Journal of Experimental Criminology
https://doi.org/10.1007/s11292-023-09551-3

Can a police-delivered intervention improve children' online safety? A cluster randomised controlled trial on the effect of the "ThinkUKnow" programme in primary and secondary Australian schools

Tony Alderman[1] · **Barak Ariel**[1,2] · **Vincent Harinam**[1]

Accepted: 6 January 2023

Abstract

Purpose Online abuse of and by children is a global concern. Methods to prevent this phenomenon are diverse; however, less is known about police-led initiatives and their effectiveness in reducing the likelihood of becoming a victim or a perpetrator of cyber abuse among children. Specifically, there are no rigorous tests of the ThinkUKnow programme, to which hundreds of thousands of young people were exposed since 2006.

Methods We present results from a cluster randomised control trial conducted to evaluate the Australian version of the *ThinkUKnow* programme, delivered to students in primary and secondary schools. The programme consists of one face-to-face classroom-based training session delivered by at least one uniformed member of the Australian Federal Police. Post-test surveys among ($n=1954$) students were used to estimate the treatment effect.

Results Exposure to the programme significantly improves knowledge about cyber abuse but marginally impacted risk perceptions, engagement with risky behaviours, or willingness to report cyber abuse to adults or others. Treatment participants are more likely to report cyber abuse to the police than control participants. The legitimacy of the police also improves following the intervention among younger but not older students.

Conclusion *ThinkUKnow* leads to desirable consequences in some but not all indictors of potentially minimising the risk of cyber abuse to and by children. Replications, preferably with diverse populations and measures of long-term effects of behavioural modifications, are needed.

✉ Barak Ariel
barak.ariel@mail.huji.ac.il; ba285@cam.ac.uk

Tony Alderman
aa2093@cam.ac.uk; Tony.Alderman@afp.gov.au

Vincent Harinam
vh315@cam.ac.uk

[1] Institute of Criminology, University of Cambridge, Sidgwick Ave, Cambridge CB3 9DA, England
[2] Institute of Criminology, The Hebrew University of Jerusalem, Mt. Scopus, 9190501 Jerusalem, Israel

Published online: 18 January 2023

Keywords Randomised controlled trial · ThinkUKnow · School-based interventions · Police · Cyber abuse

The prevalence of online crime is increasing, and with it, the importance of online safety among children has risen. More than 95% of children aged between 8 and 17 have access to the internet (Australian Communications & Media Authority, 2013). Moreover, digital platforms and devices are being increasingly integrated into the educational system. The provision of these devices to school students presents youth with both opportunities and risks associated with internet use. Risk and harm mitigation for students is often taught in school programmes to promote youth safety online and encourage help-seeking behaviours (Polanin et al., 2022; Tanrikulu, 2018). The police are often involved in this approach, such as Australia's ThinkUKnow initiative, a national cyber-safety programme delivered to students by the Australian police. However, despite being delivered in Australia for more than a decade for thousands of students, the effectiveness of ThinkUKnow is presently unclear. As the programme is a police-delivered initiative, it begs the question of whether it can backfire, like other school-based, police-led interventions before it (e.g., Klenowski et al., 2010; Perry et al., 2003). In this paper, we report the findings of a large-scale experiment to test the effect of this programme.

Literature review

Definition and prevalence of cyber abuse

Cyber abuse is any type of abuse that happens on-line. Children are particularly at risk of cyber abuse from people they know and from strangers and can take multiple forms, including cyberbullying, emotional abuse, grooming, sexting, sexual abuse, and sexual exploitation (Bryce, 2017). The prevalence of these offences varies based on personal factors, as users' characteristics play a key role regarding the likelihood of cyberbullying among different groups (Camerini et al., 2020; Park et al., 2021; Shaikh et al., 2020). Nevertheless, it is agreed that cyber abuse is a common event (e.g., Evangelio et al., 2022; Selkie et al., 2016). Mishna et al. (2011) reported that the prevalence of cyberbullying can range from 10 to 35%. Olweus and Limber (2018) estimated its prevalence between 3 and 50%. Olweus and Limber (2018) show that considerable variation in the prevalence of cyberbullying is due to different methods by which empirical studies have measured the phenomenon. To this extent, Kowalski et al. (2014) use the general aggression model in their meta-analysis to identify studies, while Brachado et al. (2017) use the terms "cyberbullying", "cyberbullying", "internet harassment", and "internet bullying" to find eligible studies. Both meta-analyses have identities research with vastly differing rates of abuse given measurements which encompass a combination of offending, victimisation, lifetime prevalence, and periodic prevalence (Zhu et al., 2021).

What is agreed is that the prevalence of online child exploitation is likely to be far higher than indicated by official sources, given offenders' proclivity to conceal their activities (Bryce, 2017) on the one hand, and under-reporting on the other.

School-based interventions against cyber abuse

Research on cyber abuse remains a comparatively new field, but has gained a great deal of attention. While the number of studies on this topic is steadily increasing, many focus on the identification of risk factors and the impact of cyberbullying. Rigorous impact evaluations of effectiveness of interventions designed to prevent or mitigate the impact of cyber abuse remain limited and often suffer from methodological considerations (Abreu & Kenny, 2018; Calvo-Morata et al., 2020; Cantone et al., 2015; Evans et al., 2014; Gaffney et al., 2018; Hutson et al., 2018; Macaulay et al., 2018; Polanin et al., 2022; Tanrikulu, 2018; Zych et al., 2015). Still, the literature remains informative.

