

Chapter 7 – Conclusions and Predictions for the Future of Online Safety Policy

Is it knowing when to tell someone if you're upset by something that happens online?

I am fond of using the above quote, from a 13 year old young woman during a discussion that took place in a school over ten years ago. It was during an invited session where I had been asked to do “a talk” to year 8 pupils about online safety. However, the class was small so I said I would just have a discussion, rather than giving them a talk. In the discussion one of the young people asked what I did for work, because they did not see how someone could have a job talking to young people in schools. I explained broadly that my work examined what young people did online, and what adults can do to keep them safe. I asked them what online safety meant to them. The quote above is what one of the group said. If young people have recognised this for well over ten years why do policy makers keep plough their furrow of platform liability?

As I said in chapter 1, during this time researching the online safety space, as new technologies come and go, as do policy makers and civil servants, one thing has remained constant. Young people want help, support and education; not heroes or saviours. And they want support from the adults around them, and they do not expect engineers in large companies far away from their lives¹.

Yet, as we have demonstrated in this exploration of policy, path dependency and isomorphism, as we have explored to try to understand why policy makers will only ever do more of the same, highlight how disconnected the policy space is to young people's needs. In 2008, the government accepted recommendations from the Byron Review (Byron 2008) about multi stakeholder approaches to tackling online safety, underpinned by education and stakeholders working to their strengths.

In 2012 a different government decided the best approach was to block content.

And as other nations adopted similar approaches, with equally hollow impact, isomorphism, it seemed, locked them on this path. Other nations were adopting a prohibitive model and trying to look tough in bring the tech sector in line, so anyone suggesting a more progressive approach might appear weak. And currently, there seems to be a legislative arms race to see who will “ban” smartphones for young people with most rapidity.

There's nothing wrong with the new laws per se. There is certainly some value in expecting platforms to provide evidence that they have thought about the potential harms that occur on their platforms and how they mitigate them. Which is what most already do to some extent. However, the political expectation that they can and should prevent harms from occurring, and the ideological obsession with holding platforms to account, with not end goal of what enough looks like, is doomed to fail.

¹ I caveat this, once again, by stating clearly that industry, as a stakeholder in child online safety, clearly has a role to play and they are responsible for things like content moderation, monitoring problematic activity and providing routes and tools for disclosure. However, that cannot prevent harm from occurring, and their role has to mesh with the practices of other stakeholders. Other stakeholder cannot simply sit back and shout “How did <company x> allow that to happen!?”.

As young people have told us, people cause upset online, and we can't expect private companies to single handedly police the behaviour of people globally.

In the reactive approach to dealing with media outrage and policy direction driven by a coercive element of isomorphism which, by its nature, driven by NGOs and media, will make use of edge cases to tackle the broader context of duty of care. And the path dependency manifest in the view that “well, this is the direction we’re going in, so we better keep going” is clear to see.

It seems, when exploring stakeholder views, that we have lost sight of why we are doing this. Rather than placing the child at the centre of any decision, by stakeholders at any level, it seems instead that if things feel too complex, or too hard, it's better instead to demand “someone” does something about this. For example, schools could, and should (speaking from a regulatory point of view) train their staff to understand the issues of online harms and how to mitigate. But with no specification of what that training looks like, and who might deliver it, coupled with the low risk of an inspector asking about it, many schools simply do not, so the professionals in their setting do what every other adult around the child does – fall back on knowledge informed by media and moral panics, and bring their own moral judgements to decisions affecting young people. And the same could be said for others in the exosystem, as highlighted in the case discussed in chapter 6. And the expectation to deliver education to young people, another statutory requirement, wanes considerably as young people move to secondary school, leaving them poorly informed about emerging online risks they are likely to experience through their teenaged years. Far better, it seems, to ask a police officer to come in instead and tell them if they send nudes they might not be able to go to Disneyland²

In conclusion, the ecosystem of child online safety remains fraught with challenges and paradoxes. Despite an expanding body of knowledge, heightened media attention, and robust policy interest, the approach to safeguarding children in the digital realm often appears misaligned with the nuanced needs and voices of young people. This book has delved into the systemic tensions and historical oversights that plague current policies, revealing a pattern of reactionary measures rather than proactive, evidence-based strategies.

