were worried about their revenue stream. On top of that we have microtargeting—the ability for advertisers and political messages to be aimed at voters based on their vulnerabilities, based on their personalities. It turns out we can infer a person's personality from their Facebook likes. If we have enough likes, we can do a better job through the machine than that person's spouse; Facebook knows your personality better than your spouse does.

So those are the problems we have to deal with structurally. And unfolding against that is the psychology of misinformation, which means it is very difficult to correct. I think at this point I will hand over to Ulli since he is the world expert in that.

Prof. ECKER: Sure. What I would like to point out is that exposure to misinformation, to conspiracy theories can actually have lasting impacts on a person's thinking. So if you are exposed to disinformation, it can affect your decision-making even after you have acknowledged that you have received a clear and credible correction, which basically means that little fixes like adding a little red flag to something published online is not really going to cut it. If you are exposed to a conspiracy theory, there is evidence that that can result in a loss of trust in institutions even if that institution is not actually implicated in the particular conspiracy theory that you are exposed to. I think loss of trust is something that is a very real threat to democracy, as I think we are witnessing globally.

The second point I would like to make is that the whole idea that there is some shared knowledge, that there is agreement on a basic set of facts, sort of provides the foundation of democracy and the common ground for any evidence-based policymaking. So what we see with this rise of disinformation and conspiracy theories is that this common ground is sort of eroded, and I think we can call that an epistemological crisis that should be of concern to all policymakers.

The last point I would like to make is that the risks are being downplayed. The platforms will say they are doing everything they can, which of course is not true. Many users will experience the environment as relatively benign—I mean, my feed does not lure me into conspiracy theory rabbit holes. But even if the majority of users are not being sucked into these filter bubbles and these echo chambers, social media algorithms are designed to increase the amount of fringe content that many people see as well as the extremeness of that content. And as Steve said, they target vulnerable people who might feel disenfranchised and who might be prone to conspiratorial thinking or psychopathy or whatever, which leads to a demonstrable radicalisation of those individuals. There are a lot of case studies now showing how exposure to those conspiracy theories has led individuals to become radicalised, and a few of them have been charged with terror-related offences in the US and so on. I am sure you are all aware of that. I will leave it at that for now.

The CHAIR: Thank you. I might go to Dr Read for the first question. He may not be there. I will go to Mr Meddick.

Mr MEDDICK: Thank you, Chair, and thank you both. I am just looking at one of the pieces in your presentation here about people continuing to believe false things even if they are proven false and how even after correction sometimes on social media our brains are not good at processing that. What do you think can be done to change that scenario? Is there anything that can in fact be done, or are we now so programmed within a social media platform as a society that it does not matter if the truth is out there, we are just not going to believe it?

Prof. ECKER: Well, it is a nuanced issue. When we say that people continue to be influenced in their reasoning by misinformation that has been corrected, it does not necessarily mean that fact-checking does not work at all. I think we have very good evidence that fact-checking does work to an extent. It can reduce the misconceptions that you hold, but often it does not generalise beyond that and it might not affect your behaviours. It might wear off over time, so you might now realise, ‘Oh, yes, that’s false’, but if I ask you again in a week’s time, you might revert back to your false belief. So it does not mean that fact-checking does not work. There is a lot of benefit of fact-checking, but the effects are sort of smallish and they do not last, and we need to do more than that.

Mr MEDDICK: In traditional media, for instance, where this occurs, newspapers, for instance, and sometimes even television programs like news programs, for instance, have to print or announce retractions. Do you think that there is a possibility—I am clearly no expert here—that some form of, say, Facebook actually puts out a post, given that they have access to everyone’s accounts, that they can actually say, ‘This is actually false and we let this go. Here’s a retraction, and this is the actual truth’. That way everybody knows. Is there a role for that?

Prof. LEWANDOWSKY: Just to pick up on that specific suggestion, Twitter actually did that after the last presidential election. They contacted all their users who were following Russian bot accounts and were exposed to disinformation. So that tells us that the technological capability is there for the platforms to contact people and tell them, ‘Oh, by the way, you were exposed to misinformation’. Now, another point I want to make, connecting to what Ulli just said, is what we do know is very effective is to try and warn people ahead of time that they might be misinformed and to tell them ahead of time how they can resist that misinformation. That technique is called inoculation, and it is just like medical inoculation. You are exposing people to a weakened dose of misinformation and you explain to them how they might be misled, and we can then show that people become more resistant to misinformation.

We also know from COVID-related misinformation that Facebook is very capable of identifying things that are clearly false. In the case of COVID, they warned users for certain content that the content was false. Now, according to Facebook’s own data, that cut sharing of that false content by about 95 per cent, so it was very effective because it was presented upfront rather than afterwards as a correction. And it was not censorship, which of course is something we all want to avoid, because people were still able to click through those

warnings to get at the dodgy content. It was just that it was much harder to get there. The moment you introduce a little bit of friction like that on social media you can dampen the flow of information.

