

SELECT COMMITTEE ON SOCIAL MEDIA AND ONLINE SAFETY

Questions on Notice – February 2022

Question 1 – What research has been done?	
<ul style="list-style-type: none"> What research have you undertaken to inform yourself on online harms? 	<ul style="list-style-type: none"> eSafety undertakes a range of research. Please see Attachment A for a list of research undertaken since the inception of eSafety Commissioner in 2016.
<ul style="list-style-type: none"> What is the annual research budget for online harms? 	<ul style="list-style-type: none"> eSafety's research budget for the 2021-2022 Financial Year is \$1.5 million.
<ul style="list-style-type: none"> How many research projects have been conducted? Please provide a list. 	<ul style="list-style-type: none"> 35 research projects have been conducted since 2016, please refer to Attachment A. A small number of projects produced infographics rather than reports.
<ul style="list-style-type: none"> How many reports have been produced? Please provide a list. 	<ul style="list-style-type: none"> 37 reports have been produced, please refer to Attachment A. This total includes unpublished reports that were used for our own purposes. Please note some projects led to multiple reports.
<ul style="list-style-type: none"> How many reports have been published? Please provide a list and copies of reports. 	<ul style="list-style-type: none"> 29 reports have been published. All published reports are available on the eSafety website. They can be accessed from https://www.esafety.gov.au/research. The table at Attachment A also cites published reports, unless specified.
<ul style="list-style-type: none"> What harms were analysed? <p>eSafety research covers a broad range of online harms in a range of contexts such as online gaming, social media platforms, messaging services, family and domestic violence situations. Harms covered by eSafety research are:</p> <ul style="list-style-type: none"> cyberbullying image-based abuse technology-facilitated abuse sexting viewing inappropriate material (such as young people accessing pornography or violent content) hate speech online abuse personal information being used in a harmful way 	

<ul style="list-style-type: none"> • accounts used without consent (including hacking) • receiving threats of real-life harm or abuse via online means • an individual's experience of someone pretending to be them online. <p><i>eSafety research also explores specific Australian populations:</i></p> <ul style="list-style-type: none"> • Australian adults • women • Australian children and young people • older Australians • Aboriginal and Torres Strait Islander peoples • people who identify as LGBTIQ+ • CALD communities, and • people with disability. 	
<ul style="list-style-type: none"> • How were these harms analysed? <ul style="list-style-type: none"> • Was any qualitative or quantitative research undertaken? • Was the prevalence and/or impact analysed? 	<ul style="list-style-type: none"> • eSafety uses a range of methodologies in our research program, including qualitative, quantitative and mixed-method approaches. • The research design used for a specific project is determined based on several factors, including the research question(s) being asked, the population of focus, existing evidence and other ethical and practical considerations. • Prevalence of online harms and their impact on citizens are explored in many eSafety research projects. Research aims and specific research questions are used to identify whether these factors will be investigated in particular projects.

Question 2: Research Gaps	
<ul style="list-style-type: none"> • What research gaps are there? 	<ul style="list-style-type: none"> • There are many areas of research that could be conducted in the online safety space given how quickly technology and issues evolve. • A current research gap is the need for rigorous, independent research into the effects of social media use particularly for children and young people, as the current evidence base is mixed. In particular, causal and longitudinal research is needed to inform our understanding of cause and effect relationships – both positive and negative. • This evidence can inform changes and solutions which elevate the positive aspects of social media use and diminish its negative aspects.

	<ul style="list-style-type: none"> eSafety will continue to evaluate solutions aimed at preventing and reducing online harms and assessing the effectiveness of interventions.
<ul style="list-style-type: none"> What harms have not been analysed? 	<ul style="list-style-type: none"> eSafety research covers a broad range of harms. We routinely track new and emerging technologies, trends from our Investigations Branch and evidence generated from academic and other institutions to inform and prioritise our research program. The extent to which algorithms have the potential to contribute to online harms is an area for future exploration subject to funding.
<ul style="list-style-type: none"> What harms is research being planned for? 	<p>The following harms research projects are underway or planned to commence in the Financial Year 2021 – 2022 to inform our programmatic work:</p> <ul style="list-style-type: none"> Australian adults online survey – digital participation and online harms Older young people and online pornography Literature review to inform online safety education for CALD families recognise and prevent harmful behaviours online, and Young people’s understanding of safety in the metaverse.

