THE UNIVERSITY OF WESTERN AUSTRALIA

TECH & POLICY LAB

SUBMISSION TO THE DEPARTMENT OF INFRASTRUCTURE, TRANSPORT, REGIONAL DEVELOPMENT, COMMUNICATIONS AND THE ARTS

STATUTORY REVIEW OF THE ONLINE SAFETY ACT

This submission responds to the Australian Government's Statutory Review of the Online Safety Act 2021 – Issues Paper, dated April 2024. The Issues Paper seeks feedback on the scope, operation, and effectiveness of the Online Safety Act 2021 (**the Act**). We appreciate the opportunity to provide evidence and expertise on this topic. Below we provide recommendations in response to consultation questions 1, 2, 3, 6, 11-12, 14, 16, 17, 21, and 30 in the Issues Paper.

RECOMMENDATIONS

1. (a) Establish a clear, adequately scoped problem statement for each of the heterogenous safety challenges the Act seeks to address. (b) Consult frontline experts to ensure that the Act is supporting **real** child safety, rather than chasing the assorted artefacts that fall under the umbrella of 'child sexual exploitation and abuse material' (CSAM).
2. Review and revise the eight industry sections to incorporate platforms that specifically cater to underage users (e.g., ed-tech platforms); regulate predatory features that emerge from advertising dependencies and the surveillance business model; and establish safeguards against opportunistic self-labelling.
3. Ensure the Act is equipped to assess predatory business practices, including product features such as: random match; social mapping; algorithmic recommender systems; algorithmic feed; and surfacing ads in private communications services.
4. (a) Work with independent research leaders to undertake landscape studies and literature reviews to build an evidence-based understanding of different human rights considerations and the complex problem space surrounding online safety. (b) Partner with, and consult, *direct support organisations* (vs. pure advocacy groups), who have meaningful holistic insight and experience into safety and resilience.
5. (a) Require safety-by-design principles and practices at the design and development stages, rather than post-implementation. (b) Place the responsibility of clearly communicating online safety norms, processes, and resources to end-users on tech companies. (c) Utilise risk assessment, independent audits, and other tools for ensuring ongoing attention to product safety and responsible design and deployment from providers.
6. (a) Avoid disproven technological solutions for age assurance. (b) Increase the availability of robust and accessible sex and intimacy education.

7. (a) Integrate resilience-building as an object for the Act. (b) Require providers to provide resilience-building skills into robust user control features

WHO WE ARE – UWA TECH & POLICY LAB

The University of Western Australia (UWA) Tech & Policy Lab is an interdisciplinary research centre focused on civic accountability in the tech ecosystem. Based at UWA Law School under the leadership of Directors Assoc. Professor Julia Powles and Professor Jacqueline Alderson, the Lab has expertise in technology law and governance, biomechanics and bioengineering, data analytics and machine learning, and augmented/virtual/extended reality technologies. This submission was co-led by Assoc. Professor Powles and the Lab's incoming Director of the Child Online Safety and Privacy Research Program (COSPR), Dr. Kate Sim. Dr. Sim brings expertise from academia, community organising, and industry on child sexual abuse and exploitation, tech-facilitated sexual violence, and tech policy. Most recently, she worked at Google where she oversaw product policy on grooming, sextortion, and non-consensual intimate imagery (NCII).

RESPONSE TO Q1: OBJECTS

The Act's objects – to "improve online safety for Australians" and "promote online safety for Australians" – fail to set a meaningful scope for the Act, and should be revised.

The Online Safety Act takes on a gargantuan task: it must bring seriousness to issues that have been dismissed (e.g., image-based abuse); it must balance an ecosystem-level approach with concrete remedies for harmed individuals; and it must hold accountable tech giants that have been highly effective in avoiding responsibility of any kind.[1] **The Act suffers from three major policy failings in seeking to meet these important challenges: (1) it collapses meaningful differences between various tech-facilitated harms; (2) it homogenises all end-users; and (3) it fixates on discrete pieces of problematic content, diluting the "systems-focused approach" it claims to adopt. In our submission, a key element underpinning these failures is the uninterrogated, animating spectre of "child sexual exploitation and abuse material", or CSAM.**

Chasing CSAM

The Issues Paper makes clear that child sexual exploitation and abuse material (**CSAM**) is a leading example of online harm that shapes how the Act formulates its problem statement and develops interventions, particularly through the Online Content Scheme, Industry Codes and Standards, and Basic Online Safety Expectations. **The resultant over-indexing of CSAM is not supported by evidence on harm prevention, nor is it conducive to the Act's mission of preventing and responding to tech-facilitated abuses like cyberbullying, cyber-abuse, and non-consensual sharing of intimate images.** For example, Part 3 of the Act begins by defining and describing the size and scope of various tech-facilitated abuses, but goes on to focus on the Online Content Scheme, which involves detecting, reporting, and removing Class 1 (including

[1] Powles, J. "The Corporate Culpability of Big Tech." In Bant, E. (ed.), The Culpable Corporate Mind (Hart Publishing, Oxford: 2023), 97-115.