The analysis underscores the critical need for a holistic, multi-stakeholder approach, as conceptualized through an adaptation of Bronfenbrenner’s ecosystem. Effective online safety policy should incorporate the voices of young people and consider the interconnected roles of family, education systems, social services, and broader societal influences. Yet, current policies, exemplified by the UK's Online Safety Act 2023 and similar international legislations, often default to simplistic solutions such as platform liability and technological interventions, which fail to address the root causes of online harm. And, when things go wrong, platforms are scapegoated as “they should have stopped it”, rather than accepting the complexities of a connected ecosystem and the needs and responsibilities of the stakeholders.

Historical reviews, such as the Byron Review, have highlighted the importance of comprehensive, collaborative strategies involving education, awareness, and industry cooperation. And the Byron Review was released in 2008 and the government of the day agreed to implement its recommendations. However, recent debates and legislative actions suggest a persistent tendency to favour regulatory and punitive measures over these more complex, but ultimately more effective, approaches.

² This particular anecdote was told to me by a professional who had observed such an assembly and then when they asked a year 8 pupil what they had learned from it, the pupil said “If I send a dick pic I won’t be able to go to Disneyland”.

This book calls for a paradigm shift in how policymakers and stakeholders view and address child online safety. It advocates for moving beyond a narrow focus on preventing harm through bans and restrictions, towards empowering young people with the knowledge, resilience, and support they need to navigate the digital world safely. Only by aligning policy with the authentic needs and voices of young people can we hope to create a safer, more inclusive online environment that truly protects and empowers the next generation.

The period under review witnessed significant legislative efforts aimed at enhancing online safety, most notably the UK Online Safety Act 2023. This legislation was presented as a landmark in the attempt to regulate digital spaces, imposing stringent requirements on online platforms to protect users, particularly children, from harmful content. The Act's introduction of a statutory duty of care, robust age verification measures, and enhanced content moderation mechanisms reflects a comprehensive approach to mitigating online risks. However, as explored throughout this book, the practicalities of implementing such measures are fraught with challenges, including the operational and financial burdens on platforms and the complexities of balancing safety with freedom of expression and privacy rights. While the statutory duty of care mandates that online platforms take proactive steps to prevent the dissemination of illegal content, such as child sexual exploitation, terrorism, and hate speech. Platforms are also required to conduct risk assessments and put measures in place to mitigate potential risks. However, the lack of precise definitions for what constitutes "effective" content moderation and the challenges of measuring compliance present significant obstacles for platforms.

Content moderation, while essential, is not a one-size-fits-all solution. The nature of harmful content varies widely, and algorithms designed to detect and remove such content often struggle with context and nuance. Furthermore, the implementation of robust age verification measures poses its own set of challenges. While age verification is crucial to prevent children from accessing harmful content, it raises concerns about privacy and data security. Collecting and storing age-related data introduces risks, such as data breaches and misuse of personal information. Balancing the need for age verification with the protection of user privacy requires careful consideration and the development of secure, privacy-preserving technologies.

Despite these legislative advancements, the effectiveness of these measures remains questionable. The persistence of moral panics and the tendency to prioritize reactive, prohibitionist policies over proactive, evidence-based strategies underscore the cyclical nature of online safety policy. This approach often results in measures that fail to address the root causes of online harms and overlook the nuanced needs and voices of young users. The comparison with the evolving drug policy landscape, which has increasingly adopted harm reduction strategies, suggests that a similar shift towards education, awareness, and support might yield more sustainable outcomes in the realm of online safety.

The ecological system model developed during this period, based on Bronfenbrenner's Ecology of Child Development, offers a framework for understanding the multifaceted responsibilities around online safety and the importance of the interactions between different actors in a child's environment—family, schools, public services, media, and policymakers—the model underscores the importance of a holistic approach to safeguarding. Each stakeholder plays a crucial role, and their collective efforts can create a supportive ecosystem that mitigates risks while empowering young people to navigate the digital world safely.

However, this analysis also highlights significant gaps in stakeholder engagement and collaboration. The focus often remains on external regulatory measures and platform liability, with insufficient

attention to the roles of families, schools, and local communities. Engaging these microsystems more effectively could lead to more nuanced and context-specific interventions that address the diverse experiences and needs of young people. Moreover, fostering a culture of digital literacy and resilience within these closer circles can complement broader regulatory efforts, ensuring that children are better equipped to handle online challenges.

