Cyberbullying in New Zealand

$1b cost in 2023

Report to Netsafe
9 September 2023

SENSE PARTNERS
DATA LOGIC ACTION

Foreword from Netsafe

Welcome to the digital era, where being online has given many opportunities for connecting, creating and collaborating.

The internet is changing how we think about how we talk to one another, our work and how we learn.

The digital age has also given rise to challenges, and one of the most pressing among them is measuring the economic cost of online harms.

As an online safety organisation dedicated to fostering safer online spaces and places, we are pleased to be able to repeat this study we last did back in 2017 to examine the economic costs of online harm. This research examines the tangible and intangible costs to New Zealand due to online harms.

This report reveals economic impacts on individuals, employers and broader institutional systems.

Beyond the emotional and psychological toll experienced by people at the receiving end of harm, the economic impacts are significant. From lost productivity at work and increased healthcare costs to legal expenses and withdrawal from participating online.

Through this report, we shine a light on the hidden costs of online harm. Awareness is the first step towards change, and with this knowledge, we can begin to make changes to create safer spaces for everyone.

I want to thank Sense Partners for this report. Their attention to this topic demands our collective attention and action.

I hope to work together to minimise the economic costs of online harms and use the money saved to build better online spaces and places for the enjoyment of all the people in New Zealand.

Key points

New Zealanders are more online, online harm is more prevalent and cyberbullying is more recognised. The cost to individuals, communities and interventions are substantial. We estimate the societal cost of cyberbullying is $1,071m a year, up from $444m in 2018. Most of it remains in the social rather than the private realm. Meaning the costs and benefits are best borne in a collective approach.

We updated the estimated cost using the same methodology used in 2018, noting that the international literature has matured substantially in that period.

There is more awareness of online harm, with more people experiencing harm themselves and know friends who have experienced it. Concern for and experience of online harm is greatest for parents, women, rainbow and ethnic communities.

Over half (58%) of those experiencing harm reach out friends and family first. While people are more likely to seek professional advice (including GPs, nurses and counsellors) than five years ago, less than 1 in 5 do so. People are more likely to use free services (particularly education and online resources) and strategies than paid services (only 1 in 5).

Online frauds and scams are not included in our estimates but are under reported harm. Netsafe reported circa $35m of financial loss in 2022. Police data shows 93% of fraud and deception is not reported to the Police, and 97% of cybercrime is not reported. Conservatively the true cost of online fraud could be over $200m-$470m a year. We need better reporting and not treat online crime differently from other crimes.

Online harm is increasing even as the online landscape changes rapidly. AI is such an emerging case, which could be leveraged for cyberbullying, frauds and scams, making them more prevalent and more harmful. We need to prepare for the harm we face today and risks of tomorrow by:

- Continue to support free online resources (most preferred strategy) to prevent and manage online harm and their credibility and accessibility. This is important for government and charities that support these activities.
- Appropriate training for health professional to deal with cyberbullying, it is a public health issue, not a tech one. This means continued investment in training to ensure online safety and online harm reduction are active components of public health training.
- Improve categorisation and reporting of online crime including frauds and scams; reporting rates remain too low for online issues, which are no longer separate from the real world. Singapore and Australian models of an independent anti scam agency should be explored.

Cyberbullying is already costing society $1b a year. Targeted and appropriate investment will pay dividends.

Contents

Introduction

Netsafe commissioned Sense Partners to update a 2018 report which estimated the societal cost of cyberbullying in New Zealand.[1] There is a growing consensus that bullying and cyberbullying are harmful, although the approach to quantifying it remains variable.

We have brought together a methodological approach, new survey data, and identified gaps. We conservatively estimate that cyberbullying costs New Zealand society $1,071m in 2023, up from $444m in 2018. This is the cost per year in harms to individuals, community and interventions.

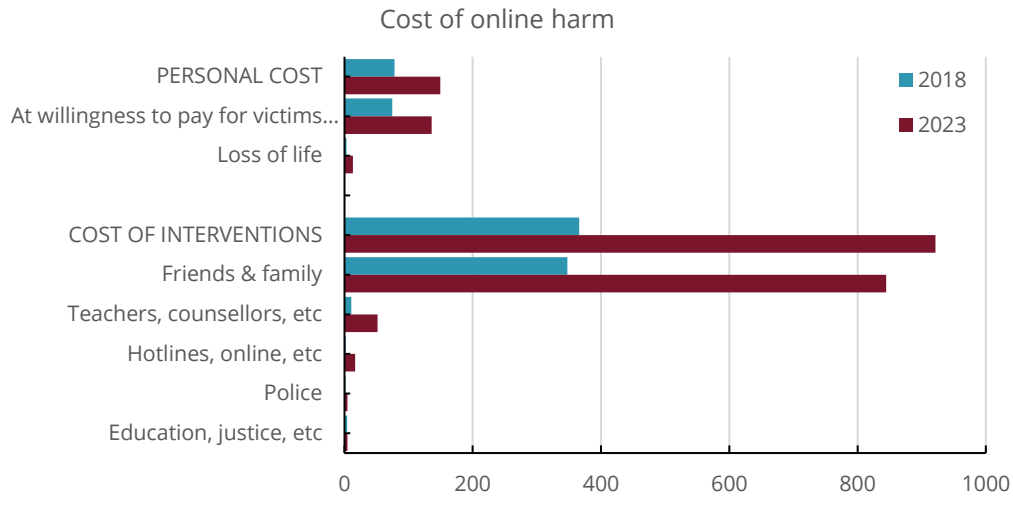
Cost of online harm $1,071 in 2023

Netsafe describe online-bullying as:

> *"Online bullying (also known as cyberbullying) is when a person uses digital technology to send, post or publish content with the intention to harm another person or a group."*

We updated our 2018 methodology with fresh survey data on New Zealanders' perceptions and experiences of online harm. We found that experience of online harm has increased substantially. This is leading to increased awareness and willingness to pay for services to prevent and mitigate online harm. People are most likely to rely on free resources and reach out to family and friends when experiencing online harm. People are also increasingly likely to seek professional support than five years ago. Increased prevalence, increased use of unpaid (but with an economic cost) and paid services, and an increase in wages and cost of public service delivery has lead to the cost of online harm to increase from $444m in 2018 to $1,073m in 2023.

FIGURE 1: ONLINE HARM COST NEW ZEALAND $1B IN 2023



Source: Sense Partners estimates

[1] https://netsafe.org.nz/cyberbullying-cost/

The harm framework

Harm reduction or damage limitation had a long history as a framework used to reduce negative impacts of drug use. Interventions to reduce harm typically operate from a set of desired goals that show options to reduce harm[2].

This typically looks at three specific areas. Costs to individuals, or personal harm. The cost to community or societal harm (here the focus is on friends and family in particular). The cost of interventions, typically by government.

The Ministry of Health publishes a Drug Harm Index using this framework to count all the harm and costs associated with drug use. We use a similar framework here to look at the cost of cyberbullying.

This approach allows us to observe the costs borne by different parts of society. It can identify under-resourcing of interventions relative to harms experienced by individuals and the community. It can also be used to assess the effectiveness of new policies or interventions.

The literature, interpretation and the harm framework we established in 2018 are summarise in Appendix A. We replicated the survey work from 2018 to ensure consistency and comparability of approach. We updated the cost of delivery from The Treasury's CBAx tool and updated the Statistical Value of Life to the latest estimate.[3]

A more detailed discussion of the methodology is presented in Appendix A and in our original paper in 2018.