Question 3: Specific Research	
What research have you done in relation to the following?	
<ul style="list-style-type: none"> Online hate speech against groups (e.g., race or religion) 	<ul style="list-style-type: none"> Online hate speech has featured as a harm category in both our youth and adult surveys over the years – tracking top line trends. In 2020 eSafety released the <i>Online Hate Speech</i> report using data from the adult survey conducted in 2019. The report was formed part of a joint international reporting effort between the eSafety Commissioner (eSafety) in Australia, Netsafe in New Zealand and UK Safer Internet Centre (as part of the European SELMA project) in the United Kingdom. Our Women in the Spotlight research also quantifies online hate speech experienced by women who are currently or have recently been employed.
<ul style="list-style-type: none"> Radicalisation (including in conjunction with the Department of Home Affairs and/or security agencies) 	<ul style="list-style-type: none"> Primary research responsibility in this area is not with eSafety. The Social Cohesion research conducted in 2016 (funded through a specific cross portfolio allocation as part of the Government’s broader Countering Violent Extremism agenda) looked broadly at issues such as exposure to hate speech and terrorist content but did not explore radicalisation.

<ul style="list-style-type: none"> Sexual assaults facilitated by social media apps (including in conjunction with other Departments/agencies such as Health, Social Services, Office for Women) 	<ul style="list-style-type: none"> No primary research has been conducted by eSafety.
<ul style="list-style-type: none"> Eating disorders and social media (including in conjunction with the Department of Health) 	<ul style="list-style-type: none"> No primary research has been conducted by eSafety. Some desktop research has been conducted along with environmental scanning.
<ul style="list-style-type: none"> Suicidal ideation 	<ul style="list-style-type: none"> No primary research has been conducted by eSafety. Some desktop research has been conducted along with environmental scanning.

Question 4 – Online Harms	
<p>What, if anything, is eSafety doing to address the following?</p>	<ul style="list-style-type: none"> eSafety takes a holistic approach to online harms across the pillars of prevention, protection and proactive and systemic change. The new Online Safety Act strengthens our ability to prevent and address a variety of harms and provides important new options for eSafety to regulate online services' systems and processes. For example, under the Basic Online Safety Expectations, providers are to take reasonable steps to ensure that end-users can use the service in a safe manner, and to proactively minimise the extent to which material or activity on the service is unlawful or harmful. eSafety can require providers to report on their compliance with the expectations, including in relation to the issues specified below, and may publish statements about the extent of their compliance. eSafety's Safety by Design Assessment tools guide and support technology companies to proactively consider and embed safety into the culture, ethos and operations of their businesses. The tools cover key online safety considerations across thematic areas of Online harms, Structure and leadership, Internal policies and procedures, Moderation, escalation and enforcement, User empowerment and Transparency and accountability. The Online Harms module was specifically developed to educate technology companies about the types of online harms that can occur, the impacts on victims, techniques and tactics – modes of online abuse, data, privacy and security, threats, manipulation and deception, contributing factors – who might be at risk?, victim factors, perpetrator factors, as well as the intersectionality of online harms and related issues.