Families play a critical role in shaping children's online behaviour and safety. Parents and guardians are often the first line of defence, providing guidance, setting boundaries, and monitoring their children's online activities. However, many parents feel ill-equipped to manage the complexities of the digital world, especially given the rapid pace of technological change. Providing parents with the necessary resources, education, and support to navigate these challenges is essential. This includes promoting open communication between parents and children about online risks and encouraging parents to take an active role in their children's digital lives.

Schools also have a significant part to play in promoting online safety. Educational institutions can integrate digital literacy into their curricula, teaching students about the potential risks of online interactions and how to navigate the internet safely and responsibly. Schools can also provide training for teachers to help them understand and address online safety issues, and to create a supportive environment where students feel comfortable discussing their online experiences and concerns.

Public services, including law enforcement and child protection agencies, are crucial in responding to serious online harm incidents. These agencies need to be equipped with the knowledge and tools to effectively address online risks and to support victims of online abuse. Collaboration between public services and other stakeholders, such as schools and tech companies, can enhance the overall effectiveness of online safety efforts.

Media narratives and moral panics have played a profound role in shaping public perception and policy responses to online safety. Sensationalist reporting on issues such as online grooming, cyberbullying, and exposure to harmful content often drives a reactive policy environment, where immediate, headline-grabbing measures are prioritized over thoughtful, long-term strategies. Through the early stages of this book we have drawn parallels with historical moral panics, such as the Satanic Panic of the 1980s, to illustrate how fear and misinformation can lead to disproportionate responses and unintended consequences, and that these moral panics rarely result in improved outcomes for those that are claimed we need to protect.

The media's portrayal of online safety issues tend to focus on extreme cases and worst-case scenarios, which can create a distorted view of the prevalence and nature of online harms. While these cases are undoubtedly serious and warrant attention, they do not represent the everyday experiences of most young people online. This skewed perception can lead to policies that are overly restrictive and fail to account for the positive aspects of digital engagement, such as social connection, learning, and creativity.

Addressing this dynamic requires a shift towards more responsible and balanced media coverage, as well as greater emphasis on public education and awareness. Providing clear, evidence-based information about online risks and the effectiveness of different safety measures can help counteract fear-driven narratives and support more rational policy-making. Additionally, involving young people in these discussions and recognizing their active role as digital citizens can help bridge the gap between adultist perspectives and the realities of youth digital experiences.

The overarching theme of this book, and indeed the majority of my work over a twenty-year period, is the need for a progressive, evidence-based approach to online safety that aligns with the lived experiences and needs of young people. Moving away from punitive, prohibitionist policies towards strategies that emphasize education, empowerment, and harm reduction can create a more sustainable framework for safeguarding children online. This approach recognizes the complexity of the digital landscape and the diverse ways in which young people engage with technology.

Key to this shift is the inclusion of youth voices in policy discussions and decision-making processes. Ensuring that young people have a say in the development and implementation of online safety measures can lead to more relevant and effective interventions. Additionally, fostering partnerships between different stakeholders, from policymakers and tech companies to educators and families, can enhance the collective capacity to address online harms in a holistic and coordinated manner.

Fostering partnerships between stakeholders is also crucial for creating a cohesive and comprehensive approach to online safety. Collaboration between policymakers, tech companies, educators, and families can facilitate the sharing of knowledge, resources, and best practices. These partnerships can also help to bridge gaps in understanding and build a shared commitment to protecting young people online.

There is a need for evidence driven, rather than ideologically driven, policy and stakeholders decision making, not “what I reckon” - a phrase I hear for too often from professionals claiming to adopt an evidence-based approach. I heard it when discussing the influence of violent video games on young people’s offline behaviour (Phippen and Street 2022). And I was recently told in a meeting by a senior practitioner that “just because there is no evidence doesn’t mean its not happening”. However, I would suggest clairvoyance is not a good foundation upon which to support young people’s needs around understanding risk online.

A Future?

Certainly, it is difficult to predict what the future for online safety policy and practice might be. However, reflecting upon the conclusions in the previous book on the safeguarding dystopia and reliance on technology to prevent online harms (Phippen 2017), the reality was perhaps starker even than anticipated. While there seem to be few trends for overreaching surveillance in the home, policy has shifted to platforms implementing the dystopia at a macro level, at the instruction of governments wishing to scapegoat them for the harms that occur online. And the legislation, which previously was considering a single use case (preventing youth access to pornography) has expanded, with the same prohibitive ideology, to any form of online harm.