New Zealand survey evidence

We updated the 2018 survey of New Zealanders on their perceptions and experiences of online harm, strategies employed to deal with online harm, and willingness to pay for prevention and mitigation. Survey details can be found in Appendix B.

Five years on, the data shows high awareness of online harm, increased experience of online harm, and greater use of unpaid and paid resources to deal with online harm, and a greater willingness to pay to be safe online.

The survey also gives us some information on population segments across gender, age, ethnicity, family makeup, housing tenure, and household income. However, our sample size is not large enough to slice and dice the data too much. The segmented data should be interpreted with care.

Nevertheless, we report concentrations in summary format which help us understand potential areas to focus on in the next iteration. The segmentation data shows that online harm is more prevalent for women, ethnic minorities, and families with children.

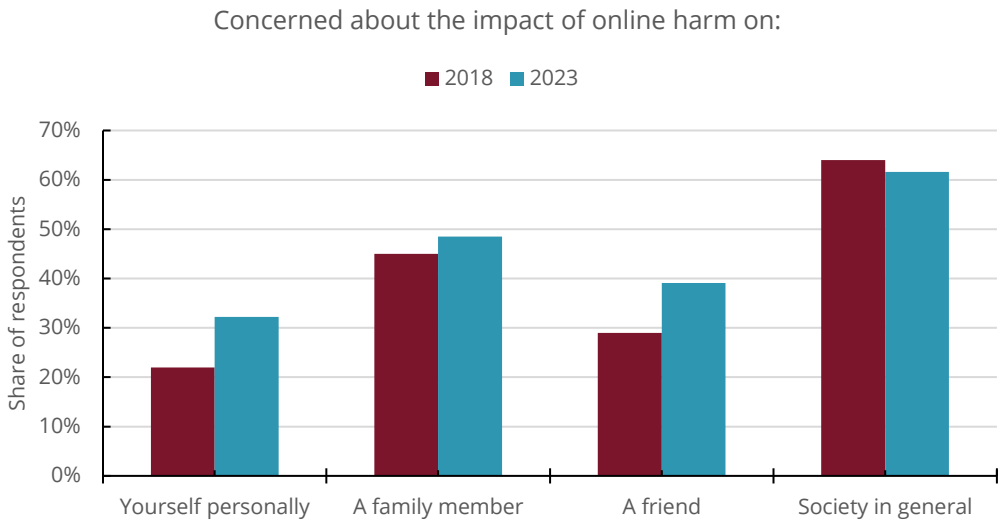
---

[2] Newcombe (1992)
[3] https://www.newsroom.co.nz/govt-to-pay-three-times-the-price-for-faster-safer-journeys

Perceptions of online harm have increased

62% of New Zealanders are concerned about the impact of online harm to society in general. This is broadly similar to 2018 (64%). However, perceptions of online harm have increased for themselves, family members, and friends. Concern is highest among women, middle aged and older New Zealanders (45+) and ethnic minorities (Māori, Pasifika and Asian). Families with children were most concerned.

FIGURE 2: INCRESASED CONCERN ABOUT ONLINE HARM CLOSER TO HOME



Source: Primary Purpose survey for Netsafe

FIGURE 3: WOMEN, ETHNIC MINORITES AND FAMILIES MOST CONCERNED

Total concerned about the impact of online harm

Shaded blue represents more prevalent

		Impact of online harm on: Yourself personally	A family member	A friend	Society in general
Segment of respondent					
Gender	Male				
	Female				
Age	18-29				
	30-44				
	45-59				
	60+				
Region	Auckland				
	Wellington				
	Other North				
	Canterbury				
	Other South				
Ethnicity	NZ European				
	NZ Māori				
	Pasifika				
	Asian				
	Other				
Housing	Renting				
	Home is owned with a mortgage				
	Home is owned freehold				
Dependent children	Yes				
	No + Unsure				
Household income	<$50K				
	$50-100K				
	>$100K				

Source: Primary Purpose survey for Netsafe

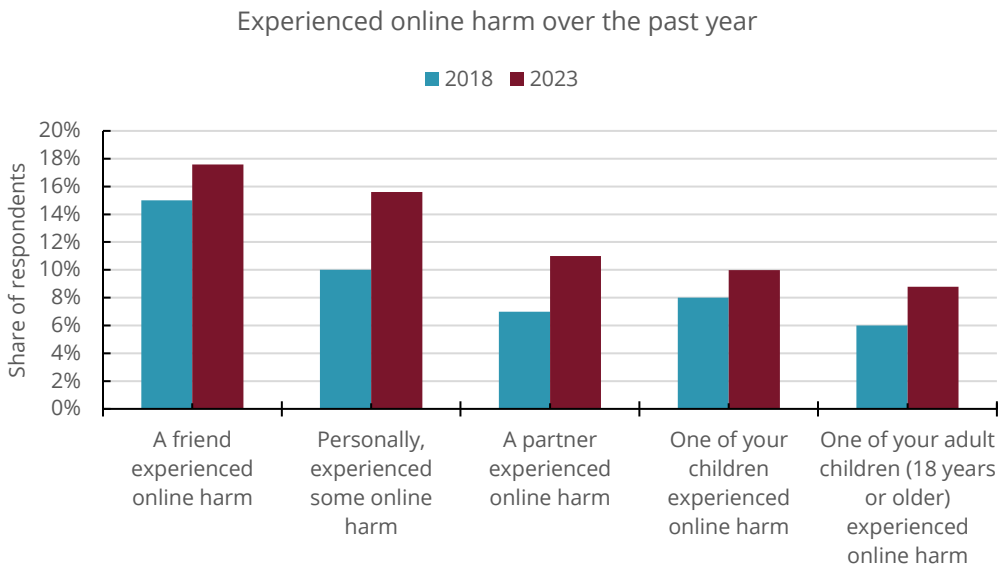
Experience of online harm has increased

Perceptions of online harm are mirrored in experience. More people are likely to have:

- known a friend who has experienced online harm (18% in 2023 vs 15% in 2018)
- experienced online harm themselves (16% vs 10%)
- a partner who has experience online harm (11% vs 7%)
- a child who has experience online harm (10% vs 8%)
- an adult child has experience online harm (9% vs 6%)

The survey shows a widespread increase in online harm experienced in peoples family and friends circles. While the majority of people still do not experience online harm, the surveys shows a worrying prevalence of online harm and a broad based increase online harm over the five years to 2023.