CSAM) and 2 material, as a response to those problems.[2] Many tech companies already employ machine learning systems to detect, report, and remove CSAM, including hash matching technologies for known CSAM,[3] and classifiers to identify suspected unknown, or novel, CSAM.

Contrary to popular belief, possessing CSAM does not necessarily indicate a threat to safety or that an act of violence has taken place. This is in part due to the definition of CSAM, which extends well beyond abuse. Originating from US federal law (18 U.S. Code § 2256), the definition of CSAM refers to any sexual depiction of apparent minors. As the eSafety Commissioner notes, this is 'a broad category of content that encompasses material that sexualises and is exploitative to the child'.[4] In operation, this broad definition captures a vast range of imagery: from a parent taking a picture of a naked baby; to adolescents consensually sharing intimate imagery; to unthinkable abuse. Similarly, sharing CSAM also encompasses a range of behaviours. For instance, some people stumble upon an incendiary image and share it with friends out of anger or disbelief; some share it to "troll" or provoke others; and others view it thinking it is adult pornography.[5] Crucially, material falling within the definition of CSAM may also involve adolescents engaged in healthy, developmentally normative behaviour of sharing intimate imagery with trusted contacts. In recognition of the inherent difficulty of assuming intent to harm a child on the basis of CSAM possession alone, Germany is currently in the process of revising its minimum sentencing requirement for CSAM possession.

Experts in child sexuality development and adolescents' media use demonstrate that **generating and sharing intimate imagery with consent is a developmentally normative behavior among teenagers**. This is not an outlier phenomenon: young people are increasingly exploring physical and emotional intimacy with others through communication technologies, because that is how people communicate with each other today. A blunt approach that requires platforms to universally report and remove suspected novel CSAM – including CSAM that represents developmentally normative behaviours – perpetuates harmful social attitudes that equate sexuality with promiscuity, police young people's bodily autonomy, and erode privacy. These are not abstract possibilities. Shaming and policing how young people discover and assert their bodily and sexual autonomy can significantly endanger young people, especially those in the LGBTQI+ community; those living with abusive and controlling guardians; or those residing in repressive regimes where sexuality is stigmatised and even punishable by death.[6]

Some stakeholders in the child safety space advocate for technological tools that purport to be able to distinguish consensual imagery from others.[7] **Privacy considerations aside, there is no**

---

[2] Issues Paper, p.19-23.

[3] Farid, H. "An Overview of Perceptual Hashing." Journal of Online Trust and Safety 1, no. 1 (October 28, 2021). https://doi.org/10.54501/jots.v1i1.24.

[4] eSafety Commissioner. "Online Content Scheme: Regulatory Guidance." (2023) https://www.esafety.gov.au/sites/default/files/2023-12/Online-Content-Scheme-Regulatory-Guidance-Updated-December-2023.pdf.

[5] The industry and law enforcement parlance for this type of suspected CSAM report is "Potential Memes." To read more on this, see "Section 8 Discussion and recommendation" here: Grossman, S., Pfefferkorn, R., Thiel, D., Shah, S., DiResta, R., Perrino, J., Cryst, E., and Stamos, A. (2024). The Strengths and Weaknesses of the Online Child Safety Ecosystem. Stanford Digital Repository. Available at https://purl.stanford.edu/pr592kc5483.

[6] See End Violence's https://www.end-violence.org/disrupting-harm#findings.

[7] Some advocacy groups are proposing "self-hashing" as a technique for early detection of non-consensual image distribution. Self-hashing tools invite individuals who have taken their intimate imagery to share their imagery with advocacy groups or law enforcement so they can add their imagery to the database of CSAM hashes. Given the novelty of self-hashing tools, it remains to be seen their uptake, efficacy, and impact on deterrence. It should also be noted that

such technology.[8] It is not technically feasible to distinguish consensual from non-consensual imagery, because images do not have intrinsically consensual qualities. Consent, or the breach of it, is relational. Here, we see a dangerous strain of techno-solutionism that ironically mirrors much of the rhetoric and presumption of the large tech firms that legislation like the Online Safety Act are meant to challenge.

We have described the definitional and operational challenges of identifying CSAM. In our submission, the Act's over-indexing of the problem of CSAM prevents it from engaging thoroughly with the particular challenges of different forms of tech-facilitated harm.

