Development and initial validation of a school self-assessment tool to measure the extent to which schools are prepared to prevent and respond to cyberbullying and other challenges to the school social environment

ORIGINAL ARTICLE

Development and Initial Validation of a School Self-assessment Tool to Measure the Extent to Which Schools are Prepared to Prevent and Respond to Cyberbullying and Other Challenges to the School Social Environment

Nina Van Dyke[1] · Fiona MacDonald[2] · Rachael Bajayo[3] · Chinh Duc Nguyen[1,4,5] · Cynthia Leung[1] · Sophie Francis[3]

Abstract

Bullying, which includes cyberbullying and face-to-face bullying, is amongst the most pervasive threats to the wellbeing of children and young people. As social environments, schools are at the forefront of managing bullying behaviours. The rapidly changing and complex nature of bullying requires schools to put in place and maintain systems to prepare for and respond to such activities, and to continually test and refine these systems to ensure optimal performance. Despite this clear need, there is a lack of school level, self-assessment tools that enable schools to assess and measure their preparedness to deal with bullying and related disruptive activities. The aim of this paper was to describe the development, and reliability and validity testing, of such a tool—the School Self-Assessment Tool. The result is a 22-item, evidence-based, reliable, and validated instrument, situated within socio-ecological theory, and drawing on theories of behaviourism, social learning theory, prevention science, and systems change in the school climate literature. Schools may use the SSAT-22 in a number of ways depending on their needs, including to monitor progress, examine areas of strengths or challenges, and/or assist in collaborative efforts with other schools.

Keywords Bullying · Cyberbullying · Self-assessment · Schools · Tool · Measurement

Introduction

Cyberbullying and traditional, face-to-face bullying (here-inafter referred to as "bullying" unless otherwise specified) are amongst the most pervasive threats to the wellbeing of school-aged children and young people (United Nations, 2021). As social environments, schools are at the forefront of managing bullying behaviours, with incidents occurring either at school, or outside of school via school networks. As a result, schools are regularly asked to respond to anti-social behaviours, related incidents, and other challenges to their social environment (Gaffney et al., 2019; Lindstrom et al., 2019; Pennell et al., 2020). COVID-19 and the shift to online teaching in Australia and around the globe has increased the online activity of school-aged children and young people. While evidence is still emerging, the incidence and nature of cyberbullying continue to have negative impacts for children and young people, placing additional pressure on schools to respond (eSafety Commissioner, 2021; Gupta & Jawanda, 2020). The rapidly changing and complex nature of bullying requires that schools put in place and maintain systems to prepare for and respond to social disruptions, testing and refining these systems to ensure optimal performance. Despite this clear need, there is a lack of school level, self-assessment tools that enable schools to assess and measure their preparedness to deal with bullying and related disruptive activities (Pennell et al., 2022).

✉ Nina Van Dyke
  nina.vandyke@vu.edu.au

1 Mitchell Institute, Victoria University, Melbourne, Australia

2 Institute for Sustainable Industries and Liveable Cities, Victoria University, Melbourne, Australia

3 Alannah & Madeline Foundation, Melbourne, Australia

4 Melbourne Graduate School of Education, Melbourne, Australia

5 Dong Nai University of Technology, Bien Hoa, Vietnam

The Alannah & Madeline Foundation's eSmart Schools Framework (the Framework) has been delivered in schools since 2011. The Framework is a long-term change program that works with schools to build supportive and connected school environments that can identify, address, prepare for, and respond to bullying and other social challenges that threaten their school environment. To assess the effectiveness of the program, the Alannah & Madeline Foundation and Victoria University researchers worked in partnership to develop an evidence-based, systems-level, self-assessment tool for schools (the SSAT).

In developing the SSAT, we acknowledged the complex nature of bullying and focused on the expectations of schools to create supportive social environments. In addition, drawing on Social Ecological Theory, we incorporated the understanding that bullying is embedded in larger social contexts (Bronfenbrenner, 1977; Gaffney et al., 2021; Pennell et al., 2020). We acknowledged that there are differences in school environments and the way they implement programs that impact the effectiveness of anti-bullying programs (Gaffney et al., 2019). Finally, we also drew on theories of behaviourism, social learning theory, prevention science, and systems change in the school climate literature.

Background and Context

Bullying is a global problem. It impacts all life stages but is widely acknowledged as having a significantly negative impact on children and young people under the age of 18—the first generation to grow up "surrounded by digital devices" (eSafety Commissioner, 2021, p. 3). Incidents of bullying can impact children and young people's mental health, including their levels of stress, anxiety, and depression, and can negatively influence their self-esteem and perceived degree of social connectedness. In some instances, bullying can lead to ongoing emotional issues, self-harm, and suicidal ideation and related behaviours (Ford et al., 2017). The issue of bullying is not confined to teenagers, with parents of children in primary years also reporting incidents of both traditional bullying and cyberbullying (Tanrikulu, 2018). Moreover, cyberbullying takes place across a variety of online mediums, with one study finding that over 19% of children and young people reported bullying on social media sites and apps, 11% through text messaging, and nearly 8% through online video games (eSafety Commissioner, 2021).