<ul style="list-style-type: none"> Online hate speech against groups (e.g., race or religion) 	<ul style="list-style-type: none"> eSafety is working to help Australians experiencing online hate using its three-p model: <p>Prevention – Research and education:</p> <ul style="list-style-type: none"> eSafety is working to better understand and address online hate through a variety of research and other initiatives aimed at protecting voices at risk online. As outlined above, this includes research conducted in 2020 in relation to online hate speech. It also includes education resources for young people, the development of tailored resources for individuals from culturally and linguistically diverse backgrounds, work to support LGBTQI+ audiences, Aboriginal and Torres Strait Islanders, and work to support Islamic women in dealing with image-based abuse that may involve a religious or cultural dimension. <p>Protection – Adult cyber abuse:</p> <ul style="list-style-type: none"> Hate speech directed at groups (rather than an individual) will be beyond the scope of the new Adult Cyber Abuse scheme. Hate speech targeting a <i>specific</i> person (i.e., an Australian adult) may meet the adult cyber abuse threshold. For example, when investigating adult cyber abuse, eSafety will consider whether menacing, harassing or offensive content was targeted at someone because of their racial or cultural background, and if it was likely intended to cause serious harm. <p>Protection – Online content:</p> <ul style="list-style-type: none"> Under the Online Safety Act, there may be grounds to consider some hate material as class 1 material. For example, hate speech against individuals or groups that promotes, incites or instructs in matters of crime or violence may fall within the Online Content Scheme and be subject to removal. <p>Proactive and systemic change – BOSE:</p> <ul style="list-style-type: none"> Where complaints reveal systemic issues – such as a failure by services to enforce their terms of use (including terms of use relating to online hate) – eSafety will have the power to drive services to do better through reporting and public statements under the Basic Online Safety Expectations.
<ul style="list-style-type: none"> The rise of right-wing extremism online including the radicalisation funnel 	<ul style="list-style-type: none"> eSafety notes that while portfolio responsibility for this policy measure is primarily the responsibility of the Department of Home Affairs, eSafety has equity in this policy space. Where right-wing or other forms of extremist online content promotes, incites or instructs in matters of crime or violence, it may fall within the Online Content Scheme and be subject to removal as class 1 material. In addition, material that advocates the doing of a terrorist act will be considered class 1 material. As our submission to this Inquiry explains, under the Online Safety Act, there are two specific regimes of relevance to

	<p>recommendation algorithms which may contribute to a ‘funnel’ of increasingly extreme content.</p> <ul style="list-style-type: none"> • The first is the development of industry codes to prevent and limit harmful online content relating to class 1 and 2 material. One of the measures recommended in eSafety’s position paper guiding the development of codes is to reduce the promotion and reach of this content within algorithmic systems, including recommendation algorithms and choice architecture. Registered codes will be mandatory and enforceable enabling eSafety to investigate potential breaches and direct services to comply. • The second is the power to require reporting from services about how they are meeting the Basic Online Safety Expectations. The effects of algorithms and how they relate to user-safety will be something we consider as eSafety implements the BOSE and reporting requirements. For example, information about the outcomes and intent of recommendation algorithms could be relevant to the expectations for minimising harmful material and making sure people can use the service in a safe manner. • In addition, our Safety by Design initiative, aims to guide and support industry to consider harms upfront and put user safety and rights at the centre of how they design and develop products and services. The next phase of our Safety by Design activities will consider the extent to which algorithms have the potential to contribute to online harms and highlight this as a factor for platforms to include in their risk-based thinking. • eSafety also participates in the Global Internet Forum to Counter Terrorism’s working group on transparency. One of the topics this group is considering is algorithmic transparency. • eSafety received a discrete funding allocation from the Attorney General’s Department in 2016 to develop programs aimed at building social cohesion. Materials such as <i>Lost Summer</i> and the <i>Young and eSafe</i> resources were developed with this funding.
<ul style="list-style-type: none"> • Sexual assaults facilitated by social media apps 	<ul style="list-style-type: none"> • eSafety has engaged with multiple dating app companies and other platforms where sexual assault may be facilitated. These discussions focused on the responsibilities to provide clear reporting pathways and the steps that companies were taking to stop repeat offenders returning to the platforms. This is an area for further Basic Online Safety Expectations consideration. • eSafety’s Safety by Design tools encourage industry change through highlighting innovative user safety practices, such as Bumble’s ‘Unmatch’ and ‘Private Detector Tool’.