FIGURE 4: ONLINE HARM IS WIDELY EXPERIENCE AND HAS INCREASED OVER THE LAST FIVE YEARS

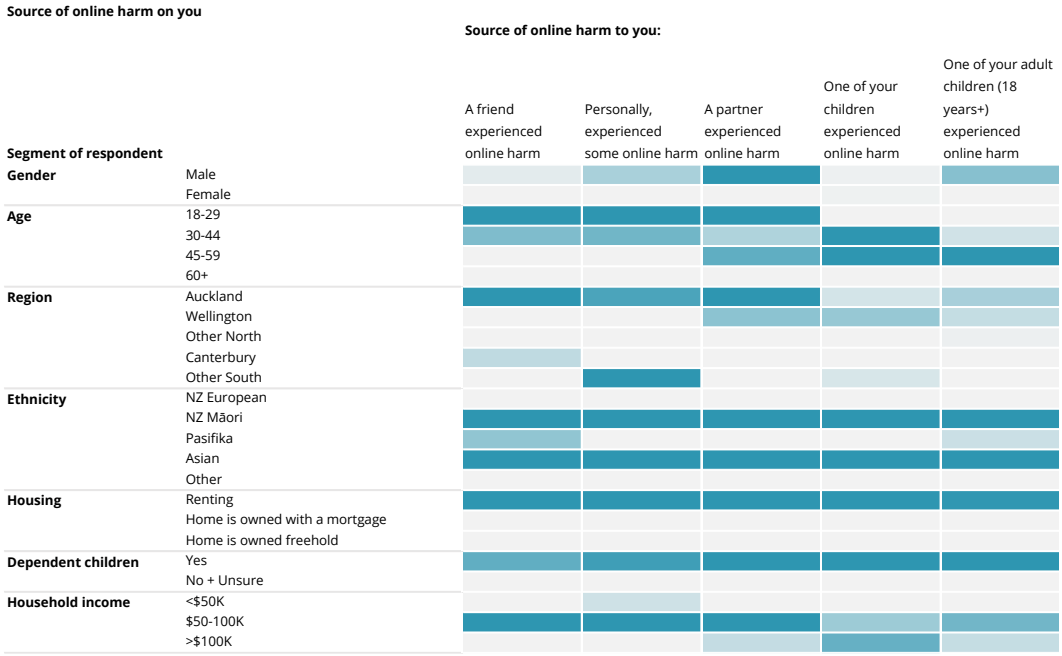


Source: Primary Purpose survey for Netsafe

Unlike perceptions of online harm, experience of online harm is borne by slightly different groups. In particular, younger people (and renters) are more likely to experience online harm, compared to perceptions of harm being higher among middle aged and older people.

Experience of online harm is highest among Māori and Asians. Pasifika are less prevalent, but it is not clear how much is related to unwillingness to report. People with children were more likely to experience online harm, suggesting a focus on young people is critical.

FIGURE 5: ONLINE HARM IS EXPERIENCED BY YOUNGER PEOPLE, MAORI AND ASIANS IN PARTICULAR, AND PEOPLE OF MODERATE INCOME.



Source: Primary Purpose survey for Netsafe

Strategies to deal with online harm is weighted toward free tools…

while perceptions and experience of online harm is increasing, the strategies employed to deal with online harm remains similar.

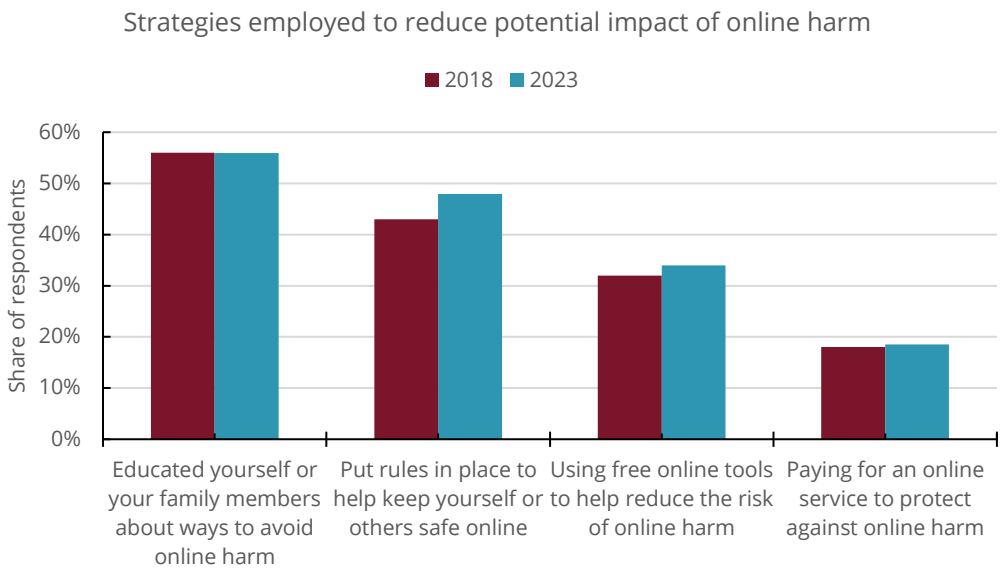
The most common approach (56% of New Zealanders) is to use educate themselves or family members about way to avoid online harm. This highlights the importance of well curated, evidence based, credible and accessible education material to support this most common strategy to stay safe online.

More people are putting rules in place to keep themselves safe, 48% in 2023, up from 43% in 2018. People are more likely to use free online tolls (up slightly from 2018). However, the use on paid online services remains the least preferred option (at 1 in 5).

The survey allows for people using multiple strategies. We can see that those who experience more online harm (women, young people, families, ethnic minorities) also deploy multiple strategies to stay safe online.

There appears to be a cost barrier to keeping safe online. This suggests existing tools and resources that educate, inform, prevent and manage online harm continue to be important. This is relevant for funding authorities, generally government funding and philanthropic organisations.

FIGURE 6: PEOPLE USE PERSONAL STRATEGIES TO KEEP SAFE ONLINE



Source: Primary Purpose survey for Netsafe

FIGURE 7: THOSE AFFECTED BY ONLINE HARM MORE LIKLEY TO PURSUE MULTIPLE STRATEGIES

Strategies employed in the last year to help reduce the impact of online harm

		Strategies employed:			
Segment of respondent		Educated yourself or your family members about ways to avoid online harm	Put rules in place to help keep yourself or others safe online	Using free online tools to help reduce the risk of online harm	Paying for an online service to protect against online harm
Gender	Male				
	Female				
Age	18-29				
	30-44				
	45-59				
	60+				
Region	Auckland				
	Wellington				
	Other North				
	Canterbury				
	Other South				
Ethnicity	NZ European				
	NZ Māori				
	Pasifika				
	Asian				
	Other				
Housing	Renting				
	Home is owned with a mortgage				
	Home is owned freehold				
Dependent children	Yes				
	No + Unsure				
Household income	<$50K				
	$50-100K				
	>$100K				

Source: Primary Purpose survey for Netsafe

Box A: Online fraud & scams: $200m-$470m a year

Online fraud and cybercrime is a grey area for the purposes of our report. However, there is now an increasingly well scoped definition of online harm. The World Economic Forum for example published the following typology of online harm in 2023:[4]

- Treats to personal and community safety: Child sexual abuse material; Child sexual exploitation material; Pro-terror material; Consent those praises, promotes, glorifies or supports extremist organisations or individuals; Violent graphic content; Content that incites, promotes or facilitates violence; Content that promotes, incites or instructs in dangerous physical behaviour; Grooming for sexual abuse; Recruitment and radicalisation; Technology facilitated abuse; Technology-facilitated gender based violence; Child sexual exploitation and abuse.
- Harm to health and wellbeing: Material that promotes suicide, self-harm and disordered eating; Developmentally inappropriate content.
- Hate and discrimination: Hate speech; Algorithmic discrimination.
- Violation of dignity: Online bullying and harassment; Sexual extortion.
- Invasion of privacy: Doxxing; Image-based abuse.
- Deception and manipulation: Disinformation and misinformation; Deceptive synthetic media; Impersonation; Scams; Phishing; Catfishing.

One aspect of online harm is deception and manipulation, under which fraud and scams feature. The data on frauds and scams is very limited, in part because it is not commonly reported officially. For example, The New Zealand Crime and Victims Survey reported 609,000 victims of Fraud and cybercrime. The same survey showed that 93% of fraud and deception, and 37% of cybercrime was not reported to the policy. The reporting is very low and international research suggests it is linked to stigma of falling victim to such crimes, often low value of frauds, lack of awareness of their rights, among other factors lead to such low reporting.

