HOUSE OF COMMONS

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

Oral evidence: Online safety and online harms, HC 620

Tuesday 14 December 2021

Ordered by the House of Commons to be published on 14 December 2021.

Watch the meeting

Members present: Julian Knight (Chair); Kevin Brennan; Steve Brine; Clive Efford; Julie Elliott; Damian Green; Simon Jupp; Jane Stevenson.

Questions 227 - 260

Witnesses

I: Julie Inman Grant, Australian e-Safety Commissioner.

Examination of witness

Witness: Julie Inman Grant.

[*This evidence was taken by video conference*]

Q227 **Chair:** This is the Digital, Culture, Media and Sport Select Committee. This is our Sub-Committee on Online Harms and Disinformation. We are joined today from Australia by Julie Inman Grant, the Australian e-Safety Commissioner. Before I open the session, I am going to make some declarations of interest. If anyone else wants to put any interests out in the public domain they are free to do so. The interest that I wish to declare is the fact that I am the Chair of the APPG for New and Advanced Technologies. Anyone else? That is fine.

Julie, thank you very much for joining us. Good morning from the UK and good evening in Australia.

Julie Inman Grant: Good evening to you and good morning.

Chair: The Online Safety Act allows you to request takedowns of certain types of harmful content. Why do you think Australia went down this route compared to the UK's focus on systems and processes rather than individual pieces of content?

Julie Inman Grant: We actually do a bit of both. Would it be helpful if I give you a quick overview on e-Safety, who we are, what we do and our history?

We very much evolved out of the kind of tragedy that you saw with Caroline Flack, a well-known presenter with very public mental health issues. She was terribly trolled on Twitter, had a nervous breakdown, came back on, was trolled on Twitter and ended up tragically taking her own life. That started a petition to the Government in 2014 to get involved and do something about online abuse and bullying.

What the Government decided to do in 2015, rather than create an adult cyber-abuse scheme at the time, they wanted to start something gingerly. So, they took other functions, including our online content scheme, which deals with illegal content, including child sexual abuse material and terrorist content. That has been in place for about 21 years. We serve as the hotline for Australia, just like IWF, the NGO, does for the UK.

What they created was a youth-based cyber-bullying scheme. We are meant to serve as a safety net for children who are seriously harassed, intimidated, humiliated or threatened. If they report to the social media site and that isn't taken down—because we know so many things fall through the cracks. If you think about content-moderation services, they are outsourced to developing countries. They are dealing with millions of

pieces of content and they have between 30 seconds and a minute to determine whether that piece of content violates their terms of service. Culture and context often fall through the cracks. We serve as that safety net and we advocate on behalf of the child, in this instance, and we bridge that inherent power gap that exists between that big tech behemoth and the young person or their parent.

That has been tremendously successful, so we layered on different schemes over time and in 2017 we became the e-Safety Commissioner. That started an image-based abuse scheme, where we have about an 85% success rate in terms of taking down intimate images and videos shared without consent on thousands of sites around the world. Almost none of the content is hosted in Australia.

Following the Christchurch atrocities, we were given new powers around abhorrent, violent material notifications. We also have the ability to call crisis events and to compel ISP blocking to prevent the virality of terrorist or criminal activities that are captured by a perpetrator or an accomplice for the purposes of inciting terrorism.

We just renewed our Act, and Parliament passed a new reform Bill in July 2021. That will come into play on 23 January. It does a few things. We know that harms don't just happen on social media, and that was the focus when this was developed. So we will cover everything from online gaming site platforms, dating sites—it recognises that search engines and app stores have a role in the ecosystem—up to servers at bottlenecks, for instance.

We will have a mandatory set of codes that will cover eight different subsectors of the high-tech sector, ranging from what we call websites and these platforms to IoT devices, also in terms of retailers that are selling technology, or point-of-sale retailers. This will require them to proactively detect and remove illegal content, but also restrict access to harmful content, such as pornography to children.

Then we have basic online safety expectations, which are clearly articulated objectives. This is a ministerial instrument that basically lays out what we expect the licence to operate. This is for systems and processes, a type of provision that does not have significant penalties, as is in your Bill, but will require me to compel transparency reports so that we can get under the hood.

We are trying to balance out the individual content takedowns with more systemic changes, because what we have seen over time are flaws in the system. The intelligence and the threat trends that we get from interacting with our citizens is much more effective than user surveys or even over-the-top behavioural analysis. We see how people are weaponising technology to commit online harms, but we are seeing systemic failures.

An example would be that we are seeing evidence that the platforms really aren't stopping the creation of fake impostor accounts. They are not really effectively preventing what we call recidivism. You see those same bad actors continuing to return to the platforms, maybe using a different SIM card or a different device ID, but the companies have the ability to pick up those signals and triangulate and prevent the return of those bad actors but there are some systemic failures there.

Another example would be around what we call volumetric cross-platform attacks or brigading and pile-ons. They have the ability to see when a particular account or set of accounts are being attacked and when habitually abusive accounts are attacking. The challenge is that we have had situations where a single person would be receiving 400 tweets, for instance. Often they are organised on Telegram or on other services. It is almost like death by 1,000 cuts. The platforms will say to us, "Tell us each one that contravenes your terms of service." We will say, "No actually, can't you see what's happening here? I can see that in the open. Why can't your stop this at scale or at least protect that account?"

