Tuesday, 22 June 2021

Committee Secretary
Legal Affairs and Safety Committee
Parliament House
Alice Street
Brisbane QLD 4000

By email: lasc@parliament.qld.gov.au

Dear Committee Secretary,

As Australia's eSafety Commissioner, I welcome the opportunity to provide a submission to the Queensland Government's inquiry into serious vilification and hate crimes in Queensland (inquiry).

I am writing to you to outline the powers and functions I have as eSafety Commissioner to protect Australians from online hate speech and vilification, as well as eSafety's overall approach to these issues, which is evidence-based and focused on prevention and proactive change

eSafety

eSafety is Australia's national independent regulator for online safety. Our core objective is to minimise harm to Australians online.

eSafety is the first government agency in the world dedicated specifically to online safety. We lead, coordinate, educate and advise on online safety issues and aim to empower all Australians to have safer, more positive online experiences.

As you may know, on 23 December 2020, the Australian Government commenced a consultation on a Bill for a new Online Safety Act (the Act). The proposed new Act will expand eSafety's regulatory remit and my functions and powers as eSafety Commissioner, improving the effectiveness, reach and impact of eSafety's work.

At-risk groups

eSafety undertakes an extensive research program to ensure its programs and resources are evidence based. This equips us with the insights and knowledge we need to understand the nature of online safety issues and design, implement and evaluate best possible solutions.

eSafety's research and experience points to the fact that online harms can disproportionately impact at-risk and diverse groups. This includes, but is not limited to, Aboriginal and Torres Strait Islander people, people from culturally and linguistically diverse communities, people with disability and people who identify as LGBTQI+, as well as, depending on the circumstances, women, older people and children and young people.

Of particular relevance to the Committee is eSafety's research into online hate speech.[1]

---

[1] eSafety, Online Hate Speech, February 2020, https://www.esafety.gov.au/sites/default/files/2020-01/Hate%20speech-Report.pdf.
The research was based on the following question: In the last 12 months, how many times, if ever, have you received a digital communication that offended, discriminated, denigrated, abused and/or disparaged you because of your personal identity/beliefs (e.g. race, e hnicity, gender, nationality, sexual orientation, religion, age, disability, etc)?

This research found that around 1 in 7 adult Australians aged 18–65 (14%) were the target of online hate speech in the 12 months to August 2019: this is around 2 million people. While this is a staggering statistic, it is also important to disaggregate this overall figure, in order to highlight the intersectional nature of online abuse.

People identifying as LGBTQI or Aboriginal or Torres Strait Islander experience online hate speech at double the national average. In general, people experiencing online hate speech identify their political views (21%), religion (20%) and gender (20%) as the top three reasons for being targeted. However, terms like race, ethnicity and nationality can be interchangeable from a respondent's perspective. Taken together, these reasons account for 32% of people experiencing online hate speech.

The findings were particularly stark in relation to sexual orientation. Sixty-one per cent of those who identify as LGBTQI report that their sexual orientation was the reason for being the target of online hate speech, compared to their gender (35%) or political views (30%). Most people were unable to attribute responsibility for their online hate speech experience to a specific person, with 47% assigning blame to a stranger and 13% reporting that they didn't know who is responsible (13%). People identifying as LGBTQI (76%) are considerably more likely to identify a stranger as the source of online hate speech than any other group. An estimated 58% of those personally experiencing online hate speech report a negative impact from their experience. Thirty-seven per cent report mental or emotional stress as a result of their experience, while 10% report reputational damage. People identifying as LGBTQI were more likely to report mental or emotional stress from online hate speech than other groups.

eSafety strategies, programs and resources are tailored to the needs of diverse people and communities. In August 2020, eSafety released our Protecting Voices At Risk Online report. This statement outlines eSafety's current initiatives and defines our strategic priorities for ongoing protection of those who are most at risk of online harm. We recognise that we need to remain innovative, agile and culturally responsive in order to meet the changing online safety needs of Australia's diverse and continually evolving population.

Current regulatory approach to hate speech

Under Schedules 5 and 7 of the *Broadcasting Services Act 1992* (Cth), the Online Content Scheme creates a report-based scheme about illegal and offensive online content. Using the same classification criteria applied to films, television programs and computer games, eSafety undertakes regulatory investigations to determine whether online content is prohibited under Australian law.

Our focus is on the most serious subsets of prohibited content, being child sexual abuse material, pro-terror content and material that promotes or instructs in matters of crime. Hate speech that incites violence against a particular group may be considered prohibited content.
eSafety is empowered to give takedown notices to the relevant hosting company where the content is found to be hosted in Australia.

Additionally, in response to the live-streamed attack in Christchurch, New Zealand on 15 March 2019, the Australian Parliament passed the *Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019*.

Abhorrent Violent Material (AVM) is defined as audio and/or visual material that records or streams a terrorist act involving serious physical harm or death, the murder or attempted murder of another person, the torture of another person, the rape of another person, or the kidnapping of another

person involving violence. It is material produced by a perpetrator or an accomplice – it does not include bystander coverage.

There are a number of defences in relation to the sharing of AVM online, including when it is done by journalists, law enforcement agencies, public officials, or for research, advocacy or artistic purposes.