If I consider what these six years might hold, it seems apparent that policy will remain driven by ideology rather than evidence. Or, where evidence is used to vindicate approach, it will ignore anything that disagrees with the approach.

The writing of this book concludes at a time, in the UK, where is Online Safety Act 2023’s implementation is still under consultation with the regulator. It is moving from a legislative goal to one that a regulator might implement in a practical manner that can be both adopted by those who will be regulated (predominantly large online platforms and search engines) and also enforced.

As already discussed, a potential fall out from stronger regulation is that there will be a lot more public information about what platforms do to tackle safety issues, borne out in their risk assessments they are expected to produce, and the transparency reports they will be mandated to provide. The value of these sorts of outputs for those developing education and interventions in

the microsystem has already been discussed, but it will also mean that it will be more difficult for some actors in the macrosystem to claim that platforms do nothing or do not care.

And potentially, given that the regulator is appointed to implement a political ideology on “bringing big tech billionaires to heel” or “prevent them from profiting from children’s misery”, and also that there is no clear definition of the end point of “doing more”, there is clearly potential that a regulatory who, as is the expected role of a regulator, works with industry to assure alignment with the statutory duties, will also attract the ire of those who wish to continue to scapegoat industry and distract for the undeniable complexities of this space. Could regulators become the next scapegoat due to their reluctance to impose “massive” fines on companies who are clearly “doing more”?

There is already rhetoric from certain media voices and political figures criticising the regulator for not going far enough and how if they don’t children will be harmed (see below). Highlighting, once again, that there is a certain motivation among some not to ensure better outcomes for young people, but to engage in high profile political and media battles that keep their own profiles high. When a policy with no end point, other than “stop anything bad happening online” look like becomes something that needs to be regulated, it is down the regulator to make sure what is demanded is possible, and the legislation in its current form makes this challenging.

Take, for example, the previously discussed age assurance/verification to prevent young people’s access to adult content. Should a platform conducting a risk assessment and put robust age assurance in place as a result then be “held to account” if a young person bypasses this verification with privacy enhancing tools or using someone else’s verification token? And if the regulator takes the, quite correct, view that the platform has demonstrated due diligence and, given the lack of perfection in age assurance, and the roles of other stakeholders in bypassing such approaches, therefore cannot be accountable should a determined young person achieve a bypass, might they equally become part of the more toxic elements of the ecosystem target for criticism for not caring about children? Might the locus of blame shift from the platform (“do more”) to the regulator (“you’re letting them get away with it!”)

Writing a book like this is hard because things are always moving. Will there ever be a case for writing a book that says “Online safety used to be hard but its great now”? There is no end point, and it would seem that policy makers and NGOs, given they have not defined what “doing enough” looks like, will never take the view that the policy has reached its end goal. Because there is no end goal other than stop online harms. Which, as young people surveyed and explored in chapter 5 very clearly tell us, are caused, predominantly, by people.

As the writing of this book was coming to an end, during the UK general election campaign 2024, the Molly Rose Foundation³, whose goal is stated to be suicide prevention, released a manifesto⁴ calling for the next government to bring in a new Online Safety Act 2023, because the current one is not strong enough:

Molly Rose Foundation is calling for the next Government to commit to five transformative policies that, taken together, would decisively protect young people from preventable harm and bring about a step change in their online safety and well-being.

We are calling on the next Government to:

³ <https://mollyrosefoundation.org/>

⁴ <http://mollyrosefoundation.org/wp-content/uploads/2024/05/MRF-Manifesto.pdf>

1 Introduce a new Online Safety Act 2023 to strengthen regulatory protections and focus the regime on achieving measurable harm reduction.

2 Take bold moves to ensure transparency and accountability from Big Tech, including a new duty of regulatory candour, and new measures to make companies report on user exposure to online harms in their corporate accounts.

3 Make the polluter pay, with a one-off harm reduction windfall tax that ensures tech companies no longer profit from harm while passing on the costs of their business model to children, families and society.