Q228 **Chair:** Julie, thank you. It is very extensive but very interesting to trot around exactly where you are and how you have got there. Something that occurred to me during your answer to the question was that you mentioned an 85% success rate in terms of taking intimate images down globally. How is that you have been able to be that successful and how did you manage to get over the fact that Australia is a country of 30 million people in a global population of 8 billion—a tiny fraction? It is a similar situation for the UK as well, with 70 million people. How do you get over that when social media companies try to have global positions and you are asking them effectively to make exceptions for Australia and to take a very different approach?

Julie Inman Grant: Most of what you would call the well-known or responsible platforms have been quite helpful. For instance, for youth-based cyber-bullying, they do not want children being bullied on their platform, so 90% of our cases have been settled informally where the content has been removed. It has to meet that serious cyber-bullying threshold but we haven't had to compel them, fine them or issue service notifications.

More of the challenges that we have are with what we would call rogue porn sites or sites that are set up for the purposes of humiliating women, usually, in the imaged-based abuse cases. We have collaborative relationships with the platforms. We have escalation paths that we have established over the years into their trust and safety teams and we are constantly reaching out to mid-tier platforms to try to build those relationships to make sure that they understand the laws and that we will apply them. I think we have been able to do that because we have been the only regulator in the world looking after online harms that has been able to do that.

They need a licence to operate in Australia, and particularly where they have invested and put sales and marketing subsidiaries in there, they want to be able to comply with local laws. We often use things as basic as terms of service violations to point out that companies are violating their own terms of service. When you have a Government seal behind it, it does help compel the takedown. And when you have threats of significant fines and other sorts of remedial actions, people do act.

Q229 **Chair:** And criminal sanction?

Julie Inman Grant: No, we are civil powers mostly but we work closely with law enforcement. There are a number of schemes where there are both criminal penalties and civil penalties. What I would say about our civil powers is that when people come to us, mostly what they want is to have the content taken down. They don't want to have to go through a criminal justice pathway and go through the court system. It is expensive, it takes a long time and they often have to share their images and they are re-traumatised by all that.

Q230 **Chair:** More generally, before I hand over to Steve Brine for his questions, what you have said today seems to talk about the speed, reactivity and being able, effectively, to point them in the direction where there is a harm taking place. In the UK, in the last 24 hours, we have had the publication of the Joint Committee report on online safety, which proposes installing an extra layer of bureaucracy between Minister and Parliament, and potential for further delays to any form of legislative action in order to widen the scope of what is defined as an online harm. Do you think that is counterproductive, and that speed should be of the essence?

Julie Inman Grant: The more expeditiously you can get the content taken down, the more relief from trauma and psychological distress. I would say that is absolutely right. We have had situations where there was a Fight Fit video targeting a very vulnerable child and Instagram was able to help us get it down in 12 minutes.

For most of our schemes with the reform legislation we have truncated the timeframes for removal from 48 hours to 24 hours. Most of them do have around-the-sun, as we call it, trust and safety operations. They are often using AI for at least initial detection, followed with human moderation. Therefore, they can move with speed, just as they move with speed when they are coming up with new products or targeting consumers with advertising.

Q231 **Steve Brine:** Hello, Julie; thank you for joining us. Your Online Safety Act, like our regime proposed here, provides a list of exemptions for journalists and public interest, because of course they would never spread misinformation. How do you think you will manage your duties against these exemptions? What should we be learning from you in this space?

Julie Inman Grant: All of us have a different regulatory environment and different bodies handling this. I would say one of the key differences is that we are an independent authority dealing with online safety and online harms. What I guess you would call our traditional media regulator, the Australian Communications and Media Authority, is working on disinformation and misinformation, and to a certain degree scams. Our ACCC, which is the competition regulator, has instigated the news media bargaining code. From my read—and I just had a quick skim of the Bill this morning and we met with Ofcom and I had a very fruitful meeting two weeks ago—it will have a much broader scope and different issues that it will be focused on. That is probably not answering your question but it has never really come up.

The only regulatory area where the issue of journalism and newsworthiness has come up is in the context of abhorrent, violent material. Whether livestreamed terrorism, rape, torture, all of the horribles, we have to make those assessments. Every time we get a report it triggers an investigation where it has to stand up in a court of law. It has to stand up to an internal review, ombudsman review and a tribunal review and in a court of law, so we have to be very diligent about how we exercise that. Whenever there is a question of newsworthiness, we will lean towards that. If it is on the news, it does not make sense to block it online.

Q232 **Steve Brine:** What is your feeling on blanket exemptions? We have received evidence that critiques that. Did you have that in the passage of your legislation?

Julie Inman Grant: Not blanket exemptions. Of course, we have discretion when we are developing our regulatory guidance, always having a little bit of room for flexibility. What we find in these kinds of jobs is that you can write the best legislation for the time of day but you can never anticipate what harms will happen or the creative and myriad ways that people will misuse technology or weaponise it, so you want to give yourselves the best possible tools while protecting a range of fundamental rights.