RECOMMENDATION 1

(a) Establish a clear, adequately scoped problem statement for each of the heterogenous safety challenges the Act seeks to address.
(b) Consult frontline experts to ensure that the Act is supporting **real** child safety, rather than chasing the assorted artefacts that fall under the umbrella of 'CSAM'.

RESPONSE TO Q2: SECTIONS OF THE ONLINE INDUSTRY

The Act's current focus on eight specified industry sections has two distinct disadvantages: (1) it enables tech companies to self-label, thereby evading the intent of the Act; and (2) it collapses meaningful differentiation of culpability between companies.

Platform companies are not a monolith, and everything that touches a computer is not a 'platform'. Many 'platforms' contain multiple features and engage users across age bands. This presents a thorny regulatory problem where policies and practices are difficult to enforce. This is true for tech-facilitated abuses as well, where certain features and certain companies bear greater responsibility than others. The Act collapses this important distinction, because its unit of enforcement focuses on platform types, rather than specific features. This creates an odd situation where Relevant Electronic Service (RES) providers like ChatRoulette and Designated Internet Services (DIS) providers like 4Chan, both of which are renowned as agents of tech-enabled abuse, are held to the same standard as Signal and Quora, respectively.

One consequence of the divisions used in the Act is that **key sectors in the online industry that have demonstrated flagrant disregard for children's data privacy and online safety are able to operate without effective scrutiny of their harmful features and policies**. Consider the "ed-tech" sector. From Google Classroom to student record management systems, ed-tech companies have a business-to-business-to-client model where they are contracted by educational institutions (e.g., schools, libraries, universities) to provide enterprise platforms for minors. This business model creates an incentive system where the provider prioritises the demands of the contractor over the end-users' needs. Tech companies that provide services intended for use in educational

---

the user interface of such tools do echo body- and sexuality-shaming rhetoric. They also pose interesting challenges to upholding and implementing children's data and privacy rights. For an example of a self-hashing tool, see https://takeitdown.ncmec.org/.

[8] For example, WeProtect Global Alliance alleged in their 2023 Annual Report recommended age assurance as a policy agenda, but did not consult technical experts on the tool's feasibility, accuracy, and integrity. See the report here: https://www.weprotect.org/global-threat-assessment-23/#full-report.

settings should be held to the highest safety standards with respect to privacy, data minimisation, transparency, and usability by default.[9]

Another stark oversight in the Act is the absence of any focus on the ad ecosystem – the bread and butter of tech giants, and what fuels their predatory business models. We find this omission confusing, given that the surveillance business model is pivotal to many of the issues the Act purports to concern itself with. The incentive to keep people seeing and clicking ads is part of the structure that incentivises algorithmically amplifying harmful or deceptive content. At the same time, the fact that advertisers are the real customers of these platforms leads to a disincentive against thoroughly policing these ads. There is a clear opportunity here for the Commissioner to regulate predatory, deceitful, and manipulative advertising practices, especially as they target young people.[10] In fact, media and child development experts agree that the negative developmental impacts of ads outweigh their positive ones.[11]

Finally, the Act does not address how segments of the tech industry use self-labelling as a tactic for avoiding accountability. A cursory look at the Apple app store demonstrates this: dating apps are categorised as either "Lifestyle" or "Social Networking." These are legacy categories that Apple has created, and app developers self-label in a way that is advantageous for them. The eight sections proposed by the Act do not fundamentally address this problem of providers' self-labelling practices, and how self-labelling is a widely-used technique for avoiding responsibility.

RECOMMENDATION 2

Review and revise the eight industry sections to incorporate platforms that specifically cater to underage users (e.g., ed-tech platforms); regulate predatory features that emerge from advertising dependencies and the surveillance business model; and establish safeguards against opportunistic self-labelling.

RESPONSE TO Q3: THINGS THAT SHOULD BE REGULATED

Following from the response to Q2, **the Act fails to address the predatory business models – enacted through certain product features and incentive structures – that lie at the heart of tech giants' disregard for human rights, responsibility to the public, and collective safety.**

Scholars of media and communication have repeatedly demonstrated how platform companies implement predatory features that manipulate user interaction and extract personal data in order to boost their engagement, privileging shareholder interest over public interest.[12] This predatory business model impacts children in particular, and this is where regulators like the eSafety