Educational messages are one method to successfully raise awareness of the risks and consequences of unsafe online behaviour among youth (Bryce 2010). The centrality of such engagement lies in framing cyber abuse as an everyday occurrence in the lives of students, clarifying that risky behaviours online can be avoided. Risky behaviours include, for example, sharing intimate images of oneself or another person, insulting others, disclosing private details to strangers, and agreeing to meet with strangers following an engagement online. Mishna et al. (2011) as well as Zych et al. (2015) conducted systematic reviews of interventions to prevent or reduce youth cyber abuse and found moderate to large effect sizes for changes in knowledge of cyber abuse and cyber safety following the programmes. However, these changes in knowledge did not extend to significant shifts in risky online behaviour (Mishna et al., 2011). The authors argued that the developers of these programmes needed to focus not only on the transmission of knowledge but changing the attitudes of students towards cyber safety.

Walsh et al. (2018) performed a systematic review of 24 trials of school-based education programmes to prevent child sexual abuse. These programmes were generally effective at increasing both the knowledge of prevention concepts and protective behaviours among students, with medium to large effect sizes observed in 18 of the 24 studies. Another systematic review (Gaffney et al., 2018) examined 24 RCTs and quasi-experimental studies of cyberbullying intervention or prevention programmes that dealt with school-aged participants. These interventions indeed reduced victimisation by 14% and perpetration by 10 to 15%, respectively, and RCT designs yielded larger effect sizes relative to quasi-experimental designs. However, these were more focused interventions, rather than programmes aimed at preventing cyber abuse more broadly.

Police-led programmes to reduce cyber abuse in schools

Police-led school-based interventions are one method of deterring future crime. Indeed, law enforcement can not only instruct students on law abiding behaviour and the pitfalls of criminal activity but build rapport with them such that students can rely

on law enforcement personnel when in need (Feinberg & Robey, 2009; Pennell et al., 2022; Radebe & Kyobe, 2021; Vandebosch et al., 2012, but *cf.* Broll & Huey, 2015). More generally, these engagements serve to improve police legitimacy among youth, which may be linked to increased reporting of future crimes (Mazerolle et al., 2013).

Shaw (2004) identified three different models of police engagement in schools: (1) officers permanently embedded in schools, (2) officers acting as educators, and (3) officers participating in comprehensive liaison schemes. Within the Australian context—where the present test takes place—one study used a waitlist cluster RCT to evaluate the effectiveness of a school-based protective behaviours programme on 611 first grade students (White et al., 2018). The treatment group demonstrated an increase in their knowledge compared to the control group. However, changes to students' intention to disclose abuse did not significantly increase. Unfortunately, additional randomised control trials were not conducted to evaluate the efficacy of any one of these programmes to reduce cyber abuse, revealing a clear research gap (see Diaz et al., 2021; Kenny et al., 2022; Nyberg et al., 2021; Solehati et al., 2022; White et al., 2019).

ThinkUKnow programme

One popular programme to prevent cyber abuse is the *ThinkUKnow* initiative, which aims at keeping children safe by producing education and training about cyber abuse. The programme has been implemented since 2006 in the UK, and it is available globally. In the UK, the programme is delivered by the National Crime Agency, by identifying the main threats to children—which presently focuses on sexual abuse as a core target area (thinkuknow.co.uk 2021). In Australia, the ThinkUKnow programme is led by the Australian Federal Police and delivered nationally in partnership with law enforcement and industry to "raise awareness and educate the community about preventing online child sexual exploitation" (Australian Federal Police, 2021) The programme offers bespoke training courses for children depending on their age ($5-7$, $8-0$, $11-13$, $14>$) and is usually administered by a policer officer. Training is provided in classroom settings, by highlighting the risks associated with online activity, how to identify abuse and how to report it.

Despite its popularity and use worldwide, little is known about the efficacy and cost-effectiveness of the ThinkUKnow programme. As far as we are aware, the only evaluation conducted by Davidson et al. (2009) concluded that whilst at the time nearly 14% of all UK students were exposed to the programme, a high proportion of students were unable to recall having it at all. Thus, whether there are benefits to this international initiative, to which a large population of students is exposed, remains unclear.

Police legitimacy and schools

As noted earlier, an adjacent issue when studying the effectiveness of school-based policing initiatives is the perceived legitimacy of the police following these engagements (Shaw, 2004). Such initiatives are meant to reduce crime, but one outcome is the potential effect on the relationship between law enforcement and young people.

Mazerolle et al. (2013) reported that school-based interventions, including police officers in schools, may increase perceptions of legitimacy through students interacting with police in a less formal setting. In turn, enhanced perceptions of police legitimacy may support compliance with laws and school policies around online safety and cyber abuse and could lead to greater reporting of cyber abuse.

School-based police interventions present an opportunity to harmonise socialisation across two of these domains. Through a randomised controlled field trial examining truancy and anti-social behaviour in students in Queensland, Mazerolle et al. (2021) demonstrated that a consensus-based programme involving police reduced self-reported anti-social behaviour. This was premised on a change in participants' perception of police legitimacy.

Methods

Research design

This study utilises a cluster randomised controlled design with two "blocks" comprised of primary or secondary school years, and a post-test measure only. These are common in education research (Donner & Klar, 2002), with randomisation occurring by class, grade, or school (Farrington & Ttofi, 2009; Gaffney et al., 2018; Ttofi & Farrington, 2011). Here, the control group also received the intervention but only *after* completing the survey (whereas the treatment group received the programme *before* completing the survey instrument). The settings for the two arms were identical but approximately 90 min apart. This procedure ensured that all students received the mandated programme while at the same time measuring its effectiveness (for more on this design, see Campbell and Stanley 2015:53–54; see also Ariel et al., 2022). This design also enhanced consistency and minimised costs, as it did not involve multiple programme deliveries to a single school.