4 Commit to a new statutory Code for app stores and operating systems, ensuring that more of the tech ecosystem has legal responsibilities to protect children and support parents.

5 Invest in mental health support and prevention schemes to ensure children and young adults get the support they need to address the risks and impacts of technology-facilitated harm.

As discussed throughout this book, at the time of writing, the Online Safety Act 2023 is being turned into regulatory practice by OFCOM, the media regulator. How can anyone confidently claim that the act needs to be strengthened when its strength has not been tested. The rest of the rhetoric in the statement equally feeds into the scapegoating narrative, with little focussed anywhere but platforms (bizarrely referred to as “the polluter”) except for point five.

Similar, in the political arena, the politicians want to remain tough on platforms, and in the Labour Manifesto for the 2024 election (which they went on to win)⁵

Children and young people face significant harm online, with inappropriate content too easily available at their fingertips on a smartphone. We have seen an increase in extreme misogynistic content online driving a culture of violence against women. Labour will build on the Online Safety Act 2023, bringing forward provisions as quickly as possible, and explore further measures to keep everyone safe online, particularly when using social media. We will also give coroners more powers to access information held by technology companies after a child's death.

Even though the current Online Safety Act 2023 remains largely untested. How do they know the current legislation will not help keep everyone safe online? And if they are confident it will not, why was it not challenged during its long and winding passage through Parliament?

Particularly significant, however, is a further statement in the manifesto:

Regulators are currently ill-equipped to deal with the dramatic development of new technologies, which often cut across traditional industries and sectors. Labour will create a new Regulatory Innovation Office, bringing together existing functions across government. This office will help regulators update regulation, speed up approval timelines, and co-ordinate issues that span existing boundaries. Labour will ensure the safe development and use of AI models by introducing binding regulation on the handful of companies developing the most powerful AI models and by banning the creation of sexually explicit deepfakes.

⁵ <https://labour.org.uk/wp-content/uploads/2024/06/Change-Labour-Party-Manifesto-2024-large-print.pdf>

The new government, it seems, want to regulate the regulators. Which does suggest that they feel regulators, at present, will not fit with the political narrative.

So, given the analysis conducted in this book, and looking across the broad range of research I have done in this area, here are some predictions I will return to in the next volume:

- More calls to “strengthen” the online safety act. By strengthen that means further scapegoat industry for failing to stop whichever new harm emerges from a media story.
- The regulator, whose roadmap to implementing the regulatory needs of the online safety act will become increasingly criticised by other stakeholders who have not read the codes of practice for failing to “hold platforms to account” and letting them off the hook” for failing to prevent online harms, even though, on the evidence presented to the regulator platforms are doing more than most people are willing to accept. They will become the second online harms scapegoat.
- Industry will continue to be told to “do more”, without defining the extent of “more”, only to be told its not enough.
- Certain NGOs will continue to fuel moral panics because their business model is based on there being a significant concern for young people and an ideology that only they can solve it.
- Schools will still fail to deliver the sort of education young people are calling for and instead will fall back on popular prohibitive messages in primary school and these almost non-existent coverage in secondary settings
- Politicians will still think being seen to be tough on platforms gets them good media coverage and reflects that they do not really have sufficient understanding of the complexities of keeping children safe online or that fact that technology, alone cannot solve it. And they will continue to ignore progressive public health or education policy that might go some way to fixing the ecosystem, because that is too hard and does not fit in with political cycles.
- And young people will still call for better education and for those around them to understand the issues they face better and provide more effective support, and not call for big tech billionaires to be brought to heel. And they will be routinely ignored because “we” know better.

This may seem like a pessimistic view. However, I would prefer to regard it as a pragmatic one. Even though we have had close to twenty years of policy discussions around online safety it is still in its infancy compared to other areas of prohibitive social policy (such as the failed ‘War on Drugs’ – Grey 2010). With this in mind and reflecting upon the last twelve years as I have in the two volumes of this analysis, it is fair to suggest that without a rapid change in direction, things are likely to get worse before they get better.

A Personal Reflection

On a personal level, writing this book has been something of a cathartic experience, but not one without significant annoyance. Toward the end of the first volume I raised concerns about the impact of technological intervention on young people’s rights and questioned whether these dystopian interventions would actually result in the intended goal, which is, we are told, to ensure young people can go online without risk of harm.