[9] On predatory business and design practices in the ed-tech sector, see https://internetsafetylabs.org/wp-content/uploads/2022/12/2022-k12-edtech-safety-benchmark-national-findings-part-1.pdf; https://www.taylorfrancis.com/books/mono/10.4324/9781003344308/public-education-digital-age-morgan-anderson.
[10] Marwick, Alice, Jacob Smith, Robyn Caplan, and Meher Wadhawan. "Child Online Safety Legislation: A Primer." Bulletin of Technology and Public Life: University of North Carolina at Chapel Hill., 2024. 10.21428/bfcb0bff.de78f444.
[11] Wilcox, Brian L. "Report of the APA Task Force on Advertising and Children: (539692009-001)," 2004. https://doi.org/10.1037/e539692009-001.
[12] See Marwick, Alice, Jacob Smith, Robyn Caplan, and Meher Wadhawan. "Child Online Safety Legislation: A Primer." Bulletin of Technology and Public Life: University of North Carolina at Chapel Hill., 2024. 10.21428/bfcb0bff.de78f444.

Commissioner can enact effective interventions to genuinely safeguard both minor and adult users from manipulative business practices.

Product features are the manifestations of how tech providers enact business decisions, which can be harmful or predatory. Consider, for example, the business decisions underlying "meet-new-friends" features that randomly connect end-users.[13] Providers like Kik, ChatRoulette, BlindMeet, and now-defunct Omegle connect end-users randomly, with few safeguards in place. Such apps make money through advertising, which is built on the surveillance business model of extracting end-users' data and leveraging it for advertising.[14] Wanting to form new connections is a deeply human desire. Meet-new-friends features instrumentalise this desire as a pretext for collecting and leveraging personal information of its end-users, many of whom are adolescents. Companies whose products are built around these features often intentionally target young users,[15] because advertisers have a commercial interest in collecting (generally indirectly, through proxies and classifications) young people's attributes, interests, and other sensitive information. Such features also place end-users at heightened risk, because the provider does not provide adequate user safety options, like blocking, muting, and filtering contacts.

RECOMMENDATION 3

Ensure the Act is equipped to assess predatory business practices, including product features such as: random match; social mapping; algorithmic recommender systems; algorithmic feed; and surfacing ads in private communications services.

RESPONSE TO Q3: THINGS THAT SHOULD NOT BE REGULATED

There is a serious absence of scientific evidence underpinning the Act, presenting an important opportunity for rectification during the present review. We note that the citations and data points referenced throughout the Act and its ancillary documents rarely include evidence from academic research. In fact, of the 129 footnotes in the Issues Paper, peer-reviewed scientific research is referenced less than five times. Similarly, the citations referenced in recently released Industry Standards reference are from reports and blogposts by advocacy groups, law enforcement, and government agencies, including the eSafety Commissioner's office itself.

This is a deeply concerning lack of rigour for a topic of high complexity, nuance, and stakes. **Making big regulatory moves – as this Act does, and which the review proposes to extend – without scientific backing, in a context of heightened emotions, is reckless and irresponsible.** There are real online safety risks that harm young people's well-being, yet these are being discounted in favour of political attention-grabs that defy scientific evidence. Consider, for example, that the scientific community, particularly from the field of Media and Communications

[13] For why "meet-new-friends" apps warrant heightened regulation, see Gillespie, Tarleton, Patricia Aufderheide, Elinor Carmi, Ysabel Gerrard, Robert Gorwa, Ariadna Matamoros-Fernández, Sarah T. Roberts, Aram Sinnreich, and Sarah Myers West. "Expanding the Debate about Content Moderation: Scholarly Research Agendas for the Coming Policy Debates." Internet Policy Review 9, no. 4 (October 21, 2020). https://doi.org/10.14763/2020.4.1512.

[14] West, Sarah Myers. "Data capitalism: Redefining the logics of surveillance and privacy." Business & society 58, no. 1 (2019): 20-41.

[15] The predatory business practice of targeting young people as objects of data harvesting is an area warranting further study.

Studies, has been vocal and clear in opposing regulatory efforts to ban or age-gate young people's access to certain websites.[16]

It should be noted that there is a **substantial lobbying push from tech companies like Thorn and Yoti that tout technological fixes** like age verification, client-side scanning, and grooming detection in the name of child safety. These are essentially companies that sell technologies for profit. In other words, they have a vested interest in propagating the magical thinking that there is an easy, technological solution to complex social problems. On client-side scanning, the expert community from cryptography, mathematics, and digital security have repeatedly demonstrated that any efforts to "scan" erodes the integrity of encryption.