Randomisation

We use the class as the unit of analysis, with all participants in a class randomly assigned to a control or treatment group. The class was assessed as the unit of analysis having an optimal balance of administrative complexity and statistical power. Within the Australian Capital Territory (ACT), there are approximately 60 public primary schools and 20 public secondary schools. Randomisation at this level would have resulted in 30 primary school clusters and ten secondary school clusters, but many of these schools have already exposed their students when the experiment was rolled out, and others have declined to participate in the experiment - thus leading to a significantly underpowered study. In contrast, randomisation at the individual level (i.e., students) would have been an arduous undertaking as it requires separating students within classes. Moreover, this approach would have diluted treatment integrity by increasing the likelihood of contamination as students in both the treatment and

control groups would likely discuss their experiences with each other. Finally, this randomisation strategy poses serious privacy concerns as it requires a substantial level of identification to allocate students into either group.

This study thus consisted of 51 clusters of primary school students in years 5 and 6 (block A) and 49 clusters of secondary school students in years 7 and 8 (block B) within the Australian Capital Territory - i.e., school years as the units of randomisation. Primary school clusters were then randomly assigned into 27 and 26 study arms, and both groups possessed an average class size of 22.6 students. Among secondary school clusters, 24 were designated as treatment groups, while 25 became control groups, with 18.2 and 18.1 students in each class, on average, respectively. All participating classrooms were allocated an identification number for the purposes of random assignment.

Sample characteristics

Gender

Among the primary school block, the treatment group comprised 49.4% females and 48.7% males, while the treatment group was 48.4% female and 50.7% male. Only 1.9% and 0.9% of those in the treatment and control groups were of unknown gender, respectively. In the secondary school block, the treatment group was 49.6% female and 49.3% male, with 1.1% declining to state their gender. Similarly, the control group was 50.6% female, 47.3% male, and 2.1% of unknown gender.

Age

For the primary school block, 50.5% and 51.6% of the treatment and control groups were born in 2008, respectively. Approximately one-third of the treatment (32.3%) and control (32.3%) groups were born in 2007. Among the secondary school block, 48.8% of the treatment group and 49.6% of the control (38.1%) groups were born in 2006. Just over one-third of those in the treatment (34.3%) and control groups were born in 2005.

Nationality

In total, 78.4% and 80.6% of students in the primary school treatment and control groups were born in Australia, respectively. Similarly, 79.4% and 76.1% of those in the secondary school treatment and control groups were born in Australia, respectively.

Online presence

For all groups, students spent more time on the internet on non-school days than on school days. Students in the year 5 – 6 treatment group averaged 2.97 and 4.01 h on the internet on school days and non-school days, respectively. Those in the year 5 – 6 control group averaged 2.93 and 3.81 h on the internet on school days and

non-school days, respectively. Students in the year 7 – 8 treatment group averaged 4.92 h on school days and 5.09 on non-school days, while students in the control group averaged 4.61 h on school days and 4.7 on non-school days.

In general, the treatment and control clusters are well balanced with a comparable number of clusters and students within each cluster. There was a near even gender distribution in both treatment and control groups for each cohort. With regard to age, the control and treatment groups have broadly similar distributions of age in each cluster. Importantly, the overwhelming majority of participants were born in Australia.

Procedure

Schools that had requested a ThinkUKnow session were contacted to determine their willingness to participate in an evaluation. Interested schools were sent an email with further information on how the trial would run, requirements of participation, time scales, and measures taken to ensure the privacy, and confidentiality of students. Schools that agreed to participate were provided with dates and times for the delivery of ThinkUKnow. This decision ensured that the evaluation was delivered in a manner that most closely resembled a business-as-usual environment. Each participating school provided a list of classes that would be used to create the control and treatment conditions.

As noted earlier, data were collected in schools, with students accessing the survey instrument with an internet-connected device before (control) or after (treatment) programme delivery. Compared to other survey administration methods, group-administered surveys have a high response rate (Bachman & Schutt, 2017), minimising non-response bias (Fowler, 2013). The online survey was administered through *Qualtrics*. Instructions and the survey link were provided to a school contact officer and loaded to a Google Classroom site that students could access. A non-uniformed officer was present to oversee the survey delivery and address any questions from the school staff. Importantly, students' participation in the online survey was voluntary.

Treatment conditions

The ThinkUKnow programme was delivered in a class-based setting by uniformed Australian Federal Police officers. At the bare minimum, the programme was delivered to an entire year group. However, at small schools with a limited number of classes across year cohorts, both years participated simultaneously. Delivery comprised a 60-to-90-min presentation that included case studies, short videos, and interactive question-and-answer segments.

Instrument

This study utilised a post-test only measure to estimate the causal effect of the treatment. Conducting pre and post measures necessitated tracking responses from individual students, which was not feasible. Furthermore, this design also reduced the

risk of instrumentation bias from repeated exposure to the survey instrument in a short period (Shadish et al., 2002).

The survey instrument (see Supplementary Materials A) was constructed to measure six outcomes: students' (1) knowledge, (2) certainty of apprehension, (3) protective behaviours they might take, (4) perceptions of risk online, (5) likelihood of reporting victimisation, and (6) perceptions of police legitimacy. Furthermore, the instrument inquired about students' age, gender, and minority status (e.g., country of birth, language spoken at home). The items used to measure minority status were drawn from Murphy and Cherney's (2011) research on the response of ethnic minorities in Australia to procedural justice-based policing. In addition to the length of time spent online, the survey instrument also included an item on whether participants had access to a variety of online devices.

Knowledge items were developed based on the ThinkUKnow curriculum, with possible responses of "true", "false", or "do not know". Years 5−6 and 7−8 had two different sets of items for this outcome based on the extended content presented to the older cohort. Both instruments contained items relating to knowledge of online safety and privacy and items relating to cyberbullying. The years 7−8 instrument also included a series of items relating to image-based abuse and the sharing of images online. The certainty of apprehension items asked participants to indicate how much risk there was of getting caught and punished for a specific activity. This dimension involved items on a four-point response scale ranging from "no risk" to "a very great risk". These were adapted from Hirtenlehner et al.'s (2013) research on deterrence and juvenile offending.