And by the way, before I forget, the reason I am sitting in front of this background is because that is a report I was lead author on for the European Commission that we just finished. It is a public document available at that link, and that goes into very much detail on all the issues we are talking about here, but it was not ready at the time I submitted. I was not allowed to circulate it back then, but now it is public. You can have a look if you want.

Mr MEDDICK: If I can suggest, Chair, then, perhaps if Professor Lewandowsky would be able to forward that, we can take that in as a paper.

Prof. LEWANDOWSKY: Sure.

The CHAIR: Yes, that would be good, absolutely. Thank you, both. I just want to follow up on that. You mentioned inoculation. Are you familiar with that being used anywhere and how effective it has been, the concept of inoculation and sending out that information ahead of a campaign being run?

Prof. ECKER: Yes. We have lab data on that. It does work. Whether it has been applied in the field large scale, I am not sure. Stephan, I am not sure if you know much more about that.

Prof. LEWANDOWSKY: Yes, well, it depends on what you mean by that. Sander van der Linden and colleagues have developed something they call the *Bad News* game, which is available online, and that has been played by—I forget now, an incredible number of people—at least hundreds of thousands, if not millions. You can do it yourself; it is the *Bad News* game. What it does is it is a 15-minute game where the person who is playing it is trying to get credit by designing misleading messages, so the game is about learning how disinformers work by effectively turning you into a person who is in a fictional world trying to mislead. You can show that once people have played that game they are better at identifying misleading content, so there is some evidence that that works. We also have a series of experiments currently underway, about halfway done, where we show that very brief videos of less than 2 minutes can teach people how to detect certain misleading techniques; for example, highly emotive content that is supposed to elicit outrage, which is one of the big problems on social media—there is all this pumping out of information that elicits outrage or hatred or other negative emotions. So we can show fairly good effects in experiments with very brief videos that could be rolled out at scale. So I think there is a lot of potential there, but of course to do all that requires an environment where the social media companies will do that.

The CHAIR: Yes. Ms Lovell.

Ms LOVELL: I am going to pass at the moment, Mr Tarlamis; my questions have actually been asked.

The CHAIR: No worries. Dr Read.

Dr READ: Thank you very much for your appearance. A lot of your work seems to apply to the public health field, and you have used COVID examples, but it seems there are a lot of parallels with the anti-vaccine battles that have been going on. Have you looked at anything in that arena that might apply also to political misinformation?

Prof. ECKER: We have actually looked at political misinformation, so maybe I can say a little bit about that in a minute, but yes, we have used materials that relate to vaccines, and we find pretty much the same effect. I do not think there is anything specific about the topic of vaccines that makes it much different from other sort of conspiracy theory content or misinformation. We have used vaccine-related misinformation and find pretty much the same effects: that inoculations tend to work, fact-checking works, but then fact-checking needs to be designed properly and well to have an influence, and the effects are not as large as we would like them to be, so little flags are not going to fix it.

In terms of political misinformation, again, we find the same kind of patterns—corrections do work, you can reduce people's political misconceptions, but what is much harder to do is change the actual support for a candidate or their voting intentions. In the US in particular it seems very difficult. Even if you point out that a politician is lying over and over and over again, you can reduce misconceptions—the beliefs in those lies—but

it is very hard to shift the support. We ran similar studies in Australia where that is not quite the case. Voters in Australia seem to be much more worried about whether politicians lie to them and they will penalise the politician for lying and reduce their support for them. That is a difference that we are still exploring but that seems to be the case, that it is a cultural difference in some regard.

Dr READ: And just quickly, are you aware of any examples of inoculation-type of messaging being conducted by, say, a respected independent authority or even by a government in an effort to counter fake information?

Prof. LEWANDOWSKY: Yes, there is in fact one in the vaccination context; at least one case. I would have to look up the details—it was Idaho or Iowa, one of those two—some time ago where they were facing I think an outbreak of measles because of low vaccinations, and so what the health department did there was to pre-emptively describe some of the misleading arguments that were bound to arise and they made that available to the media at the outbreak of the pandemic. It turns out that the media was therefore able to report on the misleading information, being informed about how it was misleading. It is a bit of sophisticated nuance, but they took on this whole notion that most of the people who got measles were in fact vaccinated, which you always get because of the base rates involved because most people are vaccinated. It is a nuanced, technical issue that lends itself to misinformers to exploit something that is just a statistical consequence that is unavoidable even if vaccinations are effective. And despite being complicated, they were able to pre-empt that to find much traction. So there is evidence for that.

Dr READ: Great, thank you.

The CHAIR: Mrs McArthur.

Mrs McARTHUR: Thank you, Chair. Can you just explain a bit more about the concept of creating friction by intervention, such as bumps and nudges, around social media sharing?