In addition to tech companies with vested interest in selling technological solutions, **online safety policy relies heavily on reports from advocacy groups**, such as WeProtect Global Alliance and the Canadian Center for Child Protections, alongside technology companies, like Thorn and Yoti. Advocacy groups' reports offer a valuable insight into how and why survivors experience tech-facilitated CSAM, but it is essential to note that they ***do not*** **provide generalisable evidence on which expansive laws and policies can responsibly be based.**

Frequently referenced in the eSafety Commissioner's Industry Standards and corollary documents are reports about offender behaviour from advocacy groups. These offer useful insights into the lifecycle of tech-facilitated CSAM but, as above, they do not provide an empirical basis from which to draw meaningful conclusions about how most people use messaging apps ("Relevant Electronic Services") or websites ("Designated Internet Services"). For example, section 1.6.1 of the Commissioner's Impact Analysis cites a 2024 report by Suojellaan Lapsia, a Finnish advocacy group, based on a voluntary survey of 30,000 CSAM seekers. A bar graph (see p.34 of Impact Analysis) in the Analysis lists messaging platforms like Telegram and WhatsApp as messaging platforms used by CSAM seekers. **A blanket targeting of digital services based on such partial reports would be the equivalent of banning nylon rope because some murderers have used nylon rope to hold their victims.** This is broad, scientifically baseless, technological determinism. Because CSAM offenders use Telegram and WhatsApp to seek CSAM, the argument goes, messaging platforms must be *responsible* for the distribution of CSAM. This false causation neglects the many common and reasonable uses of messaging platforms and exaggerates their culpability without marshalling rigorous evidence. By proffering such blunt, unevidenced approaches as solutions for the complex social issue of abuse, such efforts can redirect resources and attention away from vital and proven interventions.

Finally, **a striking feature of the advocacy groups so often invoked in the online safety policy context is that *none* offer direct support to children and adult survivors of CSAM, tech-facilitated or otherwise,[17] and therefore they have only limited, point-in-time experience with**

[16] "American Psychological Association Health Advisory on Social Media Use in Adolescence," May 2023. https://www.apa.org/topics/social-media-internet/health-advisory-adolescent-social-media-use.pdf; Marwick, Alice, Jacob Smith, Robyn Caplan, and Meher Wadhawan. "Child Online Safety Legislation: A Primer." Bulletin of Technology and Public Life: University of North Carolina at Chapel Hill., 2024. 10.21428/bfcb0bff.de78f444.

[17] Scholars examining the negative impacts of anti-trafficking efforts in the context of development and against communities of color have repeatedly argued that it is important to distinguish between service providing groups and purely advocacy groups. For further readings, see Ditmore, Melissa. Unbroken Chains: The Hidden Role of Human Trafficking in the American Economy. Beacon Press, 2023; Bernstein, Elizabeth. "Militarized Humanitarianism Meets Carceral Feminism: The Politics of Sex, Rights, and Freedom in Contemporary Antitrafficking Campaigns." Signs: Journal of Women in Culture and Society 36, no. 1 (September 2010): 45–71. https://doi.org/10.1086/652918; Kim, Mimi E. "The

the full scope of these issues. This is an important gap, because direct service providers who regularly interact with vulnerable youth populations necessarily have a holistic picture of the lifecycle of CSAM, including if, when, and how communication technologies play an outsized role.

RECOMMENDATION 4

(a) Work with independent research leaders to undertake landscape studies and literature reviews to build an evidence-based understanding of different human rights considerations and the complex problem space surrounding online safety.
(b) Partner with, and consult, direct support organisations (vs. pure advocacy groups), who have meaningful holistic insight and experience into safety and resilience.

RESPONSE TO Q6: TERMS OF USE

Service providers' terms of use should be a very minimal management tool. As any person who has used online services knows, terms of use provide a perfunctory wall of text, written in incomprehensible legalese, that tech companies employ for compliance purposes. Most end-users do not read terms of use, nor should they have to. There are two aspects to consider here in how providers enact online safety measures: (1) how they actually develop and implement online safety measures; (2) how they communicate those decisions as policies, practices, and norms to end-users.

First, there is little to no consideration of how providers design and deploy features and products as potential sites of regulatory intervention. The actual work of enacting online safety measures happens through product policy and enforcement. Currently, there exists very little guidance on how the Act will enforce responsible design and deployment of online safety policies and practices from providers.

One potentially impactful regulatory mechanism is internal risk assessment. Platform companies rarely put the time, effort, and energy needed to identify and mitigate anticipated harms of their new features and products. The Standards' risk assessment requirement is an important step in the right direction, but its current iteration is vague and tame. This is one of the areas where eSafety Commissioner can be a leader in innovating risk assessment techniques that will actually curb platform companies' predatory business practices while providing a meaningful vehicle for transparency.

Second, the onus of clearly communicating their online safety policies, processes, and resources to end-users should be placed on tech companies. Users refer to public-facing accessory documents (i.e., help center, community guidelines, FAQ) to access information about norms and guardrails about a digital service. Yet, as anyone who has tried to read company policy report an incident knows, these documents are difficult to find and follow. When they are discovered, they often contain inaccessible legalese, inconsistent updates, and broken hyperlinks.