Q233 **Damian Green:** Hello, Julie; thank you for joining us. One of the big issues that we are facing here is the use of anonymous accounts for trolling purposes. What powers do you have to look into this?

Julie Inman Grant: It is worth stating up front that, in addition to our regulatory powers, we bolster everything we do with prevention on the front end and what I call proactive and systemic change on the other end. Proactive and systemic change not only involves our safety-by-design initiative—which I can talk about later—but also what we call tech trends and challenges, so looking at the future because we know technology is always going to outpace policy. There are some really challenging issues, like anonymity and identity shielding, where there seem to be easy, blanket solutions but it is much more contextual than that. We have done a brief on that.

The vast majority of policy makers want to reduce anonymity to the extent that it provides a veil of protection or anonymity for a person who is abusing another person in some way, or violating the law. We have to be careful about blanket exemptions around anonymity and pseudonymity. It is the whole idea of taking a sledgehammer to a nut. You will have heard this from your very active, sophisticated and vocal advocacy community in the UK: that there are very good reasons for anonymity or pseudonymity, for your domestic violence victim or if you are a whistleblower or that sort of thing.

We have new powers, particularly because we have a new adult cyber-abuse scheme, where we will be able to take down serious cases of adult cyber abuse if we can prove intent to cause serious harm and that it is menacing, harassing and offensive in all cases. It will allow us to go to the platforms to collect BSI, basic subscriber information. Often the worst—I do not want to call them trolls—serial abusers who are using targeted online harassment to silence voices or cause psychological trauma, or threats of violence, will use pseudonymous accounts to hide their identity, sometimes. There are a lot of people who are happy to abuse others with their name and identity pretty clear.

What some of the platforms have started doing—and I think we should require them to do more—is, first, they should stop the creation of fake accounts. They can pick up dark patterns and signals, and when you see that there is one IP address that is creating 20 or 30 accounts a day, that should be telling the platforms that something is not right there. However, when they are suspending accounts for violations, as a condition of reinstating that account they should start collecting more identity information so that they can verify them. Some platforms, like Twitter, will capture a workable email address or a phone address. Often, when we issue a BSI request to a platform, it is because it is a fake account and we need to find out who the real person is, so that we can issue an enforcement notice to that person.

Q234 **Damian Green:** Sorry to interrupt. I am interested in what powers you have. You can issue a BSI request, which will give you the name and identity of whoever is behind this account. What can you then do?

Julie Inman Grant: If they have it, if the platform has it, and that is the rub.

Damian Green: If the platform has it. You cannot force the platforms to collect that.

Julie Inman Grant: If the platform has it.

Q235 **Damian Green:** How do the platforms respond? Do the platforms—

Julie Inman Grant: "Sorry, we do not have any information or we have an IP address" and we will use that as a basis to do a who-is reverse lookup and try to do our best investigations to identify who the person is.

The challenge here is re-identifying, say, Meta's 3.7 billion current users. This has not been a standard practice where it has collected any kind of identity information in the past. That is where the challenge is. We can compel the BSI but I would say that the platforms are quite spotty in terms of the information that they have now. That is something that probably needs to be looked at: in what cases should companies be collecting this information and should it be in response to abuse and how is that done in a privacy-protective way?

Q236 **Damian Green:** One of the arguments here—and we can all see the arguments for anonymity in the sorts of cases that you have suggested—is that at least the platform should know the identity of everyone who is using it. Do you think that would be helpful?

Julie Inman Grant: It would be more helpful. There are companies—like formally Facebook now Meta—that have a real-names policy. This is where our basic online safety expectations and your systems and processes provisions that you are proposing for Ofcom are really important. We are not seeing them effectively enforcing their own policies. I am sure you know lots of people on Meta, on Facebook, who are using pseudonyms now. There are systemic challenges that we can very clearly see with the plain eye.

We need to be thoughtful and cautious about what information we ask them to collect, how they store it and for what reasons. There are platforms, for instance, that if you sign up to the app store, either Google Play or the app store, they have credit card information. That is something that can be used through an investigation, a much easier pathway to identify who that person might be, for the purposes of issuing enforcement action.

I would just say that we are not providing that as a service to citizens. We are going to unmask the trolls so that little Johnny's mother will know who is bullying little Johnny. These powers are solely for the purpose of identifying who might be behind an abusive account so that we can issue an enforcement action.

Q237 **Damian Green:** There is an area of activity that we would all regard as undesirable that you cannot stop, is that right?

Julie Inman Grant: In terms of what kind of undesirable activity?

Damian Green: As you describe here as bullying at school or something like that. Your powers do not extend to that.

Julie Inman Grant: All I am saying is that, in situations like defamation cases, which we do not handle, or bullying or abuse cases, the person being bullied or attacked, if it is a pseudonymous or anonymous account, they might want to know who is behind that account. The powers that the Government have provided us, which will come into play in January, are solely for the purposes of collecting basic subscriber information so that we can fulfil an enforcement action, either a fine or a removal notice, not

as a service, per se, to tell a complainant who might be bullying or harassing them. There are a range of reasons for that but one of them is privacy and confidentiality. It is complex, no question.