Netsafe reported around $35m was lost to online fraud and scams in the last year. But given such low reporting, the true cost could be much larger. As a reasonable estimate we believe the true costs could be:

- $200m (assuming unreported crime unit values are 5% of reported) to
- $470m a year (assuming unreported crimes are 20% of unit values).

Very low reporting rates shows there is much work to be done in this area and should be a priority policy area.

[4] https://www.weforum.org/reports/toolkit-for-digital-safety-design-interventions-and-innovations-typology-of-online-harms

Use of professional help is increasing

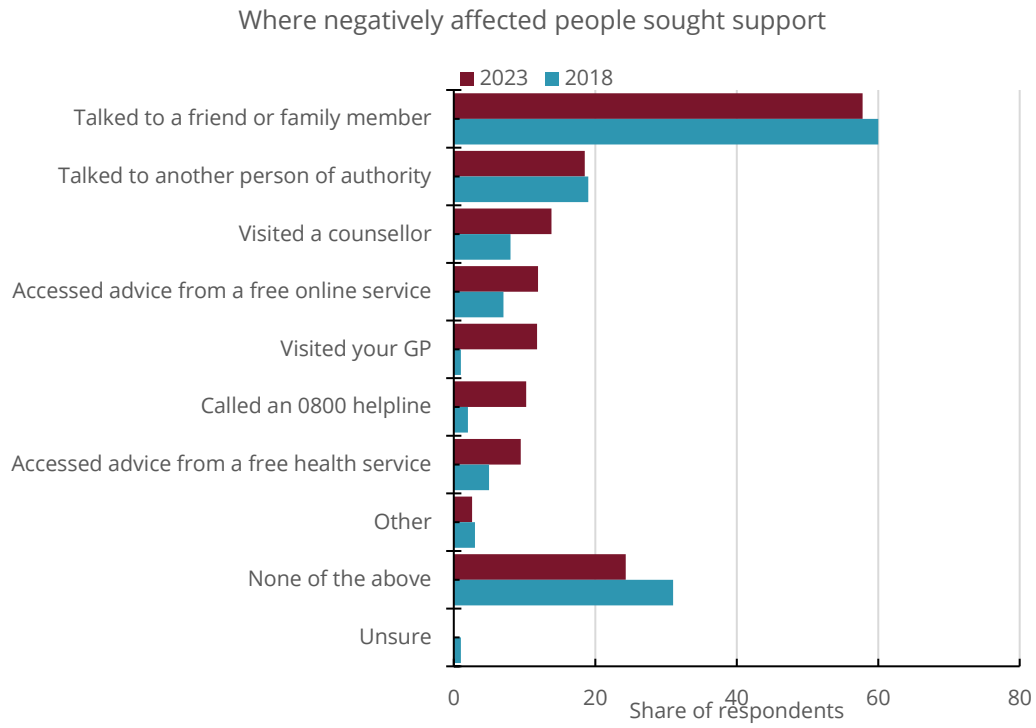
For those who experienced online harm over the past year, the most common place to seek support was friends and family. This is understandable and was also the highest in 2018 (58% talked to friends and family in 2023, versus 60% in 2018).

Fewer people were likely to ignore instances of online harm, but 24% (or one in four) did not seek any kind of support, this is still too high although an improvement from 31% not seeking help in 2018.

There has been a significant increase in the use of professional services. More people are likely to visit a counsellor, access advice from free online services, visited their GP, called an 0800 helpline, and accessed advice from free healthcare service. This is encouraging, showing greater willingness to seek professional support, but also the importance of a broad array of support points for those negatively impacted by online harm. Online harm cannot be funnelled into one solution, rather online harm is pervasive and our wider social and health infrastructure needs to be ready and able to respond.

The survey clearly shows that people are far more likely to use free strategies over paid services. Put another way, cost is a barrier to keeping safe online. High quality resources to educate ourselves, friends and family, including in curriculum for schools and health professionals will help to ensure consistent and credible advice to help prevent and manage online harm.

FIGURE 8: SINGIFICANT INCREASE IN THE USE OF PROFESSIONAL SUPPROT



Source: Primary Purpose survey for Netsafe

More willing to pay to mitigate and address online harm

The survey has shown clearly that people perceive increased online harm closer to their inner circle. Concerns about society in general remains broadly stable, but perceived harms have increased personally, for family and for friends.

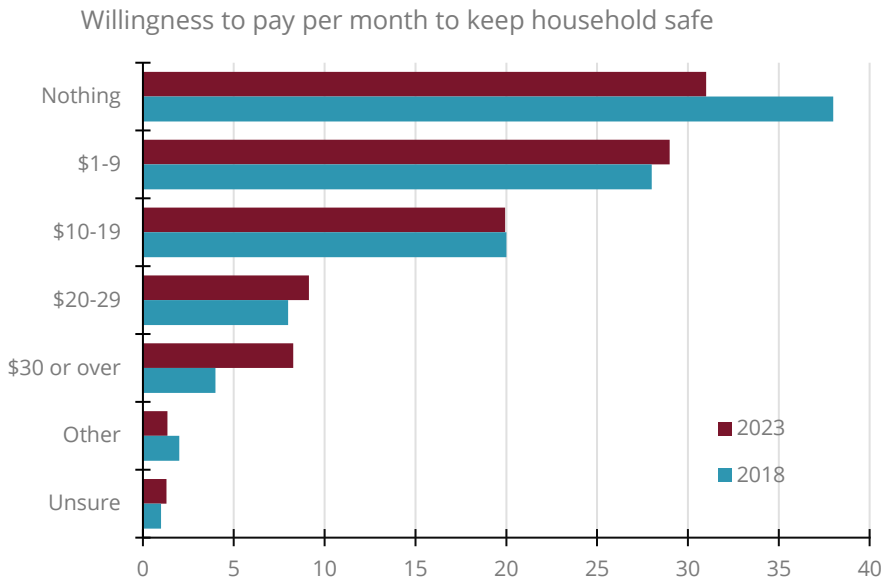
Experiences of harm also bears out these patters. More people are experiencing online harm personally, among friends, partners, and children. There is a high concentration of online harm among women, younger people, families and ethnic minorities.

This increased experience of online harm is leading to people adopting more strategies to stay safe online: getting educated, putting in clear rules, using free online tools and some paid online tool. The most favoured approaches are free, in favour of paid services.

People are most likely to reach out to friends and family for support, but there is increasing use of professional services (nurses, counsellors and GPs for example).

This maturing experience and understanding of online harm is also reflected in peoples' willingness to pay to prevent and manage these costs. Fewer people report they will pay nothing (31% in 2023 vs 38% in 2018). The biggest increase is in people willing to pay $30 or more. This is consistent with public perception that it costs and is worth more to stay safe online, than was the case in 2018.

FIGURE 9: THERE IS GREATER AWARENESS OF ONLINE HARM AND WILLINGNESS TO PAY TO STAY SAFE



Source: Primary Purpose survey for Netsafe

Conclusion

The online world exposes some people to harm from cyberbullying. The cost to individuals, communities and interventions are substantial. We estimate the cost of cyberbullying was $1,071m in 2023, up from $444m in 2018.

We used local surveys, international studies, and approaches from other fields to develop a framework and identify the costs. We believe these costs estimates are conservative, as we are yet to fully account for workplace effects and online fraud and cybercrime effects (as these are still severely under reported and unquantified).