Two outcome dimensions were operationalised for risky behaviours: scenarios reflecting various behaviours relating to perpetrating cyberbullying and a range of protective behaviours. These items and outcome measures are intended to examine whether any change in knowledge from exposure to ThinkUKnow extends to the stated likelihood to behave more safely.

The reporting of cyber-victimisation adapts a construct from a study on help-seeking intentions (Wilson et al., 2005). The items have a seven-point response from "extremely likely" to "extremely unlikely".

Fourteen items associated with the perception of police legitimacy outcome measure are used. These were adapted from Sunshine and Tyler's (2003) examination of the role of procedural justice and legitimacy in shaping support for policing and Tankebe's (2009) research into public cooperation with the police in Ghana. Responses to these items are based on a 5-point scale ranging from "strongly agree" to "strongly disagree". To measure the reliability of the instrument, we calculate the Cronbach's alpha for each outcome measure in each block (see Table 1). As is evident, the internal consistency of each outcome group is relatively high.

Response rate

The years 5−6 cohort had a response rate of 95.8%, while the year 7−8 cohort had a response rate of 90.8%. A total of 1148 and 806 valid responses from years 5−6 to years 7−8, respectively, represent 16.9% and 13.5% of the population of students in those years.

Table 1 Study dimensions and inter-rater reliability

Outcome measure	Years 5/6		Years 7/8	
	Cronbach's alpha	*N* of Items	Cronbach's alpha	*N* of Items
Knowledge: online safety	0.893	17	0.774	12
Knowledge: cyberbullying	0.749	9	0.771	7
Knowledge: image-based abuse	N/A	N/A	0.865	10
Knowledge: all	0.918	26	0.917	29
Risk perception: certainty	0.959	14	0.955	14
Risk perception: severity	0.959	14	0.958	14
Behaviours: perpetration	0.945	6	0.918	6
Behaviours: protective	0.779	7	0.843	7
Reporting: adult	0.865	6	0.908	6
Reporting: police	0.943	6	0.884	6
Reporting: friends	0.917	6	0.897	6
Reporting: handle oneself	0.881	6	0.954	6
Legitimacy	0.875	14	0.943	14

Statistical methods

Descriptive and inferential statistics were leveraged to estimate the treatment effect. Summary statistics were used to calculate the mean and standard deviation of each dimension for each cluster in the treatment and control groups. Independent samples *t*-tests were then calculated on the cluster mean scores (see Robson, 2002), and a Bonferroni correction was included due to the repeated *t*-tests (0.05 divided by number of comparisons). The estimates were computed separately for each statistical block.

Results

Main effects

Testing for newly gained knowledge

Table 1 presents a matrix of all the findings of the experiment, including a comparison of control and treatment group scores and the results of the independent samples *t*-test for both years 5 – 6 and years 7 – 8 blocks. First, we see that significant treatment effects are concentrated in the knowledge category. Across all knowledge outcomes, the treatment group clusters had a higher average percentage of correct responses than the control group clusters. However, significant improvements were recorded only for online safety in both blocks, and knowledge on cyberbullying significantly increased for the years 5 – 6 block only.

The items in which the differences were more substantial than others were as follows: “Cyberbullying is against school rules, but it isn’t against the law”; “you can control where all images of you end up”; “when I sign up for online accounts, the default privacy setting is usually ‘private’”; “there are laws that prevent companies from owning or keeping content I post or distribute through their service”; and “there are services through my school that can help if I run into issues online”. Training on image-based abuse online delivered to the years 7 – 8 block alone did not yield significant variations against control conditions ($t = 1.513$, $p > 0.10$).

Risk perceptions: certainty of apprehension and severity of punishment

Regarding the certainty of apprehension (Table 2), mean differences between the treatment and control groups for both cohorts were not statistically significant. The average perceived certainty of apprehension was high across all test groups, hovering above three on a four-point Likert scale.

Similar results are observed for the severity of punishment. While the treatment group results were marginally higher for both years 5 – 6 and years 7 – 8, the differences were not statistically significant at acceptable thresholds. The average perceived severity of punishment was quite high as well, reaching 4.5 on a 6-point scale.

Risky behaviours: perpetration and protection

While treatment group means were marginally higher for years 5 – 6 (treatment = 5.441, control = 5.370) and years 7 – 8 (treatment = 5.387, control = 5.351) for reportage of perpetration, the differences were not statistically significant. On the other hand, as it pertains to the likelihood of protective behaviour, the average treatment score was significantly higher than that of the control group for year 5 – 6 (treatment = 5.225, control = 5.050), but the difference in the years 7 – 8 group (treatment = 5.043, control = 4.943) was not statistically significant ($t = 1.020$, $p > 0.10$).

Likelihood of reporting cyber abuse

Table 2 presents the mean response scores and *t*-test results for respondents’ willingness to report cyber abuse to a trusted adult or the police, with reciprocal questions about the likelihood of handling the risk by oneself or with friends. One intention of the intervention was to increase cyber-abuse reportage and decrease the likelihood that students would handle the matter themselves or report it to friends. As such, a high mean response for parents or police and a low mean response for friends or handling the issue on one’s own both favoured the treatment.

We see that the likelihood of reporting to the police increased in both blocks ($t = 5.192$, $p < 0.001$; $t = 3.413$, $p < 0.001$), but no significant variations emerged across all other comparisons. The programme did not increase the likelihood that students would report abusive behaviour to adults or friends or handle the risk themselves any more than in no-treatment conditions.