Q238 **Simon Jupp:** Good morning. Indeed, it is very complex and you have talked about trolling already this morning. I understand that you have the powers to fine trolls. How does this process work in practice, given the information that you have shared with us about some of the challenges you have with identifying people in the first place? I think you have muted, which happened a lot in 2020 and 2021.

Julie Inman Grant: Yes, I am sorry, my service cut out and I only heard the last few words of your question. Can you please rephrase?

Simon Jupp: It is about the power that you have to fine individual trolls, bearing in mind that you have just told my colleague Damian about the challenge of identifying some people. How does this process work in practice?

Julie Inman Grant: In terms of the youth-based cyber-bullying scheme or the incoming serious adult cyber-abuse scheme, I will tell you how the adult cyber-abuse scheme will work because it is modelled on the youth-based cyber-bullying scheme.

The idea is that if a person is experiencing some form of serious online harassment, they first report it to the platform where the abuse is happening. There are a couple of reasons for that. The first is it is the platform's responsibility to remediate that harm. That is the most expeditious way to do it. We as a small regulatory agency cannot become the policemen or the censors of the internet. We are here as a safety net.

If a person applies, they think they have experienced serious cyber abuse and it does not come down within 24 hours, they can come to us and say, "I believe that this is serious adult cyber abuse." That triggers an investigation. We have about 40 investigators. We would look at the evidence that they provide us and then there is a two-pronged test, an objective test.

We have to decide whether we think that there is a serious intent to cause harm and emotional distress with that content and whether it is menacing, harassing and offensive in all cases. It is a very, very high threshold, on par with a criminal threshold. That was by design because what we are trying to target is that most harmful content—not banter, not defamation or hurt feelings or even name calling—things like serious cyber abuse, threats to kill, cyber stalking, doxing and some of those pointy-end forms of abuse.

Let's say we decide that it constitutes serious cyber abuse. What we would do is we might issue a notification to the platform saying that we think this constitutes serious cyber abuse. If they don't take it down, we

can fine them for each case up to about $500,000, so not huge. We also have powers to fine perpetrators and a range of other remedial actions.

We want to make sure that there is a deterrent effect: first, that content hosts are being responsive in taking it down and, secondly, that people do not think that they can abuse others with total impunity any more. What happens right now is that if you troll someone on Twitter, the worst thing that will happen is that you might be suspended for a few days. Even if you are permanently suspended, it is pretty easy to use VPNs and multiple SIM cards or devices to circumvent those processes. We want to make sure that people know that there is respite for the victims that come to us and, secondly, that there is a deterrent and that people cannot do this with total impunity any more.

Q239 **Simon Jupp:** Thank you for going through that in so much detail. I was not clear whether there was a judicial process in this or whether it was something completely in-house. Could you clarify, please?

Julie Inman Grant: No, this is completely in-house, through our investigations, and there is often a legal review, particularly if it is an edge case. There is recourse. In the six years that we have operated and dealt with thousands and thousands of takedowns and a range of remedial actions, nobody has challenged any decision that we have made, but that is available. We are setting up internal review but there is also the tribunal, there is an ombudsman and there is the federal court.

Q240 **Simon Jupp:** The key question is: has it worked? Has the amount of anonymous trolling on social media reduced as a result of this policy over the last five or six years?

Julie Inman Grant: Not for adult serious cyber abuse, because that is a new scheme that starts on 23 January. That would be what we hope to see. If you apply that to youth-based cyber-bullying, no, it has not stopped youth-based cyber-bullying, just like face-to-face bullying has been around from time immemorial and is still more prevalent than online bullying. But now our children have a place to go to get some relief.

The thing that is insidious about online bullying, of course, is it follows a child home in their pockets and it can be all hours of the night. It is very visible to a young person and their peers but can be almost invisible to parents and teachers. We know that it can cause a great deal of distress and often is compounded by face-to-face bullying. It is an extension of that.

I do not think that we are going to eradicate bullying or people being mean to each other or causing drama, but that is what we use our prevention, education and cultural change efforts to try to do. We are moving on multiple fronts and giving people the tools to know how to use the conversation controls, the muting and the blocking functions, how to respond, how to talk to your kids about what is happening, all those things.

Q241 **Simon Jupp:** Exactly. If I may explore the regulator-operated complaints procedure that you have in place there, my understanding is that your office will run an investigation and complaints scheme for cyber-bullying and image-based abuse. How do these schemes work in practice, handling complaints, undertaking formal investigations?

Julie Inman Grant: I explained how the serious cyber-abuse case would work. It is the same thing with youth-based cyber-bullying. Image-based abuse is slightly different. It covers all Australians and we do not require a person who has experienced image-based abuse to report to a platform or an image board or a website first. They can come directly to us with the URL where their images are.

We have an online forum, esafety.gov.au. We have reporting forums for all of our schemes. It is a pretty clean user interface. One of the first things that we try to assess is the level of distress the victim is feeling so that we can triage appropriately. We try to respond within 24 hours. I want one of the hallmarks of our agency to be responsive and compassionate citizen service and I think we have delivered that.