The results from 2023 suggest we need to continued focus on what works and new interventions to guard against emerging risks:

- Destigmatising asking for help – it is still a barrier, with a quarter of those experiencing or witnessing cyberbullying did not seek help. We need to continue our education and awareness campaigns to ensure online harm is understood in a common language and avenues for support are widely known and accessed.
- Investing in curriculum for schools. Young people are disproportionately affected by cyberbullying and we need to prepare them to deal with an ever evolving online space, where the scope for harm appears to be increasing.
- Investing in curriculum for public health professionals. People are increasingly likely to seek professional support. So, these professionals (GPs, nurses, counsellors, etc) need to be prepared with appropriate knowledge and solutions. This training needs to be updated regularly and should form a regular part of their continuing education. Making modules available for managers more generally would also help.
- Raising awareness of cyberbullying and where to seek help. Victims of cyberbullying are most likely to turn to their family and friends for help. Awareness of where to go for help would be valuable. Making sure there is free, credible and accessible resources is critical.
- International coordination of legislation and enforcement of cyberbullying remains a critical issue. The online space is global by nature and lack of coordination between global efforts has the potential to create regulatory and enforcement gaps, perpetuating or concentrating harm.
- A coordinated approach to online frauds and cybercrime. We estimate online frauds and scams cost $200m to $470m, and reporting is too low. These are not separate from other fraud and crime, but they need a different approach, as traditional approaches aren't working. Singapore and Australia have established dedicated units to do this and is worthy of consideration in New Zealand.

Appendix A: Cyberbullying harm framework

Context

The online world has become fast integrated in our lives. New technologies have become ubiquitous, but our understanding and resourcing to manage risks have not kept pace.

Communication is a key part of economic activity. Our ability to communicate facilitates trade, transfer knowledge and deepen social connections.

Successive new technologies have allowed communication to happen faster and they have become widespread more quickly. For example, the telephone took nearly 50 years to become mainstream in the US (over half of households). The home computer took nearly 20 years, the internet 13 years and smartphone 7 years. In New Zealand, internet access is now near universal.

The widespread adoption of the internet and its applications allow much wider reach and greater intensity of interactions, both positive and negative.

Like the real world, the online world also has a small population of bullies. The anonymity available online can mean cyberbullying is more intense than in person. Some surveys show greater negative impact on happiness and wellbeing from cyberbullying than social bullying.

In a survey we commissioned, conducted by UMR, only 10% of respondents had personally experienced online harm, although the impact was higher among women, young people and ethnic minorities (we did not collect more detailed information on disability, gender, etc.).

While most of the population appear unaffected by cyberbullying, some have intensely negative experiences. This motivates the rest of the paper.

Costs

Netsafe describe online-bullying as:

> *"Online bullying (also known as cyberbullying) is when a person uses digital technology to send, post or publish content with the intention to harm another person or a group."*

They note cyber-bullying is typically aggressive, often involves a power imbalance and as others have suggested, is usually repeated.[5]

Typically, we think of harm from cyber-bullying as affecting individual victims. But there can be societal and even economic impacts. Moreover, while victims are harmed, perpetrators and bystanders can also experience harmful impacts.[6]

---

[5] Patchin and Hinduja (2006) describe cyber-bullying as the "wilful and repeated harm inflicted through the medium of electronic text".

[6] See the meta-analysis in Slavtcheva-Petkova 2015 and Rice et al. 2015 who note the mental health consequences include lower self-esteem, and depressive symptoms for both victims and perpetrators, and Sinclair et al. 2012 who discuss impacts on bystanders. Janson et al. (2009) point out the complex role of the bystander.

Harm to individuals

Harm to individuals is dependent on both exposure to internet and mobile where cyber-bullying can occur and harm deriving from experiences of cyber-bullying experiences. Figure 1 shows this two-tiered process.

Cyberbullying is less prevalent than bullying in general. A meta-analysis[7] of various studies found a prevalence rate of 15% for cyberbullying – lower than a 35% prevalence for bullying in general. Assessing the prevalence of cyberbullying is difficult with different numbers reported in a variety of contexts[8]. It is hard to quantify harm when prevalence is not known precisely.[9]

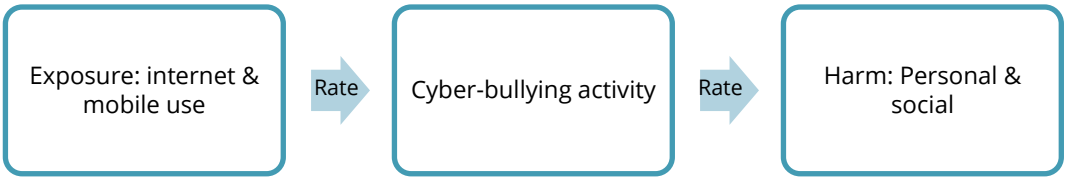
But risk and harm are linked to internet opportunities. A third of New Zealand teens now spend 4 or more hours online in an average day.[10] Being "bullied online is the risk that upsets children the most, even though it is among the least common".[11]

The New Zealand Attitudes and Values Study (NZAVS) asks participants aged 18 years and over (from a random sample of 15,822 New Zealanders) if they have experienced cyberbullying. 11.5% of the respondents to the 2014/15 NVAS had experienced cyber-bullying.[12] A whopping 46% of 18-19 year-olds had experience cyber-bullying.

Young people experience more cyberbullying. 19% of New Zealand teens experienced: "an unwanted digital communication that had a negative impact on their daily activities", based on a 2017 survey completed by 1,001 teens.[13] Internet use also differs between teenage boys and girls, so differences in exposure by gender will translate to different degrees of aggregate harm. And children that use social-networking sites are more likely to come across online risks.[14]

FIGURE 10: NOT ALL CYBERBULLYING RESULTS IN PERSONAL OR SOCIAL HARM

Responses for Turkish adolescents to cyber-bullying cases in two Turkish schools



How does cyber-bullying harm? To be clear, some victims of cyberbullying are not upset or disturbed.[15] But cyberbullying is often associated with many emotional and psychological conditions[16], including stress, lower self-esteem[17] and life satisfaction.

[7] Modecki et al. (2014)
[8] see Menesini and Salmivalli 2017
9 The Slavtcheva-Petkova 2015 meta-study suggests 17 of 63 articles about online risks related to cyber-bullying focus on prevalence.
[10] Netsafe (2018b)
[11] Livingstone (2014)
[12] see Steiner-Fox et al. (2016)
[13] Netsafe (2018a)
[14] see Staksrud et al. (2013)
[15] see Smith et al. (2013) and Ortega et al. (2000)
[16] see Harris 2009 and Hango 2016
[17] see Waisglass 2017

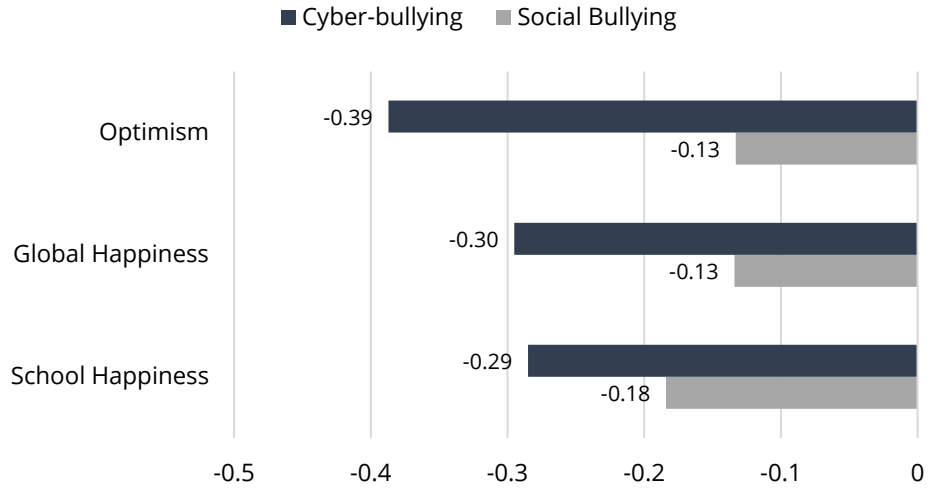
For example, Moore et al. (2012) look at cyber-bullying in a middle school is the US and report:

> *"modest, but pervasive relationships between experiences of electronic bullying and victimisation and adolescents' life satisfaction reports across a variety of life domains."*

Controlling for gender and grades confirm the correlations but note their cross-sectional data cannot confirm causal effects.[18] Similarly, a Spanish study[19] show statistically significant negative relationships between cyber-bullying and three measures of happiness and well-being (see Figure 11). Their findings a based on a large survey of Spanish 10-12 year-olds.