Carceral Creep: Gender-Based Violence, Race, and the Expansion of the Punitive State, 1973–1983." Social Problems 67, no. 2 (May 1, 2020): 251–69. https://doi.org/10.1093/socpro/spz013.

They reflect the company's operational and compliance concerns, over end-user needs. There is a missed opportunity here for companies to utilise their immense resources and tools to make safety information easily discoverable, simplify the language, humanise the process, and close the loop where appropriate, and to do so with minor-friendly language and trauma-informed approaches to information delivery.

RECOMMENDATION 5

(a) Require safety-by-design principles and practices at the design and development stages, rather than post-implementation.
(b) Place the responsibility of clearly communicating online safety norms, processes, and resources to end-users on tech companies.
(c) Utilise risk assessment, independent audits, and other tools for ensuring ongoing attention to product safety and responsible design and deployment from providers.

RESPONSE TO Q11-12: ACCESS TO VIOLENT PORNOGRAPHY/ AGE INAPPROPRIATE CONTENT (INCL. AGE ASSURANCE)

Both questions 11 and 12 share two assumptions that are crucial to interrogate. First is the assumption that pornography is inherently harmful. Much of the Act's Online Content Scheme is designed to restrict people, particularly adolescents, from accessing pornography. The Issue Paper uses terms like "violent pornography" or "age inappropriate content," but they are ill-defined (undefined, even) and used interchangeably to refer to pornography at large. In essence, the Act erroneously assumes that pornography is inherently harmful. The second assumption is the Commissioner's tacit endorsement of age assurance as the appropriate solution to this ill-defined problem.

Pornography is not inherently harmful. There is little scientific evidence that 21st century pornography is "on the whole more violent than that of two, three, or four decades ago."[18] There is also no scientific evidence to suggest that pornography causes unhealthy sexual behaviours in young people.[19] There is a dangerous "magic bullet" thinking here that believes individuals who encounter harmful or objectionable content will enact such harmful or objectionable acts.[20]

[18] See McKee, Alan, Kath Albury, and Catharine Lumby. The Porn Report. 1st edition. Carlton, Vic: MELBOURNE UNIVERSITY PUB, 2008.

[19] In fact, research increasingly shows that consensual exchange of intimate imagery with trusted recipients in adolescents may have positive benefits, like sexual and bodily autonomy, positive body image, comfort in building intimacy with others, sexual arousal, and social connection. See Maes, C., Trekels, J., Impett, E., & Vandenbosch, L. (2022). The Development of the Positive Sexuality in Adolescence Scale (PSAS). The Journal of Sex Research, 60(1), 45–61. https://doi.org/10.1080/00224499.2021.2011826; Hasinoff, Amy Adele. Sexting Panic: Rethinking Criminalization, Privacy, and Consent. University of Illinois Press, 2015. https://www.jstor.org/stable/10.5406/j.ctt13x1kx6.

[20] "Magic bullet" thinking, or "media effects" theory, is a pseudoscientific paradigm that persists to this day, in spite of repeated debunking by the scientific community. On how the magic bullet theory informs the debate about access to social media, see Vuorre, M., Johannes, N., & Przybylski, A. K. (2022, July 15). Three objections to a novel paradigm in social media effects research. https://doi.org/10.31234/osf.io/dpuya. On video games, see Magnusson, K., Johansson, F., & Przybylski, A. K. (2023, May 28). Harmful Compared to What? The Problem of Gaming and Ill-defined Causal Effects. https://doi.org/10.1111/add.16516. For an overview of the scientific debate around this theory, see Berger, A. A. (1995). Media. In Essentials of Mass Communication Theory (pp. 53-86). SAGE Publications, Inc., https://doi.org/10.4135/9781483345420.

Efforts to regulate access to pornography should first wrestle with the question of *why* young people encounter porn.[21] Some adolescents seek out pornography voluntarily for reasons ranging from curiosity and information gathering to sexual arousal.[22] Interview studies of adolescents suggest that young people want to be able to receive information and discuss sex and intimacy practices, but are rarely provided with the right spaces to engage in such conversation.[23] Sex education programs in schools are clinical and uninteresting;[24] parents are afraid to have an honest conversation or don't know how to;[25] and peers are equally confused about the world of sex and intimacy.[26] While pornography is not sex education, evidence shows that adolescents are likely seeking it out because it is an accessible way to gather information and resolve their curiosity. The policy response to this should focus on providing robust and engaging comprehensive sex education that helps young people understand healthy sexuality and intimacy, and reject coercive and unhealthy sexual advances, situations, and content, instead of outright banning access through a method that is prone to error (more on this below). Where adolescents are actively seeking out graphic and violent pornography, experts recommend prevention programs centered around identifying and mitigating unhealthy sexual behaviour.[27]

