

Harm Blindness in Digital Child Protection Policy

A Systematic Stakeholder, Sentinel and Legal Analysis of Australia's Under-16 Social Media Ban

The ban does not treat the poisoning. It removes the population whose symptoms were the early warning.

10 DECEMBER 2025

The first country to shut every child under sixteen out of social media

Australia's amendment to its online safety law exposes platforms to civil penalties reaching forty-nine and a half million Australian dollars for systematic failures to stop underage accounts being created.

It has already reshaped the argument in the United Kingdom, the European Union, Malaysia and South Africa, where the responsible minister has said comparable measures are under active consideration.

This paper argues the ban is not merely an ethically expensive policy. It is a legally fragile one, and it fails on its own terms before any question of rights arises.

THREE INSTRUMENTS, ONE ASSESSMENT

How the argument is built

1 The Harm Blindness Framework

A method for finding preventable damage at four decision points: having the idea, designing it, testing it and launching it. It names harm blindness as a systematic failure to see who gets hurt, driven by looking only at the obvious stakeholder, by handing the ethics to somebody else, by treating precedent as permission, and by moving too fast.

2 The Paediatric Sentinel Effect

Borrowed from toxicology and environmental health. Children function as a sentinel population for algorithmic systems, meeting harms earlier and more visibly than adults, for three reasons: they are exposed by institutions with little room to refuse, their development makes them more open to manipulation, and the most legally protected population generates the most behavioural data.

3 A proportionality lens

Article 3 of the children's convention and General comment No. 25 require any measure restricting a child's digital access to be necessary, proportionate and the least restrictive option, and not to end in children being excluded from the digital environment. South African law makes a child's best interests paramount, applied to the actual child rather than as an abstraction. Together these turn a policy critique into a legal assessment.

WHERE THIS SITS IN THE LITERATURE

Two accounts of prejudice against children, and what each one lacks

One account treats hostility toward children as a prejudice comparable to racism or sexism, legitimating policies that damage children while claiming to protect them. The Australian ban fits that description exactly. Protective language covers the exclusion of queer young people, abused children and disabled users from communities they depend on, and nobody asked them.

The other account treats childhood as a critical theory on the model of feminism, in which children's experience reshapes ethics and rights rather than being fitted into arrangements built by adults.

This work sits between them without being captured by either. The Harm Blindness Framework supplies the diagnostic apparatus the first account never had. The Sentinel Effect supplies the predictive engine the second points toward.

Who Was Not Counted

nine groups beyond the intended beneficiary, two of them critical



Checkpoint by checkpoint, the same failure appears at every stage of the decision.

THE FIRST CHECKPOINT

Two groups lose the thing that was keeping them alive

Nine stakeholder groups sit outside the intended beneficiaries, and two are at critical severity.

Queer young people in households that do not support them lose communities that affirm who they are, lose their friends, and lose the crisis resources they reach for when they start thinking about ending their lives.

Children living with abuse lose the outside line, the channel of communication their household does not control, and their route to the adults who are legally required to report what they hear. The policy leaves them shut in with the people harming them.

THE SECOND CHECKPOINT

Four incentives that run against the policy's own purpose

1 Platforms come out ahead

They acquire biometric verification data on everybody, while the exploitative design that generated the harm is left entirely alone.

2 Teenagers move

Into less regulated spaces: chat services, messaging apps, and network workarounds that put them outside anybody's view.

3 Parents relax

A false sense of security sets in, and supervision falls off precisely where it was doing some good.

4 Governments get to announce something

Decisive action, claimed and reported, without anybody having to confront a platform business model.

Children as the Early Warning

what happens to them first is what happens to everybody later



This ban is a second-order harm. It removes the sentinel rather than treating the toxicity.

THREE CASES WHERE THE WARNING CAME FROM CHILDREN

The pattern is documented, and each time the adult version followed

Biometric surveillance installed in schools in Lockport ran for eight months before the first wrongful arrest of an adult from the same technology.

A family screening tool built for child welfare established the precedent for predicting outcomes from administrative data, and that precedent is now migrating into contexts involving adults.

A social platform's own internal research documented what its image service was doing to teenage girls well before adult platform addiction was clinically recognised at all.

WHAT THE BAN DOES INSIDE THAT MODEL

It removes the population whose experience was the warning

Read against the sentinel account, the Australian ban is a harm of the second order. It does not address the algorithmic toxicity that children were experiencing.