Effects on victims of bullying can be long-lasting.[20] Bullying also reduces educational achievement.[21] Cyber-bullying has been linked to poorer health outcomes using Canadian data.[22]

FIGURE 11: CYBER-BULLYING CORRELATES WITH LOW SELF-REPORTED WELL-BEING
Correlation between bullying and indicators of subjective well-being, Spanish 10-12 year-olds



Source: Navarro, Raúl, Roberto Ruiz-Oliva, Elisa Larrañaga and Santiago Yubero (2015)

[18] Waisglass (2017) reports similar impacts on well-being from bullying among a survey of female university students.
[19] Navarro et al. (2015)
[20] see Heydenberk and Heydenberk (2017)
[21] see Ponzo (2013)
[22] See Soyeon et al. (2017)

What precisely drives these types of behaviour? Slonje et al. (2013) discuss different classifications of cyber-bullying behaviour and note cyber-bulling has been classified by:

(i) type of media,

(ii) (ii) action, and

(iii) (iii) the content of messages.

It is the aggregate impact of these types of factors that drives the negative impact.

TABLE 1: CLASSIFYING CYBER-BULLYING

Types	Media type	Action type	Content of messages
Factors	• mobile phone calls • text messages • picture/video clip bullying • emails • chatroom • instant messaging • websites	• flaming • online harassment • cyberstalking • denigration • masquerade • outing • exclusion	• threat of physical violence • abusive or hate-related • name calling • death threats • ending of platonic relationship(s) • sexual acts • demands/instructions • threats to damage existing relationships • threats to home/family • menacing chain messages.

NB. Media-type is Smith et al. (2008), action-type Willard (2006) and content-type Rivers and Noret (2010).

Source: Slonje et al. (2013)

Societal/community

But harm from cyberbullying can be more pervasive than individuals. It can affect friends and family, fray community cohesion, and place additional demands on services like health.

People subject to bullying are likely to change their behaviour, typically to riskier behaviours associated with depression and other risk factors.[23] Cyberbullying is also associated with mistrust of people.[24] Bullying can derail the subjective well-being of school communities.[25] These costs are often difficult to observe, measure and establish causal links to the root causes.

The drug harm literature suggests public agencies find it difficult to assess social costs[26] and this is likely to be also true of cyber-bullying. Sir Peter Gluckman notes:

[23] Juvonen and Galvan 2009
[24] Hango 2016
[25] See the three-year study conducted by Heydenberk and Heydenberk (2017)
[26] McFadden 2009

> *"the cyber world is less amenable to third-party prevention or interruption; thus, large ripple effects are possible across on-line communities. - Office of the Prime Minister's Chief Science Advisor (2017).*

Interventions

Who do victims turn to?

Any interventions to prevent harm from cyber-bullying needs to be calibrated to effectiveness and cost. One of the key features of cyberbullying is the reluctance of victims to seek help.

Technology can mean children think adults won't understand. Amichai-Hamburger (2013) makes it plain:

> *"…paradoxically in this field, the older generation, which is meant to lead, has less knowledge and less exposure than the more expert and technologically experienced youth."*

Victims of cyber-bullying do turn to friends but are reluctant to tell others – a point documented in a study of Turkish teenagers (see **Error! Reference source not found.**).[27] Victims tends to tell friends first and teachers last.[28]

What types of interventions work?

But there is a long history of successful intervention programmes for bullying.[29] Intervention programmes that address general bullying can sometimes be adapted to deal with cyberbullying.[30]

Children typically report technical ways of responding to cyber-bullying, for example:

- blocking certain people from contacting you online,
- changing passwords, user names or e-mail addresses
- deleting anonymous text messages without reading them.

Different types bullying generates different coping mechanisms.[31]

Curriculum materials are at an early stage – but there is encouraging evidence from traditional bullying. Effective curriculum material can reduce social bullying by around 20 percent.[32] While the digital world can be a challenge when it comes to identifying cyberbullying, it also offers potential solutions to better deliver resources. These can be used to provide new techniques

[27] Topçu et al. (2015)
[28] also see Smith et al. (2008)
[29] Olweus and Limber (2010)
[30] Slonje et al. (2013)
[31] Pieschl et al. (2013)
[32] Smith et al. (2012)

(including for example, cybermentoring, film clips and websites),[33] for dealing with cyber-bullying.[34]

Legislative responses

Many jurisdictions have also introduced new laws to fight cyberbullying including Australia, Europe, New Zealand, the UK and the US but there is no common legislative response across countries.[35]

Australian research[36] to advise their government on the desirability to create new cyberbullying offences found:

- most Australians were unclear about the circumstances that make cyberbullying criminal
- suggest there is a strong role for schools to provide information on the relationships between cyberbullying and the law
- most young people wanted a new civil enforcement regime but not adults
- laws that criminalise impulsive acts from young people who cannot work through the impact of their actions would be unhelpful
- stakeholder groups recognised that social media sites should conform with new laws but noted the difficulties in terms of enforcement[37]
- there is a strong role for schools to play in disseminating information on cyberbullying and the law.[38]

Towards assessment

Non-market impacts

There is no market for mitigating cyber-bullying so unlike consumer markets with regular information on prices of goods and services, valuing the costs and benefits associated with cyber-bullying is more difficult.

There are at least three methods of non-market valuation that are relevant:

(i) Stated preference methods that use specially designed surveys to extract people's willingness to pay for a programme or outcome;

[33] Slonje et al. (2013) note the success of the computer-based KiVa program in Finland.
[34] Slonje et al. (2013)
[35] Spears et al. (2014)
[36] Spears et al. (2014)
[37] Ultimately a Senate Inquiry found existing legislation to be "adequate", but made nine recommendations to address cyber-bullying
[38] Campbell et al. (2008) note the requirement for Australian and New Zealand schools to have anti-bullying policies that may or may not include reference to cyber-bullying.

(ii) Revealed preference methods that build up values of non-market goods or services based on the choices people make. Hedonic price methods, such as using house prices information to reveal preferences over school quality, are the standard technique.

(iii) A well-being or life satisfaction method, that tries to directly estimate the benefit of non-market goods or services based by modelling impacts of often self-reported measures of well-being.

The cyber-bullying literature that tries to quantify the cost of cyber-bullying, tends to connect cyber-bullying to indicators of well-being including self-reporting well-being.

To our knowledge there are no New Zealand studies that tightly link cyber-bullying to self-reported well-being. So, we pursue survey-based methods to uncover willingness to pay.