Most of these cases are very complex, particularly with image-based abuse where there may be what we call relationship retribution. We are taking a lot of reports on sexual extortion schemes. The scheme also covers deepfakes. Essentially what we do is they will provide the URL, and some people come to us with up to 400 URLs. If you have a really determined predator who is trying to humiliate you, they will put it on multiple places. We will investigate, we will figure out where it is, we will issue any kind of notice, but often in terms of a service violation notice. Like I said, we have an 85% success rate.

We can also look at a range of remedial actions that we can take. We have issued a range of formal and informal notices, which are educative to the perpetrator. We can fine the perpetrator and the content host if they do not remove the content. We also have some other remedial actions. We had a case where we knew that somebody who was repeatedly placing images of his former partner online had it stored in his iCloud account and he was about to be deported, so we issued a notice so that he had to comply with deleting all of those photos. That can be verified.

Q242 **Simon Jupp:** Can I ask you, in an average 12 months, about the number of complaints that you receive as part of these processes, or expect to receive?

Julie Inman Grant: We have seen a huge spike in reports over Covid, as you can imagine. For child sexual abuse material we had about 21,000 reports, the largest in our 20-year history. That was a 90% increase over 2019. We have seen the same exponential growth for image-based abuse. We have seen a 114% increase and we have had about 3,500 this past year. With youth-based cyber-bullying, we have helped several

thousand children, a 30% increase, and a 40% increase in the adult cyber-abuse scheme. They are all on the upwards trajectory.

Q243 **Simon Jupp:** Just briefly, are there problems managing the scale of infringing content? How much of a challenge is this for you? This must be quite a Herculean task.

Julie Inman Grant: It is, but it is a service that is vital to our citizens. We would like to be able to help more. We have set up what we call surge support. We have four different investigative teams but they all sit together in a hotline room. For instance, over the Easter long weekend there were four different variants of a sextortion scam that went out and we saw a 600% increase in image-based abuse reports. Then we can draw on our investigators either in cyber abuse or in our cyber-report team, which is child sexual abuse material, to help. We are dedicated to try to respond to people within 24 hours and give them relatively rapid and successful outcomes.

If we do not get the content taken down, we try to provide a range of other supportive services, talking them through how they protect themselves online, referring them on to mental health services, legal services or other helpful information. We want to make sure that people leave feeling like somebody cared and had their back and is doing something to help remediate the harm.

Q244 **Kevin Brennan:** Thank you for joining us, Julie. Good evening. We appreciate you giving evidence. To follow up on something that my colleague was asking on this new adult scheme that you said was coming in in January, do you have a working assumption on what percentage of complaints you are likely to receive that will meet your very high threshold that you described?

Julie Inman Grant: I think that we are going to get high volumes for this because we have never seen such a rising tide of hate and polarisation on the internet as we see today. We have an informal scheme now, as I said, and within three months of instituting that in July 2017 it has far exceeded the number of reports that we have had from youth-based cyber-bullying. It is at a very high threshold. Again, it depends on the nature of the reports that we get in and that is going to be the challenge. I am hoping it will be a good 20% that we will be able to help people on, but it is hard to forecast, given that the scheme is not in place.

Q245 **Kevin Brennan:** On the pilot scheme that you have been running, what is the rate that has met that threshold? Has it been around that 20% figure?

Julie Inman Grant: It has been much lower because we have not had formal powers. We have reserved the asks. Of about 3,500 reports we have had just over 75 actions that the platforms have voluntarily taken. We have said, for instance, "This is a domestic violence victim and she is being cyber-stalked by her former partner who has an AVO." In cases like

that, when we can also make out that it is violating their own terms of service, they will act, but that was a very small proportion of reports that hit that bar. Again, we are not going to push the envelope without formal powers because we do not have any recourse or teeth.

Q246 **Kevin Brennan:** Under the proposed UK regime, Ofcom will be required to run a super-complaints process to consider systemic failings rather than individual content. What is your view about that? Is that something that might be adopted in Australia and with these kinds of schemes, in your view—because you have vast experience in this area, obviously—what is the best way to approach it? Is it via having some kind of super-complaints process about systemic failings or is it about focusing on the individual harms that people experience?

Julie Inman Grant: I think it has to be, for us, a combination of both. With our basic online safety expectations and our mandatory codes, we are moving to a more systems and processes-focused approach with a lot less in terms of penalties but the ability to compel transparency.

What I would say about the perceived shortcomings of only relying on a super-complaints scheme is that if you are not at the coalface working, the threat trends and harms analysis, which we are able to garner as a result of taking complaints from the public, is not necessarily something that you can achieve through a survey or even analysis of tweets and that sort of thing. This might be the type of thing that we would talk to Ofcom about. We are all very dedicated and we know that we do not want to work in isolation. Each of our regimes is going to have relative strengths and weaknesses. For instance, if this goes through Ofcom it will have some very, very potent tools at its disposal.

On our end, I think we need to continue evolving and strengthening our law. Where the Online Safety Bill and Digital Services Act is going is very much in that systems and processes area, so I believe we need a bit of both.

Q247 **Kevin Brennan:** If you were to make a judgment about what level of resources ought to be devoted to each regulator in an ideal situation, would it be a 50:50 division of resources between those two approaches?