	<ul style="list-style-type: none"> • Future Safety by Design activities include targeted consideration of dating apps, specifically seeking to address user needs and expectations. • The Australia New Zealand Policing Advisory Agency has a working group on the issue of online facilitated sexual assault. • eSafety has attended several meetings with this group, most recently in September 2021, to discuss relevant eSafety work streams and areas for potential collaboration.
<ul style="list-style-type: none"> • Eating disorders and social and traditional media impact 	<ul style="list-style-type: none"> • eSafety offers advice for young people and parents about dealing with content that may have negative impacts on young people's mental health. This includes advice on signs that online behaviour may be contributing to eating disorders. • eSafety advice covers matters such as applying critical reasoning skills to content accessed online, dealing with the pressures of social media, and encouraging help seeking. We also provide links through to support organisations such as the Butterfly Foundation, Kids Helpline and Headspace. • Where eating disorder content is reported to us which consists of detailed, gratuitous images of highly emaciated individuals such that it meets the Refused Classification threshold under the National Classification Scheme, our removal powers could be enlivened under the Online Content Scheme
<ul style="list-style-type: none"> • The impact of social media on youth mental health 	<ul style="list-style-type: none"> • The impact of social media on youth mental health is complex and more research is needed to establish whether or not there is a causal nexus. • Social media use by young people is neither inherently good nor bad, but a balancing act unique to every young person and their circumstances. • eSafety provides assistance to children and young people experiencing cyberbullying and image-based abuse through our reporting schemes. This includes connecting children and young people with counselling and support services, such as Kids Helpline. • eSafety exists as a safety net to prevent these incidents occurring in future, by providing national leadership in online safety education and working with industry to make sure we consider safety at product inception and throughout the product lifecycle. • While eSafety continues to work on making platforms safer for people's mental wellbeing, we are also working to educate young Australians on communicating respectfully, showing care, empathising with one another, and understanding the impact our words and actions have on others.

	<ul style="list-style-type: none"> • We also need to further educate young people and parents on identifying broader online harms which may compound the adverse effects of mental ill-health and empower them to make positive choices online that enhance their protection from negative experiences. • For example, in 2021, eSafety delivered a parent webinar titled “Digital technologies and mental health”, which outlined scenarios related to young people's experiences online, such as a friend sharing content about an eating disorder, self-harm, or suicidal thoughts. The content covered practical strategies for starting conversations and links to mental health providers like The Butterfly Foundation and headspace for further support and resources. • eSafety also consults directly with young people to hear more about their online experiences and how we can shape education programs to meet their needs. We are establishing a Youth Advisory Council to provide a youth voice for government on online safety policies and programs.
--	--

Question 5 Additional QON from the Hearing	
<p>CHAIR: Thank you. You mentioned in relation to self-harm and the sending it through to the platforms—in the other areas I think there's probably a bit of a scale, if you like. In terms of the number of complaints that you received, or the volume, roughly what proportional percentage met the threshold of—I think it was referred to before as a high threshold—take down requirements?</p> <p>Mr Dagg: I might, if I could beg the indulgence of the committee, take that question on notice. Like I said, our figures for the last week and a half haven't been fully settled. We've been tracking complaints on a daily basis and tracking major issues so that we can understand how all the schemes have been impacted by the commencement. I'm very happy to take that on notice and provide you with some more specific information after speaking with the teams.</p> <p>Since the commencement of the OSA, we estimate, noting that not all matters have completed the investigation process, that less than 10% of complaints may meet the Adult Cyber Abuse threshold.</p>	

Attachment A

eSafety Research Projects – 2016- 2022

The following research projects have been undertaken since the inception of the eSafety Commissioner in 2016. eSafety uses a range of methodologies in our research program, including qualitative, quantitative and mixed-method approaches. The research design used for a specific project is determined based on several factors, including the research question(s) being asked, the population of focus, existing evidence and other ethical and practical considerations.

Published research is all available on the eSafety website: <https://www.esafety.gov.au/research>