Table 2 Effect of ThinkUKnow training for children: knowledge, perceptions, behaviours, likelihood of reporting, and police legitimacy perceptions— treatment vs. control clusters X year schoolyear blocks

		Years 5 – 6 block (primary school)				Years 7 – 8 block (secondary school)			
		Treatment	Control	*t*	$p^{¥}$	Treatment	Control	*t*	$p^{¥}$
Newly gained knowledge (percent correct)	Online safety	80.31 (6.75) *	73.39 (7.42)	3.555	**< 0.001**	78.5 (5.81)	73.53 (5.77)	3.025	**0.004**
	Cyberbullying	73.23 (8.16)	65.84 (6.73)	3.586	**< 0.001**	79.31 (9.70)	77.39 (7.11)	0.793	0.432
	Image-based abuse	–	–	–	–	83.27 (7.35)	79.59 (9.55)	1.513	0.137
Risk perceptions (mean scores)	Certainty of apprehension (scales 1– 4)	3.122 (0.229)	3.046 (0.194)	1.312	0.195	3.109 (0.220)	3.041 (0.182)	-1.194	0.238
	Severity of punishment (scales 1–6)	4.668 (0.313)	4.662 (0.236)	0.085	0.932	4.593 (0.286)	4.536 (0.334)	0.641	0.524
Risky behaviours (mean scores)	Perpetration (scales 1–6)	5.441 (0.310)	5.370 (0.305)	0.844	0.403	5.387 (0.323)	5.351 (0.313)	0.396	0.694
	Protective behaviours (scales 1–6)	5.225 (0.223)	5.050 (0.245)	2.716	**0.009**	5.043 (0.289)	4.943 (0.390)	1.02	0.313
Likelihood of reporting (to:) (mean scores)	Adults (scales 1–7)	5.997 (0.335)	6.035 (0.265)	0.458	0.649	5.519 (515)	5.402 (0.338)	0.903	0.371
	Police (scales 1–7)	4.892 (0.370)	4.195 (0.580)	5.192	**< 0.001**	4.456 (0.456)	4.046 (0.382)	3.413	**< 0.001**
	Friends (scales 1–7)	4.419 (0.430)	4.601 (0.393)	1.605	0.115	5.282 (0.428)	5.348 (0.448)	0.526	0.601
	Handle oneself (scales 1–7)	3.487 (0.393)	3.470 (0.451)	0.147	0.884	4.444 (0.771)	4.616 (0.490)	0.936	0.354
Perceptions of police legitimacy (mean scores)	(Scales 1–5)	4.084 (0.157)	3.990 (0.142)	2.301	**0.026**	3.904 (0.251)	3.799 (0.260)	1.439	0.157

*Standard deviations in parentheses *N* = cluster size (years 5 – 6 block: treatment = 27, control = 26) (years 7 – 8 block: treatment = 24, control = 25); ¥Bonferroni-adjusted *p* values

Perceptions of police legitimacy

The mean treatment score was higher than the control group for both years 5 – 6 (treatment = 4.084, control = 3.99) and years 7 – 8 (treatment = 3.904, control = 3.799). However, statistically significant differences emerged for years 5 – 6 ($t = 2.301$, $p = 0.026$) but not in the years 7 – 8 block ($t = 1.439$, $p = 0.157$).

Discussion

The ThinkUKnow programme has been delivered in Australia for over 12 years. Beyond the potential utility of the training module in promoting online safety, most police services across Australia provide elaborate resources to deliver ThinkUKnow. The ACT police alone delivered more than 250 sessions to schools in 2018–2019 (ThinkUKnow 2019). Therefore, this study is of national interest from an education and a policing perspective.

We tested the effect of the training under controlled conditions. We observed variations between years 5 – 6 and 7 – 8 students (separately) in terms of multiple dimensions targeted by the training. Based on our findings, ThinkUKnow significantly improves students' knowledge of online safety, but knowledge related to cyberbullying improved for younger but not older students. Substantial effects on knowledge by ThinkUKnow are consistent with meta-analyses of interventions on cyber safety (Mishna et al., 2011) and school-based programmes to prevent child sexual abuse (Walsh et al., 2018). However, considering the specific measure for image-based abuse for older children did not suggest that the programme led to knowledge the students already had without exposure to the training materials.

Insofar as risk perceptions are concerned—in terms of the severity of punishment or likelihood of apprehension—we find no evidence to reject the null hypothesis of the no-treatment effect. It is possible that children may already have had the necessary tools to grasp the odds of getting caught or the punishment for cyber abuse, and therefore did not perceive these probabilities as more or less severe compared to no-treatment conditions. In contrast, the programme had a far more significant positive effect on young participants' stated likelihood of employing protective behaviours. This finding is consistent with Walsh et al. (2018) and Zych et al. (2019), who identified the increased use of protective behaviours following interventions to prevent child sexual abuse.

Furthermore, the programme did not lead to changes in the likelihood of reporting or handling cyber abuse alone, except when reporting abuse to the police. It may be an artificial effect, as students may have been affected more by the presence of the police officers as part of the programme than the content of the training module. This measurement bias should be more tightly managed in future tests of similar programmes delivered by the police.

Finally, it is essential to ensure that the delivery of ThinkUKnow does not reduce perceptions of policy legitimacy. Legitimacy is crucial, serving as a critical gateway to the willingness of (young) people to engage with and report incidents to the police. In both blocks, ThinkUKnow created a slight increase in students' perceptions of police legitimacy, especially among younger students. While we cannot

link reductions in actual cyber abuse to ThinkUKnow (as our instrument was not intended for this purpose), increased police legitimacy could indirectly lead to better cooperation with the police to prevent cyber abuse. This result would be consistent with findings from Mazerolle et al. (2021) that demonstrated how increased perceptions of police legitimacy reduce levels of self-reported anti-social behaviour.