Fujiwara and Campbell (2011) provide a substantive overview of these methods including the difficulties with each approach. Stated preference methods suffer since answers often dependent on the framing of questions, often related to how respondents process risk and probability. On the other hand, econometric methods can be prone to difficulties such as teasing out causality, functional form specification, measurement error and isolating and attributing value to the policy issue using hedonic methods. As Fujiwara and Campbell note:

> *"…it is clear that both the revealed preference (for hedonic market studies) and the life satisfaction approaches will work best for policies with significant impacts on market prices (eg, the housing market) or life satisfaction. When this is not the case, stated preference may be the only viable method for valuation of the policy impact.*

Willingness to pay

One approach to assessing the benefits of a programme to mitigate or arrest cyber-bullying is to first survey willingness-to-pay for attributes of the programme and then sum over of the population to show societal preference for the programme.[39] The approach centres by probing for choices that reveal societal willingness-to-pay for a given programme. The approach rests heavily on survey design and researchers have made clear some of the difficulties that include respondents' understanding of risk[40] and bias over choices over hypothetical outcomes.[41]

Nevertheless, the methods are common where other price-based methods to value benefits are unavailable. Moreover, the method has precedence in the bullying literature. Persson and Svensson (2013) use a survey of residents in Örebro, Sweden who were introduced to bullying and the likelihood of bullying at schools in Örebro before being asked to make a series of discrete choices over hypothetical programmes. These programmes had varying degrees of effectiveness in reducing bullying (see Table 2).

[39] A stated preference technique in the characterisation by Fujiwara and Campbell (2011).
[40] see Hammitt and Graham, 1999
[41] Willingness-to-pay has also been estimated using hedonic regressions (see Griffith and Nesheim 2008 and Mandell and Wilhelmsson 2010) and a variety of other methods. see Murphy et al. 2005 for further critique.

FIGURE 12: DISCRETE CHOICES CAN BE USED TO UNCOVER WILLINGNESS-TO-PAY

Question 1. Currently about 4,800 pupils attend grade 7 to grade 9 in the municipality of Örebo. You have the opportunity to vote on a program that would reduce the number of pupils being bullied according to the information given below.

Pupils being bullied without the program	480 pupils
Pupils being bullied if the program is implemented	100 pupils
Cost per tax payer	500 Swedish kronor

Are you in favour of the municipality implementing this program at the given cost?

☐ Yes ☐ No

Source: Persson and Svensson (2013)

TABLE 2: ATTRIBUTES AND ATTRIBUTE LEVELS IN WILLINGNESS-TO-PAY FOR BULLYING PREVENTION IN PERSSON AND SVENSSON (2013) STUDY

Attribute	Attribute levels
Risk reduction	(1) 100 less bullied (per school/academic year)
	(2) 240 less bullied (per school/academic year)
	(3) 380 less bullied pupils (per school/academic year)
Cost	(1) 200 Swedish kronor
	(2) 500 Swedish kronor
	(3) 1,000 Swedish kronor
	(4) 2,000 Swedish kronor
	(5) 5,000 Swedish kronor

Source: Persson and Svensson (2013)

Then a variety of socio-economic and demographics features of the respondents were gathered.

Summing up across the population's willingness-to-pay for the intervention programme, Persson and Svensson (2013) estimate the aggregate willingness to pay per statistical bullying victim is 585,090–835,280 Swedish kroner or $97,900 to $139,800 New Zealand dollars in todays' terms.

Bringing it all together

The literature does not present a consensus approach to identifying and collating the total cost of cyberbullying. We have made several judgements to present a broad-brush picture of likely social costs of cyberbullying.

We looked at the costs of cyberbullying in three parts.

- First, the cost to individuals at the willingness to pay rate we observed in our survey and loss of life. Survey respondents may expect additional public expenditure to reduce cyberbullying. As an example, public health spending is 6x private health spending.
- Second, we count up the cost of time and resources spent on the victims of bullying (including by family and friends, counsellors and so on). We estimate most of this is related to the time of friends and family in prevention and mitigation of cyberbullying.
- Third, we looked at the long-term cost of cyberbullying on mental health, physical health and productivity. But the research is still nascent, and we could not make confident estimates. We left this aside for now, to pick up in future iterations of this work.

Future work

Our analysis was limited in scope, with an intention to expand in a future iteration. We have identified a range of gaps and limitations and have identified a range of areas to focus on.

In a future iteration of this work, we would like to better estimate personal harm and people's willingness to pay by replicating the work by Persson and Svensson (2013) for New Zealand. We would like to collect much finer grained data over a larger sample to better tease out intensity of harm and identify potential mitigation tools. Such a detailed survey is too costly for this exercise and requires a more academic or government led approach to accomplish.

We would like to better understand and quantify costs of interventions by various agencies and what harms they specifically relate to, for example across GPs, Ministry of Justice, police, teachers, etc. This ideally requires greater collaboration with government agencies to better understand how these costs are experienced.

We would like to explore the lifetime effect of cyberbullying on personal health, education and employment outcomes. This could be done by expanding work done on workplace bullying by WorkSafe, by including an online harm module in the next iteration.

We would also like to explore the costs associated with bullies as well, rather than the victims as we have done in this paper. In some literature, bullies have been linked to higher probability of future criminal risks. However, this field of analysis remains too immature to incorporate into this work.

References

Amichai-Hamburger, Y. (2013). Youth internet and well-being. *Computers in Human Behavior*, 29, 1-2.

Campbell M, D A Butler and S M Kift (2008), "A school's duty to provide a safe learning environment: Does this include cyberbullying?", *Australia and New Zealand Journal of Law and Education*, 13(2): 21-32.

Czajkowski, Mikołaj, Nick Hanley and Jacob La Riviere, (2016), "Controlling for the Effects of Information in a Public Goods Discrete Choice Model," *Environmental and Resource Economics*, Springer, European Association of Environmental and Resource Economists, vol. 63(3), pages 523-544, March.

Fujiwara, Daniel and Ross Campbell (2011), "Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches A Discussion of the Current Issues", HM Treasury, Department for Work and Pensions, London

Griffith, Rachel and Lars Nesheim, (2008), "Household Willingness to Pay for Organic Products," CEPR Discussion Papers 6905, C.E.P.R. Discussion Papers.

Hammitt, J. K and J D Graham, (1999). Willingness to pay for health protection: Inadequate sensitivity to probability? *Journal of Risk and Uncertainty*, 18(1), 33 – 62.

Hango, Darcy (2016), "Cyberbullying and cyberstalking among Internet users aged 15 to 29 in Canada", Insights on Canadian Society, Release date: December 19, 2016, Statistics Canada,

Harris, M. (2009). *Bullying, rejection and peer victimization: A social cognitive neuroscience perspective,* New York, NY: Springer

Heydenberk, R A, and Heydenberk, W R (2017), Bullying reduction and subjective wellbeing: The benefits of reduced bullying reach far beyond the victim. *International Journal of Wellbeing*, 7(1), 12 – 22.

Janson, G, J Carney, R Hazler, and I Oh (2009), "Bystanders' reactions to witnessing repetitive abuse experiences", *Journal of Counseling and Development*, 87(3), 319-326

Juvonen, J., & Galvan, A. (2009). "Bullying as a means to foster compliance". In M. Harris (Ed.), *Bullying, rejection and peer victimization: A social cognitive neuroscience perspective* (pp. 299-318). New York, NY: Springer.

Livingstone, Sonia (2014), "Risk and harm on the internet", In Jordan, A and D Romer, (eds.) *Media and the well-being of children and adolescents*, Oxford, Oxford University Press, 2014, pp. 129-146.