Julie Inman Grant: You are going to need a lot more money and a lot more resources and a lot more specialised staff to do the systems and processes work. I was reading the Ada Lovelace paper, an excellent paper, where they were thinking about what technical audits and technical regulation might look like. From what I can tell, Ofcom has done a really great job at preparing and hiring talented staff. Just as there is a shortage out there right now, in terms of technical challenge over all, because this is a new frontier I think it is going to be hard to find people with those skillsets.

Q248 **Kevin Brennan:** Have they been trying to headhunt you?

Julie Inman Grant: I would not be analysing the black box. I think we all need to have more questions about what that means, what regulating algorithms means and whether we are doing that for the outcomes and the harms or we are trying to analyse the source code and the training sets. Analysing source code is not going to help you get to where you need to be. So, there is a lot more exploration. There is probably going to be a lot of experimentation that happens. That is what we have done and that is why we have had a basis for improving and updating and modernising our legislation, by seeing the gaps. Implementation is often very different than policy.

Q249 **Kevin Brennan:** On the issue of violence against women and girls, it is not on the face of the UK Bill. Are there lessons that we can learn from Australia on that subject? I know you have mentioned this already, but to be clear about what the position is in Australia about altered and manipulated images in abuse defences—could you cover that?

Julie Inman Grant: Our image-based abuse scheme does cover deepfakes and manipulated images and, of course, a scenario might be someone's face morphed on to a porn star's body. We will have to test our powers on deepfakes that might be used for political purposes. You may have seen a news article where there is a deepfake of Jacinta Ardern, the Prime Minister of New Zealand, smoking a crack pipe. We would have to look and see if deepfakes of that nature would either contravene some of our criminal laws or our adult serious cyber-abuse scheme. There would have to be a lot of mitigating factors.

In terms of women and vulnerable communities, we know that women and those of intersectional backgrounds are three times more likely to be targeted with online abuse. It is often rooted in misogyny or racism and prejudice, and the abuse targeting women manifests itself differently than it does targeting men, frankly. It is sexualised; it is violent; all focused on things like supposed virtue, fertility and appearance. It is designed to silence women's voices. We can use some of these intersectional factors to establish whether serious harm is being intended, but we don't have hate speech laws, per se. What we do have are a range of programmes.

We have a social media self-defence programme called Women in the Spotlight. That is for female journalists, politicians, sporting figures, entertainers, professional women who are in the spotlight, because 35% of professional women have told us that they self-censor and 25% have said they won't go into public life or take their promotion if it requires them to have an online presence because they almost expect that that is going to happen. The more we normalise that behaviour, the more we are going to further entrench gender equality.

We also have a programme called eSafety Women that trains domestic and family violence workers in terms of how mostly male predators are using coercive control and technology-facilitated abuse to further coerce,

control and surveil their partners. This happens in 99.3% of domestic and family violence cases.

Q250 **Julie Elliott:** Good evening, Julie, and thank you for joining us. I am interested in how important it is for the regime to cover content where the subject has previously agreed to that content being used but then changes their mind. You have talked a little bit about relationship retribution but it is a much wider area than that. Could you tell us that, please? Thank you.

Julie Inman Grant: Again, we assess every case based on the specific facts of the case. There are definitely scenarios where somebody may have been in a loving, or what they thought was a trusting, relationship and when that has gone sour, relationship retribution has been the result. They may have said yes once but it is still image-based abuse if it is shared online without consent. There are certainly scenarios we see with, say, younger people like teenagers, where they may be big-noting themselves or they think it is a laugh. However, when you are talking about adults, the intent behind sharing those intimate images is usually malicious, and what we have to do is look at the impacts.

Where we have some greater challenges is, for instance, if somebody who was on OnlyFans and might have created material for commercial means. Sometimes that content can get taken out and spread virally, which means it undercuts their revenue stream. We have to look at the specifics of the situation but that is a little bit harder to make out.

Q251 **Julie Elliott:** Do you think that any regime should cover the ability of anyone to change their mind about content being changed in whatever circumstances, so they have agreed at one point and changed their mind?

Julie Inman Grant: Yes, I think that is fair—context changes; situations change—if their intimate image is out there online without their consent or being shared. In most cases, people are consenting to the image or the video to be shared between the intimate couple, not more broadly. It is still a violation of consent and a huge invasion of privacy, so we would always err on the side of the victim survivor.

Q252 **Jane Stevenson:** Thank you, Julie, for a very interesting evidence session this morning. I want to go back to cyber-bullying and children. You have spoken about how important it is that bullies don't have the ability to follow children home from school and continue abusing them. You have also mentioned remedial actions in removing content, and you did touch on punitive measures on punishments. Because we are dealing normally with children, how did you find a balance when you find it is a child who has been abusing another child? What punishments did you consider and what ones are in place now?

Julie Inman Grant: That is a great question and very intuitive. We know, based on research, that one in five Australian children are cyber-bullied online. The average age is 14. Girls are bullied more than boys and, in almost all cases of youth-based cyber-bullying, it tends to be

peer-to-peer or an extension of conflict that is happening within the school gate. Of course, that contrasts with the vast majority of adult cyber abuse. You see a combination, but a lot of adult cyber abuse is strangers or people that are not really known to you.