Policy implications

Given the substantial gains in some but not other outcome measures, an overall conclusion is challenging to gauge. The least we can confidently say is that the programme does not cause a clinically meaningful backfiring effect across nearly all comparisons. However, it may be that our approach to estimating the treatment effect was too onerous. For a police programme intended to expose students to cyber abuse in a 90-min session, obtaining significant gains in terms of perceptions, attitudes, and behavioural aspects may be underpowered, especially for more mature students at years 7 – 8 levels. At the same time, the impact of ThinkUKnow on students' knowledge cannot be understated. Among year 7 – 8 students, the knowledge that cyberbullying is against the law increased by over 40%. If ThinkUKnow were delivered to the entire years 7 – 8 public school population in the area, more than 1300 additional students in years 7 – 8 would be made aware of this fact.

It is, nevertheless, evident from this study that ThinkUKnow can benefit from some changes in content and style. The lack of treatment effect on reporting should be managed more closely, as feeling safe to report cyber abuse to an adult or the police is a critical consequence of the training in field settings. Furthermore, modifications should also be considered in terms of new or emerging risks. One area which illustrates the need to prioritise certain elements over others relates to online safety and protective behaviours as these were generally more effective than deterrent elements. To this extent, it is worthwhile to consider whether sections relating to perpetration deterrence need to be reconsidered and overhauled.

Study limitations

As this study measured outcomes at a single point in time, it is not clear if any of the beneficial effects measured would be retained over time (see review in Sutherland et al., 2017). Furthermore, the post-intervention measurement focuses on perceptions and intentions for behaviour rather than actual variations in behaviour resulting from the intervention. Were a change in intention to commit cyber-abuse reported in the post-tests, it is unclear whether this would translate into actual changes in behaviour (see Antrobus et al., 2018; Ariel et al., 2019). Measuring an actual change in behaviour would need to go beyond the survey instrument design to include a metric for reporting infractions, honeypot experiments, or other relevant approaches.

In addition, there were no controls put in place or information collected regarding schools having received other cyber-safety programmes, such as those delivered under the auspices of the Children's e-Safety Commissioner (www.esafety.gov.au/). This issue was identified in the design stage, and an

active decision was taken to not employ any exclusion criterion for past exposure to cyber-safety programmes other than ThinkUKnow in the current calendar year. The RCT aimed to explore the effect of ThinkUKnow in as close to a real-world context as possible and estimate the treatment effect above and beyond any other existing programme. More attention should be considered in the future for interaction effects with other interventions.

Finally, our design could control for most of the known threats to internal validity, except history effects (see Campbell & Stanley 2015). Treatment participants completed the survey after the intervention, while control participants completed the survey before the intervention (and not after). Given operational restrictions, this design was optimal. However, it does not account for any possible biases associated with the timing of the measurement for the control group. While we assume that approximately 90 min create minimal confounding parameters on the validity of our causal estimates, this assumption is evidence-free. Future research should consider a more robust design with two parallel groups with simultaneous post-only measures whenever possible.

Conclusions

This study provides experimental evidence supporting the hypothesis that the ThinkUKnow programme for school children in Australia improves knowledge on cyber abuse. However, we find no consistent evidence of enhancing students' risk perceptions or decreasing the likelihood of involvement in risky behaviours online. Students exposed to the programme are more likely to report cyber abuse to the police, but not to other adults or friends. Participation in the programme increased perceptions of police legitimacy but primarily for younger rather than older students. While this study examined two cohorts (years $5-6$ and $7-8$), the results did not differ significantly between these two groups. This is perhaps due to these students being in the same age range. As such, future research should consider a more diverse age range. How the age of a student affects the efficacy of the treatment is a question that has not been answered by this research. More research, preferably with more diverse student populations, is needed.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11292-023-09551-3.

References

Abreu, R. L., & Kenny, M. C. (2018). Cyberbullying and LGBTQ youth: A systematic literature review and recommendations for prevention and intervention. *Journal of Child & Adolescent Trauma, 11*(1), 81–97.

Antrobus, E., Thompson, I., & Ariel, B. (2018). Procedural justice training for police recruits: Results of a randomized controlled trial. *Journal of Experimental Criminology, 15*(1), 1–25.

Ariel, B., Bland, M., & Sutherland, A. (2022). *Experimental designs*. Sage.

Ariel, B., Mitchell, R. J., Tankebe, J., Firpo, M. E., Fraiman, R., & Hyatt, J. M. (2020). Using wearable technology to increase police legitimacy in Uruguay: the case of body-worn cameras. *Law & Social Inquiry, 45*(1), 52–80.

Australian Communications and Media Authority (2013) *Like, post, share: Young Australians'* experience of social media. Retrieved from http://www.acma.gov.au/~/media/mediacomms/Report/pdf/LikepostshareYoungAustraliansexperienceofsocialmediaQuantitativeresearchreport.pdf

Australian Federal Police (2021) AFP Annual Report 2020-21. Australia: Australian Federal Police. https://www.afp.gov.au/what-we-do/crime-types/child-protection/thinkuknow

Bachman, R., & Schutt, R. (2017). *The practice of research in criminology and criminal justice* (6th ed.). Sage.

Brochado, S., Soares, S., & Fraga, S. (2017). A scoping review on studies of cyberbullying prevalence among adolescents. *Trauma, Violence & Abuse, 18*(5), 523–531.

Broll, R., & Huey, L. (2015). "Just being mean to somebody isn'ta police matter": Police perspectives on policing cyberbullying. *Journal of School Violence, 14*(2), 155–176.

Bryce, J. (2010). Online sexual exploitation of children and young people. In Y. Jewkes & M. Yar (Eds.), *Handbook of internet crime* (pp. 320–342). Routledge.

Bryce, J. (2017). ICTs and child sexual offending: Exploitation through indecent images. In M. R. McGuire & T. J. Holt (Eds.), *The Routledge handbook of technology, crime and justice* (pp. 96–112). Routledge.