It eliminates the population whose experience of that toxicity was the signal that it was coming for everybody else.

This also advances children's rights scholarship past a welfare framing. Harm to children stops being only a harm to children and becomes a leading indicator of risk across the whole population.

The Legal Fragility

what the welfare critique misses



A statute that cannot be enforced against its primary duty-bearers is signalling rather than regulation.

FIRST DEFECT

The enforcement depends on collecting from companies with nothing here

Everything rests on the commissioner's ability to impose and then collect penalties against platforms domiciled abroad with no local legal presence. South Africa faces the identical problem.

In the age verification context, none of the available instruments closes the gap. A subpoena will not. An enforcement notice will not. A mutual legal assistance treaty will not, because all of them run into dual criminality requirements, into delay, and into limits on scope, inside a scheme that is civil or administrative rather than criminal.

A statute that cannot be enforced against the people it primarily binds is signalling rather than regulation.

SECOND DEFECT

The protective logic runs backwards

Australia's own age assurance technology trial confirmed that every available method carries measurable error, and that each requires either a government identity document or biometric data from every single user.

So children's data ends up processed more intensively than before, not less. That inverts the protective logic written into South Africa's data protection statute and into the Information Regulator's own 2021 guidance on processing children's information.

The paradox is unavoidable. There is no version of this scheme in which a child's information is handled more lightly than it was before the scheme existed.

THIRD AND FOURTH DEFECTS

A South African copy would have to survive limitation analysis, and would deepen existing crimes

Replicating this in South Africa engages the constitutional limitation provision, requiring a justifiable restriction of rights to expression, to association, to the child's best interests, and to equality.

It would also worsen harms already criminalised under the cybercrimes statute and the films and publications amendment, while cutting visibility to the very people legally obliged to report what they see.

The instrument engages the children's convention across five articles: non-discrimination, participation, expression, information access, privacy, and protection from violence. Two articles of the African Charter are directly engaged. General comment No. 25 is explicit that protection measures must not end in children being excluded from the digital environment, and that states should address the root causes of harm rather than restricting access. The ban fails that standard.

What Would Work Instead

and where each piece already attaches in existing South African law



Every alternative here hangs off a statute that is already in force.

FIVE INSTRUMENTS, ALL OF THEM ALREADY AVAILABLE

Alternatives that reach the design rather than the child

1 Algorithmic transparency

Attaches to section 71 of the data protection statute, read with section 5.

2 A ban on addictive design

Attaches to sections 48 and 49 of the consumer protection statute, and to the Information Regulator's power under section 35(2).

3 Age-appropriate design standards

Implemented through a section 60 code of conduct, on the model of the United Kingdom's children's code.

4 Platform liability

Extends sections 18F to 18H of the films and publications amendment.

5 Digital literacy obligations

Attach to the Children's Act and to the Schools Act.

SOURCES POPIA §§ 5, 35(2), 60, 71; Consumer Protection Act 68 of 2008 §§ 48–49; Films and Publications Amendment Act 11 of 2019 §§ 18F–18H; Children's Act 38 of 2005; South African Schools Act 84 of 1996.

REFERENCES

References

Travis Gilly & Rene Mattheus, *Harm Blindness in Digital Child Protection Policy: A Systematic Stakeholder, Sentinel and Legal Analysis of Australia's Under-16 Social Media Ban*, in *Children's Rights Under Pressure in a Digital World* (2026). The authors contributed equally to this work.

Commonwealth of Australia. (2025). *Age assurance technology trial: Final report*. Department of Infrastructure, Transport, Regional Development, Communications and the Arts.

Consumer Protection Act 68 of 2008 (South Africa).

Information Regulator (South Africa). (2021). *Guidance note on processing of personal information of children*.

South African Schools Act 84 of 1996.

UN Committee on the Rights of the Child. (2021). *General comment No. 25 (2021) on children's rights in relation to the digital environment* (CRC/C/GC/25).

Wall, J. (2008). Human rights in light of childhood. *International Journal of Children's Rights*, 16(4), 523–543. <https://doi.org/10.1163/157181808X312122>

Wall, J. (2010). *Ethics in light of childhood*. Georgetown University Press.

Young-Bruehl, E. (2012). *Childism: Confronting prejudice against children*. Yale University Press.