We do have what we call an end-user notice. In some corners we have been criticised for not using them but we look at every case. We don't just go in there and say, "Hey, we are going to take this content down. Our job is done." In a number of cases, because we know it involves members of the student body, the school community, often educators are involved, or parents will come to us. We will sit down with the school community and try to get to the root of the problem and surface it up.

Once we take down the content and the root of that is settled, that is not something we can scale massively. Also, when there is a serious cyber-bullying attack or an incident that needs to be managed, we have an education outreach team that will go in. We go and deliver presentations to schools and teachers but we will do special interventions to make sure that we are going into that particular school community where something pretty toxic has happened and talk about people's options, but also talk about repercussions.

We have not yet used the end-user notice because we know that, if we are issuing an end-user notice, it is likely going to a child. Once you get into the upper age ranges, 17 and 18, when they are closer to adulthood and they probably know the difference and are trying to cause harm, we will look at all these factors. We will look at mental health issues. There was a time we were looking at issuing an end-user notice and then, in speaking to the school community, we learned that this child was in out of home care, had been abused and had come from a very difficult home. Our thinking was that if we issued a notice that would have spill-on effects that would be worse for them. Therefore, we will use other methods than something as menacing as a notice from a Government entity.

We have seen some pretty insidious ways that young people are bullying other people. I am thinking of a case—it is called Phoenixing—where a child will create up to 50 to 60 accounts with that child's name in it, so every time one is suspended, the next one will pop up and they can keep targeting the child. One child experienced that for weeks and months on end and it was relentless for that child.

We saw last month somebody going to the effort of writing a song targeting another child, recording it and putting it up on Spotify. Spotify took it down quickly, but here I go back to endlessly creative ways for people to target others. Now 30% of all youth-based cyber-bullying is happening on private encrypted messaging, so on DM.

Q253 **Jane Stevenson:** In the most serious cases, no punishments have been considered or financial penalties that would apply to parents, or did you ever consider criminalising over-16s, or prosecuting some method in

really serious sustained cases of cyber-bullying?

Julie Inman Grant: We have something called 474.17 of the Criminal Code, which is criminal sanctions for a person who uses a carriage service provider to menace, harass, or cause offence. There have been cases and, again, we have MOUs with all the state territory and federal police, so we will refer some of those more serious ones where we think it requires criminal follow-up. That is pretty rare.

The other thing that has been great about having our civil schemes is in the grand scheme of things, where law enforcement is concerned it has to triage too when dealing with murders, drugs and rapes. Fighting online crime is very hard and taking on cases is hard, so a large proportion would fall through the tracks and people would not receive any support without the services that we provide.

Q254 **Jane Stevenson:** The complaints procedure that is open to children isn't open to adults. Do you think that is something that may change or do you think adults need a different route to go through?

Julie Inman Grant: On 23 January, we will be implementing our new law and introducing our serious adult cyber-abuse scheme, which is at the higher threshold and is for any serious cyber abuse targeting an adult. We have to prove serious intent to harm and that it is menacing, harassing or causing offense in all cases. There is a remedy there. We have to recognise that adults have more resilience so it is at a much higher threshold, but it will be there and, hopefully, we will have some success.

Q255 **Jane Stevenson:** I hope so, yes. Finally, you mentioned BSI requests for anonymous accounts. If that is part of a serious adult criminal scenario, if that request comes back and it is a 14-year-old child who has been doing some really serious abuse, where does that travel? Does that go back to a children's—

Julie Inman Grant: No, if we glean that information, I think it would be unusual. We have not seen a case where, say, a teenager has been targeting and seriously causing distress or harassing an adult, but we would look at the whole range of powers we have at our disposal. Just because someone comes in an image-based abuse route, if it does not meet that threshold, it may be youth-based cyber-bullying, or it may become child sexual abuse material if it is a child under the age of 18; we can use this broad suite of powers for the different schemes to meet the particular case or the circumstances of a given case.

Q256 **Clive Efford:** Thank you for giving evidence to us. It is much appreciated. Do you think your Safety by Design framework can meaningfully embed safety into the culture of big tech companies and riskier niche sites?

Julie Inman Grant: I certainly hope so but, listen, I think the process is very important. I don't know if you know this about me but I spent 22

years in the technology sector. I brought Safety by Design to Microsoft when I was working there 10 years ago. I had the eye roll and, "We are not becoming a social media company; we are becoming an enterprise company" and I think that sentiment has changed.

We sat down with 80 companies to write up a set of principles that were actionable and meaningful. They were all around the issues we are talking about: what is service-provider responsibility? What does user empowerment and autonomy look like? So, how are you empowering users? What does radical transparency and accountability look like, rather than selective? Principles are great, we got them behind. However, there are a lot of principles out there and principles are only effective if they are being implemented, so we developed a series of risk-assessment tools over an 18-month period with 180 organisations, what we called the Safety by Design Interactive Assessment Tools, one for start-ups and one for enterprise companies.