Calvo-Morata, A., Alonso-Fernández, C., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2020). Serious games to prevent and detect bullying and cyberbullying: A systematic serious games and literature review. *Computers & Education, 157*, 103958.

Camerini, A. L., Marciano, L., Carrara, A., & Schulz, P. J. (2020). Cyberbullying perpetration and victimization among children and adolescents: A systematic review of longitudinal studies. *Telematics and Informatics, 49*, 101362.

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio books.

Cantone, E., Piras, A. P., Vellante, M., Preti, A., Daníelsdóttir, S., D'Aloja, E., ... & Bhugra, D. (2015). Interventions on bullying and cyberbullying in schools: A systematic review. *Clinical practice and epidemiology in mental health: CP & EMH, 11*(Suppl 1 M4), 58.

Davidson, J., Martellozzo, E., & Lorenz, M. (2009). Evaluation of CEOP ThinkUKnow internet safety programme and exploration of young people's internet safety knowledge. Other. *Centre for Abuse & Trauma Studies and Kingston University*. https://eprints.mdx.ac.uk/5976/

Diaz, M. J., Wolfersteig, W., Moreland, D., Yoder, G., Dustman, P., & Harthun, M. L. (2021). Teaching youth to resist abuse: Evaluation of a strengths-based child maltreatment curriculum for high school students. *Journal of Child & Adolescent Trauma, 14*(1), 141–149.

Donner, A., & Klar, N. (2002). Issues in meta-analysis of cluster randomized trials. *Statistics in Medicine, 21*, 2971–2980.

Evangelio, C., Rodriguez-Gonzalez, P., Fernandez-Rio, J., & Gonzalez-Villora, S. (2022). Cyberbullying in elementary and middle school students: A systematic review. *Computers & Education, 176*, 104356.

Evans, C. B., Fraser, M. W., & Cotter, K. L. (2014). The effectiveness of school-based bullying prevention programs: A systematic review. *Aggression and Violent Behavior, 19*(5), 532–544.

Farrington, D. P., & Ttofi, M. M. (2009). School-based programs to reduce bullying and victimization. *Campbell Systematic Reviews, 6*, 1–148.

Feinberg, T., & Robey, N. (2009). Cyberbullying: Intervention and prevention strategies. *National Association of School Psychologists, 38*(4), 22–24.

Fowler, F. J. (2013) *Survey Research Methods*, 5th ed., Sage Publications.

Gaffney, H., Farrington, D. P., Espelage, D. L., & Ttofi, M. M. (2018). 'Are cyberbullying intervention and prevention programs effective? *A Systematic and Meta-Analytical Review', Aggression and Violent Behavior, 45*, 134–153.

Hirtenlehner, H., Pauwels, L. J. R., & Mesko, G. (2013). Is the effect of perceived deterrence on juvenile offending contingent on the level of self-control? Results from three countries. *The British Journal of Criminology, 54*(1), 128–150.

Hutson, E., Kelly, S., & Militello, L. K. (2018). Systematic review of cyberbullying interventions for youth and parents with implications for evidence-based practice. *Worldviews on Evidence-Based Nursing, 15*(1), 72–79.

Kenny, M. C., Long, H., Billings, D., & Malik, F. (2022). School-based abuse prevention programming: Implementation of child safety matters with minority youth. *Child Abuse Review, 31*(3), e2742.

Klenowski, P. M., Bell, K. J., & Dodson, K. D. (2010). An empirical evaluation of juvenile awareness programs in the United States: Can juveniles be "scared straight"? *Journal of Offender Rehabilitation, 49*(4), 254–272.

Kowalski, R., Giumetti, G., Schroeder, A., & Lattanner, M. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychology Bulletin, 140*(4), 1073–1137.

Macaulay, P. J., Betts, L. R., Stiller, J., & Kellezi, B. (2018). Perceptions and responses towards cyberbullying: A systematic review of teachers in the education system. *Aggression and Violent Behavior, 43*, 1–12.

Mazerolle, L., Bennett, S., Davis, J., Sargeant, E., & Manning, M. (2013). Procedural justice and police legitimacy: A systematic review of the research evidence. *Journal of Experimental Criminology, 9*(3), 245–274.

Mazerolle, L., Antrobus, E., Cardwell, S. M., Piquero, A. R., & Bennett, S. (2021). Harmonizing legal socialization to reduce antisocial behavior: Results from a randomized field trial of truanting young people. *Justice Quarterly, 38*(1),152–179.

Mishna, F., Cook, C., Saini, M., Wu, M. J., & MacFadden, R. (2011). Interventions to prevent and reduce cyber abuse of youth: A systematic review. *Research on Social Work Practice, 21*(1), 5–14.

Murphy, K., & Cherney, A. (2011). Fostering cooperation with the police: How do ethnic minorities in Australia respond to procedural justice-based policing? *Australian and New Zealand Journal of Criminology, 44*(2), 235–257.

Nyberg, A., Ferm, U., & Bornman, J. (2021). School-based abuse prevention programs for children with disabilities: A qualitative study of components and methods. *Australasian Journal of Special and Inclusive Education, 45*(2), 252–267.

Olweus, D., & Limber, S. P. (2018). Some problems with cyberbullying research. *Current Opinion in Psychology, 19*, 139–143.

Park, M. A., Golden, K. J., Vizcaino-Vickers, S., Jidong, D., & Raj, S. (2021). Sociocultural values, attitudes and risk factors associated with adolescent cyberbullying in East Asia: A systematic review. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 15*(1). https://doi.org/10.5817/CP2021-1-5

Pennell, D., Campbell, M., & Tangen, D. (2022). The education and the legal system: Inter-systemic collaborations identified by Australian schools to more effectively reduce cyberbullying. *Preventing School Failure: Alternative Education for Children and Youth, 66*(2), 175–185.