Prof. LEWANDOWSKY: Okay. You mean the friction idea? The fact that we can—yes. Well, let me give you one example from real life. Sometime ago—a couple of years, two or three years ago—there was a string of horrific murders in India where mobs would descend on completely innocent people who were travelling through a village and would kill them. The reason for that, it turned out, was fake videos of child kidnappings that were shared on WhatsApp. They were shared together with completely made-up information saying, ‘Oh, these guys in that white car over there, they’re the ones behind this kidnapping’, and bang people would congregate, and on occasion it would lead to terrible violence. Something like 60 people got killed in that manner repeatedly—I mean, there were quite a few incidents. Now, what WhatsApp did as a consequence of that was two things. Number one, they limited the number of chats that you could forward a message to. It used to be unlimited; they cut that down to five. The second thing they did was to identify information that was being forwarded as having been forwarded. Prior to that, if you got a message from your friend that they had forwarded, you could not tell that they had forwarded it. It looked as though it came from your friend. Now, those are tiny changes, and of course it is very difficult to establish causality, but since those changes were made, which has now been two or three years, there has not been a recurrence of these mob killings, because it just became technologically far more difficult to create this blizzard, this viral spread of misinformation. So the intervention was tiny and it was not censorship, but it was sufficient, it appears, to cut out the mechanism by which these mobs formed. Now, again, causality is always difficult to establish—I cannot guarantee that is what caused the cessation of those matters—but I do think it is a striking example of how clever little changes can scale up in society to make a problem worse or, in this case, to make it better.

Prof. ECKER: Another example recently is Twitter. If you retweet something—so you share something that you have received with others—and there is a link in that to an article or whatever, when you now try to retweet that, if you have not opened that article, it asks you, ‘Hey, do you want to read it first?’. A lot of people then do not share or they actually open the article and read it or at least look at it before they share it. You can still share it without looking at it, but it is just that little bit of friction to go, ‘Oh, maybe I should look at it before I share it’. They only put that out I think temporarily during the US election, so it might go away soon again, but the initial evidence I have seen suggests that a lot of people open the article and less people share potentially misleading information. The example Steve mentioned earlier where Facebook at least—I think in the US, I am not sure but somewhere—tried this where they actually hid the misleading content and you had to click a little button to view the content. So it is not censorship. Again, you can still look at it, but a lot of

people just do not. If it is labelled as misleading, it can be hidden from you. If you know that it is probably misleading, you are most likely not going to actually look at it.

The CHAIR: Do any other committee members have any questions?

Ms LOVELL: Yes. I just want to follow up on that one you were just talking about where now they are putting on a thing that this may be misleading. How do you feel about the social media platforms becoming the censors—you know, self-appointed censors—of whether something is misleading or truthful?

Prof. LEWANDOWSKY: Well, thank you. I mean, that is the elephant in the room, isn't it? Who is making those decisions? You are absolutely right; it is a serious issue. And as I said in my opening remarks, at the moment everything is done by the social media platforms, and to my knowledge they are not accountable to any democratic bodies. They are accountable to their shareholders at most, and that is a serious problem. That is precisely the space where we have to take some regulatory action or we have got to have a conversation about who decides that. Now, in the case of COVID and vaccinations, it is relatively simple because, you know, there are the CDC and the WHO and all these acronyms of professional bodies who know medical science. In politics of course it becomes far messier. So that is precisely where we have to have the conversation and, yes, that is where you come in.

Prof. ECKER: Even in politics, though, there are claims being made in political advertising that are just outright false. They are just clearly false. There is no debate about them being false. There are lies being peddled in political advertising that are clearly false, so it should not be too difficult to find an independent arbiter who can tell you whether it is true or false. There is another development in that there is something called the wisdom of the crowd where you get the opinion of many, many, many people. You can also approach truth. I cannot recall exactly where I saw that, but taking user trustworthiness ratings into account could help those algorithms actually up rank trusted content and down rank not trusted content, which then relies not on an individual determining whether something is true or false but a majority. Now, that comes with its own problems. I am not saying that is going to be the solution, but Steve is right. I think it is just something where there needs to be a conversation had.

There are a couple of other things that can be done that I might just mention. People are very sensitive to social norms, so if you, as the electoral committee, now wanted to do something apart from considering regulation, which I think in this area is necessary but still—apart from regulation, applying or fostering social norms, so creating campaigns to highlight the fact that most people do not like liars and fearmongers and spreaders of hatred and that everyone has a responsibility to act responsibly on social media to protect our democracy, that is one thing that could be done. Or to encourage voters to not listen to social media so much and to make up their own mind and make it as easy as possible for them to go to maybe one source that will tell them what all the parties and all the candidates that they are voting on actually stand for to get them more engaged with the policies that you guys put forward, that they are voting on, rather than just relying on the stuff they see on social media.

I saw just recently a 'how shall I vote' quiz that I took, which I thought was a clever way. That might be particularly suitable for younger voters. The website is called ISideWith. It is based in the US but they have an Australian spin-off, and as far as I can understand it is independent. They ask you a bunch of questions and you can decide who to vote for. Getting voters engaged with those kinds of things rather than just absorbing what they are fed on social media I think could be positive steps forward.

The CHAIR: All right. Thank you very much for your time today, Professor Lewandowsky and Professor Ecker, and for your submission also. I apologise that we were running a little bit late as well. Thank you very much for your time and your insights.

Prof. LEWANDOWSKY: Okay. Thank you for having us.

The CHAIR: All right. That ends this session. Thank you.

Witnesses withdrew.