Project	Year of Publication
Women in the spotlight: How online abuse impacts women in their working lives	Scheduled for 2022
Teacher Professional Learning program evaluation	Scheduled for 2022
Mind the Gap: Parental awareness of children's exposure to risks online www.esafety.gov.au/research/mind-gap-parental-awareness-childrens-exposure-risks-online	February 2022
Youth engagement and online safety www.esafety.gov.au/research/youth-engagement-and-online-safety	January 2022
eSafety core women program evaluation (interim report)	Unpublished 2021
Anonymity	Unpublished 2021
Community perceptions of adult cyber abuse	Unpublished 2021
Public perceptions of age verification for limiting access to pornography www.esafety.gov.au/research/public-perceptions-age-verification-for-limiting-access-pornography	Summary 2021
Online safety issues encountered by health, education and law enforcement workers www.esafety.gov.au/research/online-safety-issues-encountered-health-education-and-law-enforcement-workers	2021
Protecting LGBTIQ+ voices online https://www.esafety.gov.au/research/protecting-lgbtq-voices-online	2021
Technology-facilitated abuse among Aboriginal and Torres Strait Islander women in remote and regional areas www.esafety.gov.au/research/technology-facilitated-abuse-among-aboriginal-and-torres-strait-islander-women	2021
Technology-facilitated abuse of women with intellectual or cognitive disability	2021

www.esafety.gov.au/research/technology-facilitated-abuse-women-intellectual-or-cognitive-disability	
Best Practice Framework for Online Safety Education report (Stage 2) www.esafety.gov.au/research/best-practice-framework-for-online-safety-education-reports	2021
Evaluation of eSafety's Early Years program www.esafety.gov.au/research/evaluation-early-years-program	2021
Digital lives of Aussie Teens https://www.esafety.gov.au/research/digital-lives-aussie-teens	2021
Best Practice Framework for Online Safety Education report (Stage 1) www.esafety.gov.au/research/best-practice-framework-for-online-safety-education-reports	2020
Building Australian adults' confidence and resilience online www.esafety.gov.au/research/adults-confidence-and-resilience	2020
Children and technology-facilitated abuse in domestic and family violence situations www.esafety.gov.au/research/children-and-technology-facilitated-abuse-domestic-and-family-violence-situations	2020
Online safety for young people with intellectual disability www.esafety.gov.au/research/online-safety-for-young-people-intellectual-disability	2020
Adults' negative online experiences www.esafety.gov.au/research/adults-negative-online-experiences	2020
COVID-19 impacts on Australian adults' online activities and attitudes www.esafety.gov.au/research/covid-19-impacts-on-australian-adults-online-activities-and-attitudes	2020
Online hate speech www.esafety.gov.au/research/online-hate-speech	2020
Online safety for Aboriginal and Torres Strait Islander women living in urban areas www.esafety.gov.au/research/online-safety-for-aboriginal-and-torres-strait-islander-women-living-urban-areas	2019
Understanding the attitudes and motivations adults who engage in image-based abuse. www.esafety.gov.au/research/image-based-abuse-perp-motivations	2019
Parenting in the digital age report (which includes an infographic series) www.esafety.gov.au/research/parenting-digital-age www.esafety.gov.au/research/digital-parenting	2019

Encouraging digital participation of older Australians through mentoring www.esafety.gov.au/research/mentoring-older-australians	2019
eSafety for Women from Culturally and Linguistically Diverse Backgrounds https://www.esafety.gov.au/research/women-from-diverse-backgrounds	2019
Reducing the impact of trolls online	Unpublished 2019
Parenting and pornography: findings from Australia, New Zealand and the United Kingdom www.esafety.gov.au/research/digital-parenting/pornography	2018
Understanding the digital behaviours of older Australians www.esafety.gov.au/research/digital-behaviours-older-australians	2018
State of Play – Youth, kids and digital dangers www.esafety.gov.au/research/youth-digital-dangers	2018
State of Play – Youth and online gaming in Australia www.esafety.gov.au/research/youth-online-gaming-state-of-play	2018
Young people and sexting—attitudes and behaviours: research findings from the United Kingdom, New Zealand and Australia www.esafety.gov.au/research/young-people-sexting	2017
IBA: national survey summary report www.esafety.gov.au/research/image-based-abuse	2017
IBA: qualitative summary report www.esafety.gov.au/research/image-based-abuse	2017
Lost Summer evaluation	Unpublished 2017
Cyberbullying complaints literature review	Unpublished 2017
Digital participation survey project which developed infographics and videos www.esafety.gov.au/research/digital-participation	2016
Infographic series on social cohesion www.esafety.gov.au/research/young-people-social-cohesion	2016
Infographic on Social media use and cyberbullying www.esafety.gov.au/research/social-media-cyberbullying	2015