Certainly, I am not the only academic voice who questions the validity of prohibitive policy, and I have referenced the work of excellent progressive voices in writing this book. There is a growing

consensus (and it is no coincidence that all these scholars engage in qualitative research directly with young people) that this single stakeholder model of accountability is failing and education and progressive, harm reduction focussed, response professionals are far more important than issuing fines to global platforms.

When conducting the policy and legal analysis, exploring data sets and reading political and media debates, it is difficult to remain entirely rational and objective because it is clear how broken the ecosystem is.

During the writing of this book, speaking to the young woman whose case I explored in detail in chapter 6, brought home in a very personal way how vulnerable young people can be utterly let down by systems that are staffed by professionals who do not understand the issues, are “trained” by media narratives, who project their own moral biases onto safeguarding decisions and, bluntly, are overworked and look for the lowest effort “solution” to the issue in front of them. In the young woman’s case many adults let her down, because it was easy to dismiss her behaviour as criminal, and divert, which meant “it wasn’t really too serious for her”, rather than dealing with the far more complex issues in supporting a young person who was clearly groomed, whose views were ignored by the systems would add to an already overstretched workload for the professionals and one that potentially left an abuser free to abuse further.

In the last few weeks, a couple things remind me of how far we still need to come.

In a recent panel I was invited to speak on, there was discussion of emerging legislation around online safety and children’s privacy. One participant spoke of how “a child having a smartphone is like having a loaded gun”, while another said, “we have to legislate because there is proof that children are addicted to smartphones⁶”. These were (very) senior people, one from an NGO and another a regulator. Whose rhetoric was far more about emotion than it was about rational discussion.

As a somewhat bizarre aside (although as I draw this book to a close this seems appropriate in a reflective section), I recently had to visit the minor injuries unit at my local hospital after a disagreement with a razor blade. As is typical in these situations one of the medical practitioners asked of what I was a professor. When I explained a little detail of my job and work (while they worked on my damaged finger) they agreed with the importance of the need for this sort of thing and then stated, “My son is 25 and he spends all day in his room playing Fortnite, I think they should ban it”.

It seems, wherever I go I meet people who think online technology should be banned. It is far rarer to meet someone who talks about the need for better education/mental health services/reporting routes on platforms.

Talking further with the practitioner, I did ask whether there might be other issues affecting her son (such as unemployment, living in an extremely deprived part of the country and mental health issues) and when we explore this, she agreed things might be more complex than just banning a particular game. She went on to say the most upsetting aspect to her was hearing the language used when he is playing these games. I suggested you heard similar language on the touchline at local football matches. She agreed and said she had observed that at their local football club. As we

⁶ I have not covered much on the addiction narratives around smartphones aside from the Haidt book. Suffice to say there are many publications and little consensus, reflecting the challenges in understanding peer reviewed literature among policy makers, who are more likely to explore research that aligns with their own confirmation biases.

departed, my finger back to something resembling a finger, she said she had enjoyed our chat and it had helped her see things differently.

It's amazing what a little discussion will achieve. Which is why it is so frustrating to see the political classes pursuing their ideological obsession and sunk costs, and refusing to engage in conversations that suggest there might be a better way.

Finally, in a recent stakeholder workshop, which was a coming together of several progressive voices in the sector to discuss how we might achieve better outcomes for children, a senior industry practitioner talks about some tools they had co-created with young people. He was saying how much he learned from listening to the young people.

It is something I hear a lot.

I can remember many years ago attending a class in a local school for a discussion with young people about online safety. The head of the school asked if she could sit in on the session. Once the students had got over the discomfort of the head being sat there (none of them had spoken to her before) they were as open and honest as most of the groups of young people I have spoken to and eventually the joined in with the discussion and asked questions while being honest about her own ignorance of many of the issues the young people were raising. After the session had finished, she thanked me and said how much she had learned "by listening to young people".

It amazing what happens when we listen to young people.

References

Bryon Review (2008). Safer children in a digital world: The report of the Bryon review. London: Department for Children, Schools and Families.

Gray, J. (2010). Why our drug laws have failed: a judicial indictment of war on drugs. Temple University Press.

Phippen, A. (2017). Children's online behaviour and safety: Policy and rights challenges. London: Springer.

Phippen, A., & Street, L. (2022). Online resilience and wellbeing in young people. London: Springer.