McFadden Consultancy (2016), "Research Report: The New Zealand Drug Harm Index" (2nd edn). Wellington: Ministry of Health.

Mandell, Svante and Mats Wilhelmsson, (2010), "Willingness to pay for sustainable housing,", Working Papers 2010:6, Swedish National Road & Transport Research Institute (VTI).

Matthews, Mark and Tom Kompas, (2015), "Coping with Nasty Surprises: Improving Risk Management in the Public Sector Using Simplified Bayesian Methods," *Asia and the Pacific Policy Studies*, Wiley Blackwell, vol. 2(3), pages 452-466, September.

Menesini, Ersilia and Christina Salmivalli (2017), "Bullying in schools: the state of knowledge and effective interventions", Psychology, Health & Medicine, 22:sup1,240 – 253

Modecki, Kathryn L, Jeannie Minchin, Allen G Harbaugh, Nancy G Guerra, and Kevin C. Runions, (2014) "Bullying Prevalence Across Contexts: A Meta-analysis Measuring Cyber and Traditional Bullying", *Journal of Adolescent Health* 55 602e611

Moore, Page Malmsjo, E Scott Huebner and Kimberly J Hills (2012), "Electronic Bullying and Victimization and Life Satisfaction in Middle School Students", *Social Indicators Research*, 107:429–447

Moreno, Elias, Francisco Javier Giron, Francisco Jose Vazquez-Polo and Miguel Negrin, (2009), "A Bayesian Net Benefit Approach to Cost-effectiveness Analysis in Health Technology Assessment" *International Journal of the Economics of Business*, Taylor & Francis Journals, vol. 16(3), pages 323-345.

Murphy, J., G P Allen, T H Stevens and D Weatherhead (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30(3), 313–325.

Navarro, Raúl, Roberto Ruiz-Oliva, Elisa Larrañaga and Santiago Yubero (2015), "The Impact of Cyberbullying and Social Bullying on Optimism, Global and School-Related Happiness and Life Satisfaction Among 10-12 year-old children", *Applied Research in Quality of Life*, March.

Netsafe. (2018a). Harmful digital communications in New Zealand: Annual Population Survey 2017. Wellington, NZ: Netsafe. Retrieve from https://www.netsafe.org.nz/annual-populationsurvey-2017

Netsafe. (2018b). New Zealand teens and digital harm: Statistical insights into experiences, impact and response. Retrieved from https://www.netsafe.org.nz/NZ-teens-and-digital-harm_statisticalinsights_2018.pdf

Newcombe, Russell (1992), "The reduction of drug-related harm – A conceptual framework for theory, practice and research, in P O'Hare et al, "*The reduction of drug-related harm*, London Routledge,

Office of the Prime Minister's Chief Science Advisor (2017), Youth Suicide in New Zealand: A Discussion Paper, 26 July.

Olweus, Dan and Susan P. Limber (2010), "Bullying in School: Evaluation and Dissemination of the Olweus Bullying Prevention Program", *American Journal of Orthopsychiatry,* 80(1), January, 124 – 134

Patchin, J W,and S Hinduja (2006), "Bullies move beyond the schoolyard: A preliminary look at cyber-bullying", *Youth Violence and Juvenile Justice*, 4(2), 148–169.

Persson, Mattias and Svensson, Mikael (2013), "The willingness to pay to reduce school bullying", Economics of Education Review, Elsevier, vol. 35(C), pages 1-11.

Pieschl, S., Porsch, T., Kahl, T., & Klockenbusch, R. (2013), Relevant dimensions of cyberbullying—Results from two experimental studies. *Journal of Applied Developmental Psychology*. Volume 34, Issue 5, September–October 2013, Pages 241-252

Ponzo, Michela (2013), "Does bullying reduce educational achievement? An evaluation using matching estimators," *Journal of Policy Modeling,* , vol. 35(6), pages 1057-1078, Elsevier.

PWC (2018), "The economic cost of bullying in Australian Schools," Alannah and Madeline Foundation.

Rice, Eric,Robin Petering, Harmony Rhoades, Hailey Winetrobe, Jeremy Goldbach, Aaron Plant, Jorge Montoya, and Timothy Kordic (2015), "Cyberbullying Perpetration and Victimization Among Middle-School Students", *American Journal of Public Health*,105(3): e66–e72. March

Sinclair, K O, Bauman, S, Poteat, V P, Koenig, B., Russell, S T (2012). Cyber and bias-based harassment: Associations With academic, substance use, and mental health problems. *Journal of Adolescent Health*, 50.

Slonje, R., P K Smith, and A Frisén (2013), "The nature of cyberbullying, and strategies for prevention", *Computers in Human Behavior*, 29, 26-32.

Slavtcheva-Petkova, Vera, Victoria Jane Nash and Monica Bulger (2015), "Evidence on the extent of harms experienced by children as a result of online risks: implications for policy and research, *Information, Communication and Society*, 18:1, 48-62

Smith, P K (2012), "Cyberbullying and cyber aggression". In S.R. Jimerson, A.B. Nickerson, M.J. Mayer, M.J. Furlong (Eds.), Handbook of school violence and school safety: International research and practice (pp. 93–103). New York, NY: Routledge

Soyeon Kim, Michael H Boyle and Katholiki Georgiades, (2017), "Cyberbullying victimization and its association with health across the life course: A Canadian population study", *Canadian Journal of Public Health*, 108(5-6):e468–e474

Spears, B, C Taddeo, T Swirski, M Keeley, I Katz, P Collin, T Daly and S Bates (2014), "Research on youth exposure to, and management of, cyberbullying incidents in Australia: Part C – An evidence-based assessment of deterrents to youth cyberbullying (SPRC Report 11/2014)". Sydney: Social Policy Research Centre, UNSW Australia

Staksrud, Elisabeth, Kjartan Ólafsson and Sonia Livingstone (2013), "Does the use of social networking sites increase children's risk of harm?", *Computers in Human Behavior,* 29, 40–50

Steiner-Fox, H. W., Dutt, S. J., Christiansen, S. J., Newton, H. J., Matika, C. M., Lindsay, C., Sare, M. H. , Kapeli, S. A., & Stronge, S. (2016). Rates of cyberbullying among women and men in New Zealand in 2015. NZAVS Policy Brief, 3.

Topçu, Çigdem, Özgür Erdur-Baker and Yeşim Çapa-Aydin 11 (6) (2008) "Examination of Cyberbullying Experiences among Turkish Students from Different School Types" *Cyberpsychology and Behavior*

Waisglass, Nicole, (2017), "The Lasting Effects of Cyber Bullying on Well-Being", Brescia Psychology Undergraduate Honours Thesis.

Appendix B: Survey Methodology

We reference several statistics that use an omnibus survey we commissioned via *Primary Purpose*. The process replicates the survey undertaken by UMR's national survey in 2018. UMR has since ceased operations. The methodology of their survey as the following:

- Results in this report are based upon questions asked in a nation-wide omnibus survey. This is an online survey of a nationally representative sample of 1,000 New Zealanders 18 years of age and over.
- Fieldwork was conducted from the 14-23 April 2023.
- The margin of error for sample size of 1,000 for a 50% figure at the 95% confidence level is ± 3.0%.
- Weighting was used to ensure an accurate reflection of the general population. The data was weighted by region, gender, age and ethnicity.
- In this study negative impact from online harm is defined on a scale where 0 = 'no impact at all' and 10 = 'a lot of impact.
- 6-10 defines negative impact, 5 is the mid-point and 0-4 is defined as no impact.

SENSE PARTNERS

DATA LOGIC ACTION