The AVM regime gives the eSafety Commissioner the power to issue a notice to any website publishing AVM and/or the service that hosts that website. Rather than require the eSafety Commissioner to monitor the entire internet for AVM, it is predominantly a complaint-based regime.

Importantly, this is not a power to take down material. Further, the notices do not require the AVM to be removed. However, if a service is later criminally prosecuted for failing to expeditiously remove or cease hosting AVM, the notice can be used in legal proceedings to establish the relevant fault element under the Commonwealth Criminal Code of recklessness.

We have found these notices to be a very effective tool for alerting services to the presence of AVM so they can take appropriate action to protect their users from harm. To date, the eSafety Commissioner has issued 23 notices in relation to content depicting beheadings, shootings and other murders, with the content taken down or restricted for Australian users in 93% of the cases.

Hate speech and new Online Safety Act powers

There are several additional powers being provided to eSafety through the new Act that may assist in addressing online hate speech and serious vilification.

The Online Content Scheme is being strengthened to allow stronger action to be taken in response to class 1 material. The Act defines class 1 material as material that is likely to be Refused Classification (RC) under the Classification (Publications, Films and Computer Games) Act 1995. This includes publications, films or computer games that directly or indirectly counsel, promote, encourage or urge the doing of a terrorist act; directly or indirectly provide instruction on the doing of a terrorist act; or directly praise the doing of a terrorist act in circumstances where there is a substantial risk that such praise might have the effect of leading a person to engage in a terrorist act.

The draft Bill empowers the eSafety Commissioner to give a removal notice to a social media service, relevant electronic service, designated internet service or hosting service requiring them to take down class 1 material within 24 hours, regardless of whether the service is provided from Australia or overseas.

The Adult Cyber Abuse scheme establishes a report-based scheme to tackle menacing, harassing or offensive online content directed at a specific Australian adult, in a way that is likely intended to cause serious harm. Hate speech directed at groups (rather than an individual) will be beyond the scope of the Adult Cyber Abuse scheme. However, hate speech and racism may be relevant to eSafety's assessment that material is intended to cause serious distress or other forms of serious harm to the target.

The Act provides for the development of Basic Online Safety Expectations (BOSE) for social media services, relevant electronic services, and designated internet services. The proposed BOSE includes the expectation that providers of services will take reasonable steps to minimise the extent to which class 1 content, AVM and material that promotes, incites or instructs in abhorrent

violent conduct is provided on the service. It also includes the expectation that the service has clear and readily identifiable mechanisms that enable end-users to report such material.

The Act enables the eSafety Commissioner to require online services to provide specific information about their compliance with the BOSE. This could include requiring services to explain how they are working to minimise hate speech that may constitute cyber abuse, class 1 material or material that depicts, promotes, incites or instructs in abhorrent violent conduct on their service.

Whole of community approach

eSafety's investigative functions provide a means to respond to serious harm online. At the same time, eSafety is also very focused on preventing harm online.

Especially in the context of this inquiry, it is important to consider the structural, systemic and social factors that underpin serious vilification and hate online. A whole of community approach and systems approach is therefore needed to understand and address the underlying drivers of this behaviour.

eSafety also focuses on digital capacity building: giving individuals the skills and strategies to prevent and respond to harmful experiences online and engage online in ways likely to promote safe and positive online experiences. Capacity building should be a lifelong process that begins at the earliest age possible. It should occur at the individual and community level and at a societal and cultural level. In other words, it needs to focus on building the capacity of the individual, but also of communities and society to understand, recognise and respond to harm online, including serious vilification and hate online, and promote safer and more positive experiences.

eSafety has an extensive education and outreach program to support this stream of work. The four Rs of online safety — respect, responsibility, resilience and reasoning — are a basis for examining online information and making an informed judgement on an issue. eSafety has also developed a Best Practice Framework for Online Safety Education based on an evidence review. The review found that a sound online safety education should cover the full range of potential issues, risks and harms that children may encounter and should be delivered in supportive school systems with strong partnerships with other agencies.

eSafety also has a world-leading Safety by Design (SbD) initiative, which aims to drive up standards of user safety in the technology community. SbD emphasises the need to address online harms, alongside user safety and rights, in the design, development and deployment of technological solutions. Digital platforms and services can take preventative steps to guard against their services being used to facilitate, inflame or encourage illegal and inappropriate behaviours, including serious vilification and hate online. Additionally, these platforms can be designed to enable, optimise and support users to have greater control over their activities online.

Coordinated approach to online safety

As Australia's leader in online safety, eSafety develops and delivers a national strategic response to combat online harm. eSafety is pleased to work with state and territory governments to ensure their efforts support and strengthen a nationally coordinated approach and avoid duplicative or fragmented approaches.

Online vilification and hate are complex issues with a range of social, cultural and behavioural underpinnings and drivers. As with all our work, eSafety supports a multi-faceted approach

that explores how technology can both be a tool for, and extend upon, these underlying issues, while also serving as a means for positive and safe online experiences.

I commend the Committee for undertaking this important inquiry and look forward to the outcomes it will produce.

Yours sincerely,

Julie Inman Grant
eSafety Commissioner