Some adolescents do encounter pornography accidentally in unexpected spaces. Once again, advertisers are a culpable actor for intentionally instrumentalising graphic content to grab viewers' attention for advertising purposes. See our response to Q2 on the advertising ecosystem. Beyond ads, when young people do come across pornographic content, their response ranges from annoyance and dismissal to shock and discomfort. Responsible support is not to fearmonger or shame, but rather "to help them be resilient to the fact that it does and will exist, and help them not to become anxious or warped because they see such content."[28] Age gating and other related strategies to ban access altogether is therefore a misguided effort that does not address the reality of how and why young people encounter unwanted pornographic content.

The Act erroneously assumes that age assurance is an inevitability, and that it is a surefire technological solution. Age assurance is technologically faulty, laden with gender and racial bias, and puts young people's (and everyone else's) sensitive data at risk. A recent study of facial

---

[21] Center for Democracy and Technology. "Unraveling the Complex World of Youth Risk Experiences - Insights & Policy Implications," May 28, 2024. https://cdt.org/insights/from-our-fellows-unraveling-the-complex-world-of-youth-risk-experiences-insights-policy-implications/.

[22] See Wisniewski, Pamela, Heng Xu, Mary Beth Rosson, Daniel F. Perkins, and John M. Carroll. "Dear Diary: Teens Reflect on Their Weekly Online Risk Experiences." In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 3919–30. San Jose California USA: ACM, 2016. https://doi.org/10.1145/2858036.2858317.

[23] Cooper, Spring Chenoa, Kateryn Ferreira, Raz G. Edwards, Julia Keegan, Nika Norvila, Larissa Lewis, Kath Albury, and S. Rachel Skinner. "A Qualitative Exploration of Young Australians' Lived Experiences of Social Media Use and Sexual Agency." Sexuality & Culture 28, no. 2 (April 1, 2024): 534–53. https://doi.org/10.1007/s12119-023-10131-w.

[24] Albury, Kath. "Porn and Sex Education, Porn as Sex Education." Porn Studies 1, no. 1–2 (January 2, 2014): 172–81. https://doi.org/10.1080/23268743.2013.863654; Byron P, McKee A, Watson A, et al. (2021) Reading for realness: Porn literacies, digital Media, and young people. Sexuality and Culture 25(3): 786–805.

[25] Wisniewski, Pamela, Heng Xu, Mary Beth Rosson, and John M. Carroll. "Parents Just Don't Understand: Why Teens Don't Talk to Parents about Their Online Risk Experiences." In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, 523–40. CSCW '17. New York, NY, USA: Association for Computing Machinery, 2017. https://doi.org/10.1145/2998181.2998236.

[26] Maes, Chelly, Paul J. Wright, and Laura Vandenbosch. "Adolescents' Preferences for Mainstream and Paraphilic Pornography and Sexual Health Components: Attention to Within-And Between-Person Dynamics Over Time." Health Communication 0, no. 0 (2024): 1–13. https://doi.org/10.1080/10410236.2024.2335691.

[27] See Letourneau, Elizabeth J., William W. Eaton, Judith Bass, Frederick S. Berlin, and Stephen G. Moore. "The Need for a Comprehensive Public Health Approach to Preventing Child Sexual Abuse." Public Health Reports 129, no. 3 (May 1, 2014): 222–28. https://doi.org/10.1177/003335491412900303.

[28] Wisniewski et al., 3926.

recognition systems underpinning age assurance systems found that it performs more accurately with lighter skin tones and female faces. The authors conclude, "If our commitment is to supporting healthy sexual development in a digital context, focusing on automated processes is a distraction from the positive, practical, evidence-based steps that would make a difference."[29] This step, they advise, is what we already know: robust and accessible sex and intimacy education available to all.

RECOMMENDATION 6

(a) Avoid disproven technological solutions for age assurance.
(b) Fund and support communities providing robust and accessible sex and intimacy education.

RESPONSE TO Q14: REPORTING BY 'BYSTANDERS'

Bystanders are already able to report objectionable content. We want to caution that this presents social and operational challenges. At the societal level, this encourages bystanders to report what they perceive to be objectionable – a tactic often used to silence and police speech from marginalised communities.[30] At the operational level, encouraging bystanders to report "harmful" content with poorly defined parameters will likely raise serious capacity questions for the Commissioner's office. Even tech companies with expansive human- and machine- driven content moderation ecosystems in place struggle to triage and process "noise" in their user reports. If the Act expands reporting abilities for bystanders, this should be accompanied by clear parameters for what is reportable, who is responsible for receiving and processing reports, and possible outcomes.