Perry, C. L., Komro, K. A., Veblen-Mortenson, S., Bosma, L. M., Farbakhsh, K., Munson, K. A., … & Lytle, L. A. (2003). 'A randomized controlled trial of the middle and junior high school DARE and DARE Plus programs', *Archives of Pediatrics & Adolescent Medicine, 157*(2): 178–184.

Polanin, J. R., Espelage, D. L., Grotpeter, J. K., Ingram, K., Michaelson, L., Spinney, E., ... & Robinson, L. (2022). A systematic review and meta-analysis of interventions to decrease cyberbullying perpetration and victimization. *Prevention Science*, 23(3), 439–454.

Radebe, F., & Kyobe, M. (2021). The response of social crime prevention police to cyberbullying perpetrated by youth in rural areas of South Africa. *International Journal of Environmental Research and Public Health, 18*(24), 13421.

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers* (2nd ed.). Blackwell.

Selkie, E. M., Fales, J. L., & Moreno, M. A. (2016). Cyberbullying prevalence among US middle and high school–aged adolescents: A systematic review and quality assessment. *Journal of Adolescent Health, 58*(2), 125–133.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, Ma, Houghton Mifflin.

Shaikh, F. B., Rehman, M., & Amin, A. (2020). Cyberbullying: A systematic literature review to identify the factors impelling university students towards cyberbullying. *IEEE Access, 8*, 148031–148051.

Shaw, M. (2004) *Police, schools, and crime prevention: A preliminary review of current practices*, Montreal, Canada: International Crime Prevention Centre. https://www.publicsafety.gc.ca/lbrr/archives/cnmcs-plcng/cn3768-eng.pdf

Solehati, T., Fikri, A. R., Kosasih, C. E., Hermayanti, Y., & Mediani, H. S. (2022). The current preventing of child sexual abuse: A scoping review. *Social Sciences, 11*(11), 508.

Sunshine, J., & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law & Society Review, 37*(3), 513–548.

Sutherland, A., Ariel, B., Farrar, W., & De Anda, R. (2017). Post-experimental follow-ups—fade-out versus persistence effects: The Rialto police body-worn camera experiment four years on. *Journal of Criminal Justice, 53*, 110–116.

Tankebe, J. (2009). Public cooperation with the police in Ghana: Does procedural fairness matter? *Criminology, 47*, 1265–1293.

Tanrikulu, I. (2018). Cyberbullying prevention and intervention programs in schools: A systematic review. *School Psychology International, 39*(1), 74–91.

ThinkUKnow Australia (2019) *Corporate Report 2018–19*, Australia: ThinkUKnow Australia. https://www.thinkuknow.org.au/sites/default/files/2021-02/ThinkUKnow%20Corporate%20Report%202018-19.pdf

Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology, 7*, 27–56.

Vandebosch, H., Beirens, L., D'Haese, W., Wegge, D., & Pabian, S. (2012). Police actions with regard to cyberbullying: The Belgian case. *Psicothema, 24*(4), 646–652.

Walsh, K., Zwi, K., Woolfenden, S., & Shlonsky, A. (2018). School-based education programs for the prevention of child sexual abuse: A Cochrane systematic review and meta-analysis. *Research on social work practice, 28*(1), 33–55.

Weisburd, D., & Hinkle, J. C. (2012) 'The importance of randomized experiments in evaluating crime prevention', in B. C. Welsh and D. P. Farrington (eds) *The Oxford handbook of crime prevention*, Oxford: OUP.

White, C., Shanley, D. C., Zimmer-Gembeck, M. J., Walsh, K., Hawkins, R., Lines, K., & Webb, H. (2018). Promoting young children's interpersonal safety knowledge, intentions, confidence, and protective behavior skills: Outcomes of a randomized controlled trial. *Child Abuse & Neglect, 82*, 144–155.

White, C., Shanley, D. C., Zimmer-Gembeck, M. J., Walsh, K., Hawkins, R., & Lines, K. (2019). Outcomes of in situ training for disclosure as a standalone and a booster to a child protective behaviors education program. *Child Maltreatment, 24*(2), 193–202.

Wilson, C. J., Deane, F. P., Ciarrochi, J. V., & Rickwood, D. (2005). Measuring help seeking intentions: Properties of the General Help Seeking Questionnaire. *Canadian Journal of Counselling, 39*(1), 15–28.

Zhu, C., Huang, S., Evans, R., & Zhang, W. (2021). Cyberbullying among adolescents and children: A comprehensive review of the global situation, risk factors, and preventive measures. *Frontiers in Public Health, 9*, 634909.

Zych, I., Ortega-Ruiz, R., & Del Rey, R. (2015). Systematic review of theoretical studies on bullying and cyberbullying: Facts, knowledge, prevention, and intervention. *Aggression and Violent Behavior, 23*, 1–21.

Zych, I., Farrington, D. P., & Ttofi, M. M. (2019). Protective factors against bullying and cyberbullying: A systematic review of meta-analyses. *Aggression and Violent Behavior, 45*, 4–19.

Tony Alderman is the Manager Policing Development and Innovation at the Australian Federal Police. Tony has a Master of Studies in Applied Criminology from the University of Cambridge, and sits on the Executive Committee of the Australia New Zealand Society of Evidence Based Policing.

Barak Ariel is a Professor of Experimental Criminology at the Institute of Criminology, University of Cambridge and an Associate Professor of Criminology at the Institute of Criminology, Faculty of Law, The Hebrew University of Jerusalem. He is the immediate past Chair of the Division of Experimental Criminology of the American Society of Criminology.

Vincent Harinam is a data scientist and law enforcement consultant.