I just focus on the one for start-ups right now. Think about Zoom that we are on now. In December 2019 they had 10 million daily active users. After Covid hit in March 2020, they scaled up to 200 million users but I don't know if you remember at that time, we started having Zoom-bombing attacks and there were questions about their privacy, security, and safety so they took it offline. My Government agency and I are still not allowed to use Zoom as a result of that loss of trust.

It was really telling when the CEO said after that event, "Oh, I never really thought about online harassment before." You have to think to yourself, "How do you create a video conference platform where people interact? You have human behaviour in a frame, and you would not think that anything could possibly go wrong," but that is how technology founders think. They are thinking about their technology, how they get it out to market and how they keep users on their platforms.

I think there has been a groundswell and a change. I think Safety by Design is taking off. These companies need to be able to assess their risks and make safety a forethought rather than an afterthought. Safety by Design is fixing the foundations of the platforms and the internet, just like we pushed car manufacturers 55 years ago to invent seatbelts. This technology exceptionalism has to end and we need to move from that moving fast and breaking things mantra to an ethos of mindfully building in safety protections. With actions like this—Governments have been calling for it—they are going to be in a much better place if they are embedding Safety by Design to be able to comply with these stronger laws that are coming into place.

Q257 **Clive Efford:** Is there a danger that, because it is a voluntary code, the Safety by Design becomes safety driven by what is in the interest of the platforms rather than customer safety, if I can put it that way?

Julie Inman Grant: I have been pretty impressed. There are a range of companies that are really embracing it and I go in and I look at how they

are designing innovative new safety features all the time. Not all companies are created equal and a lot of that has to do with the leadership. If Mark Zuckerberg, for instance, decided he wanted to have a really safe platform and a really safe metaverse, he would be applying privacy and Safety by Design and he would achieve it.

If you can target people with deadly precision the way they do with advertising, you can certainly do the same in terms of eradicating harm and engineering out misuse. There is some carrot in there because I do not believe that, as Governments, we are going to be able to tell them the best way or these specific protections or innovations to put in there. We do need some carrots. We do need to bring them along because we need to change the culture and that has to start at the top.

Q258 **Clive Efford:** How does the news media bargaining code sit in your toolkit of regulations alongside the eSafety regime?

Julie Inman Grant: That sits in the ACCC, which is our competition authority. I think the UK also has a system where CMA, Ofcom and ICO, the regulators in the digital space, sit together and talk about the spectrum of harms and where their respective areas sit. We do the same thing with our Privacy Commissioner and the ACCC. The media bargaining code in Australia was very innovative, hard-hitting and effective. I think you saw it did push Facebook to the brink and it cut things off for a short period of time, and that created quite an uproar, globally. It has gotten on with it and both Google and Facebook have done a number of deals with a range of media companies, so it has worked.

Q259 **Clive Efford:** That was an extraordinary period of time when Facebook took that action back in February. What does that say about big tech companies and their attitude to these codes of conduct and regulations?

Julie Inman Grant: I was at Twitter when the Online Safety Act was first created and I am a pragmatist. The Minister at the time invited all of the major platforms in to have a view of the Bill before it was released. For some reason, I was the only one that showed up. I thought, "Well, if they are going to write a Bill, I might as well be there to shape it" but Google and Facebook, I think you are looking at a tiered scheme.

Facebook and Google declined to openly co-operate with the eSafety Commissioner when they had the opportunity to sign on. My belief is that is because they did not want to create a domino effect or say that they endorse working co-operatively with a regulator, but I think that the horse has really bolted. They said that innovation would be quashed and "We will leave the market" and all sorts of Armageddon was going to happen. None of it has come to pass.

We heard some of the same arguments around the updating and strengthening of the Bill, so you are hearing positive signs, companies saying, "Oh no, we need to be regulated. We do not need to be arbitrators of the truth." What I hear when I hear that is, "We know we

are going to be regulated, but regulate us the way we want to be regulated."

Q260 **Clive Efford:** Do you think we will get similar pushback here in the UK as we attempt to regulate, and does that suggest, if we are going to take on these big tech companies and get them to act reasonably, that we are going to have to co-operate internationally?

Julie Inman Grant: Absolutely. That is one of the conversations that I had these past couple of weeks with Ofcom and with the European Commission. Think of the collective wealth and power just of the five top companies in terms of nation states. We are going to have to work together and we are not under any illusion. We think we are making a difference for our citizens and we have seen Safety by Design taking off.

We are seeing companies adopting it and doing some good things with it. I think we need to encourage them to continue doing better because we are never going to make the internet a safer place if they are not erecting the digital guardrails and embedding the virtual seatbelts. We will just be playing a big game of Whac-A-Mole and making these big enforcement actions. We will always be responding by definition, so we need to get ahead of it. We need to shape the future. Let us think about the metaverse and the decentralised world that everyone is talking about. We need to take off our utopian glasses and think about how we are going to look at internet governance for these new worlds. How are we going to build in safety, privacy and security now, rather than ending up in a world that is just rife with harms?

Chair: That is a very appropriate moment to leave it. Julie Inman Grant, Australian E-Safety Commissioner, thank you very much for your evidence today. We managed to get through our technical difficulties. I imagine your voice is now quite hoarse having had to shout for the last hour, but thank you very much. It is much appreciated. That concludes our final session before Christmas.