RESPONSE TO Q16: RESEARCH, EDUCATION, AWARENESS

This Act focuses on detecting and removing harmful content, and keeping young people away from potentially harmful content. In doing so, there is very little consideration for fostering safety through resilience building. As child development experts insist, risk is not inherently harmful and plays an important role in fostering a sense of agency, the ability to hold and navigate nuance, empathy for others, and other desirable skills.

RECOMMENDATION 7

(a) Integrate resilience-building as an object for the Act
(b) Require providers to provide resilience-building skills into robust user control features.

[29] Stardust et al. p.14.
[30] See A. Marwick. "Morally Motivated Networked Harassment as Normative Reinforcement." Social Media + Society, n.d. https://doi.org/10.1177/20563051211021378; Matias, J. Nathan, Amy Johnson, Whitney Erin Boesel, Brian Keegan, Jaclyn Friedman, and Charlie DeTar. "Reporting, Reviewing, and Responding to Harassment on Twitter." arXiv:1505.03359 [Cs], May 13, 2015. http://arxiv.org/abs/1505.03359.

RESPONSE TO Q17: INVESTIGATION, INFORMATION, ENFORCEMENT

The Act's naming convention for harmful material is likely to present an enforcement challenge in one key respect. The recent Industry Standards refer to illegal and harmful content as "Class 1A and 1B" material. This naming convention of combining numbers with alphabet letters (i.e., "1A") is the convention used by legal authorities worldwide to classify the egregiousness of CSAM (1A/B, 2A/B, 3C). Australia's adoption of its own naming convention to refer to illegal material at large may pose operational and enforcement confusion and challenges in implementing and enforcing the Standards.

RESPONSE TO Q21: INTERNATIONAL APPROACHES

We want to take this question as an opportunity to reiterate serious concerns within international scientific communities about regulators' reliance on technological fixes, including client-side scanning, age verification, and grooming and CSAM classifiers, among others.

In spite of technology experts' repeated debunking of pseudoscientific claims about device-level scanning[31] and age verification,[32] the cottage industry of tech solutions for online safety continues to grow. Such technological tools are touted by tech companies and lobby groups with vested interest in profiting from selling their products. We are heartened by the eSafety Commissioner's recognition that it is not technically feasible to interfere with e2e encryption in its recent Industry Standards, echoing the UK Online Safety Act, in response to the technical expert consensus. However, we anticipate that the cottage industry will continue to grow and tout technological solutions. There is a lot of money in this space – a lucrative opportunity for bad faith actors to take advantage of people's lack of technical understanding and leverage emotional responses from regulators and members of the public.

Continuing to litigate the technical feasibility of spruiked technological "fixes" is not only misguided, but irresponsible. The statutory review of the Act presents an opportunity to affirm the importance of safety **and** privacy for young people. There continues to be a bifurcation between children's safety and adults' privacy, as if children do not have a right to privacy and adults do not need safeguarding. This dichotomy prevents regulators from engaging with online safety challenges with rigour, principle, and care. In the US, investigators used private Facebook messages to charge Jessica Burgess for supporting her teenage daughter's abortion. Jessica is currently living in prison, after being convicted based largely on these messages. Advocates in the Congo who are campaigning to end the use of child soldiers use encryption tools to document human rights abuses, and communicate with human rights advocates in other jurisdictions. LGBTQI+ people around the world, especially in jurisdictions where same-sex relationships are punishable by death, rely on encrypted communication tools to form communities and gather life-saving information.

[31] H. Abelson, R. Anderson, S. M. Bellovin, J. Benaloh, M. Blaze, J. Callas, W. Diffie et al. "Bugs in our pockets: The risks of client-side scanning." Journal of Cybersecurity 10, no. 1 (2024): tyad020.

[32] Z. Stardust, A. Obeid, A.McKee, and D. Angus. "Mandatory Age Verification for Pornography Access: Why It Can't and Won't 'Save the Children.'" Big Data & Society 11, no. 2 (June 2024): 20539517241252129. https://doi.org/10.1177/20539517241252129.

RESPONSE TO Q30: IS THE ACT ACHIEVING ITS OBJECT?

Much of our response to this question has already been addressed in the submission above.

In closing, we reiterate the importance of taking a balanced approach that not only safeguards young people, but also fosters resilience and autonomy. This should be baked into the Act's approach, so that safety measures promote and incentivise healthy behaviours, in addition to mitigating harmful behaviors.