The onset of COVID-19 increased concerns about incidents of cyberbullying for children and young people (Gupta & Jawanda, 2020). Education systems have been severely disrupted, and many nations switched to online delivery for periods of time between 2020 and 2022, increasing the online activity of children and young people (United Nations, 2021). In Australia, the focus of this paper, as many as 44% of Australian teenagers reported having a negative online experience in the first six months of COVID-19 (March–August 2020) (eSafety Commissioner, 2021). Such online incidents included contact from someone they did not know, receiving inappropriate or unwanted content, and being deliberately excluded from their social group or events organised within their social group (eSafety Commissioner, 2021).

While traditional bullying and cyberbullying are often reported separately, the two overlap (eSafety Commissioner, 2021; Hu et al., 2021). It has been argued that traditional bullying consists of three key elements: (1) "purposeful unwanted negative (aggressive) behaviour that (2) typically implies a pattern of behaviour that is repeated, and (3) occurs in an interpersonal relationship characterized by an imbalance of power or strength, favouring the perpetrator(s)" (Olweus & Limber, 2018, p. 139). Cyberbullying has been defined as encompassing these same three elements, but with specific nuances around the imbalance of power, which can include "technological know-how, relative anonymity, social status, number of friends, or marginalised group position" (Smith et al., 2012, p. 36). Repetition, which includes the length of time that images and messages remain online, should be used as a "subsidiary criterion" but not "an essential criterion" for cyberbullying (Smith et al., 2012, p. 37).

Systems-level Approaches to Bullying

Schools are called upon to play a significant role in addressing bullying, with expectations that they understand the offline and online behaviours of their students and remain current with new online platforms (Gaffney et al., 2019; Lindstrom et al., 2019; Pennell et al., 2020). While schools—and the staff therein—are well placed to educate children and young people about bullying and other anti-social behaviours (Sivaraman et al., 2019; Tanrikulu, 2018), their role in managing this behaviour is complicated by criminal and legal requirements, expectations of parents, their primary role as educators, and the complex and rapidly changing nature of cyberbullying (Gaffney et al., 2019). Current research demonstrates that policies, targeted programs, and legal responses implemented in isolation cannot adequately address bullying (Pennell et al., 2022; Rapee et al., 2020; Sabia & Bass, 2017). As in many nations, Australian schools have developed policies, programs, and strategies to address traditional bullying and cyberbullying (Gaffney et al., 2019, 2021; Pennell et al., 2022; Sabia & Bass, 2017), largely at the behest of state level government departments. In some instances, schools have adopted a whole-school approach, drawing on a socio-ecological model that acknowledges the complex and interrelated systems at play (Chalmers et al., 2016; Gaffney et al., 2019, 2021; Sivaraman et al., 2019). A whole-school approach fits with school level actions and responses to violence and anti-social behaviours more broadly (Chalmers et al., 2016).

The assessment of targeted programs is critical to measuring the extent to which they are effective and "there have been many previous attempts to establish what works in bullying intervention and prevention" (Gaffney et al., 2019, p. 16). There are, to our knowledge, no self-assessment tools for schools to measure their progress towards being fully prepared to prevent and respond to bullying and other anti-social behaviours amongst their school community, although Gaffney et al.'s recommendation that "a pre-intervention survey to explore the specific manifestations of bullying in their respective schools" speaks to the importance of schools evaluating their preparedness (2019, p. 27).

Theoretical Frameworks

Schools are successfully reducing bullying by developing programs based on socio-ecological system theory (Allen, 2010; Boyd & Lawes, 2018; Gaffney et al., 2021; Hawkins et al., 2021; Hornby, 2016; Njelesani et al., 2020). Introduced by Bronfenbrenner (1977), socio-ecological theory contends that individuals are situated within a range of complex, interrelated system levels that influence their development. While the theory was developed in the twentieth century, schools are drawing on socio-ecological theory in the 2020s as they respond to bullying at a systems level. In doing so, they are creating a whole-school practice, such as establishing reporting systems, that engages a range of components and individuals in prevention and intervention activities (Gaffney et al., 2019, 2021; Sivaraman et al., 2019). This approach acknowledges that whole-school approaches should be augmented with individual components, such as focusing on the perpetrators and victims, with the two employed in concert (Rapee et al., 2020). Cross et al. (2018) argue that whole-school approaches to bullying include socio-ecological strategies delivered in classrooms, schools, and to families. As such, whole-school approaches have been seen to be more successful than single-component approaches, such as traditional disciplinary actions or classroom curriculum learning (Cross et al., 2018). Gaffney et al.'s (2021) analysis of effective intervention components found that:

The presence of a whole-school approach, classroom rules, information for parents, formal peer involvement, co-operative group work, and CBT approaches resulted in larger mean effect sizes in comparison to studies where these components were absent (2021, p. 51).

In whole-school approaches to bullying, the relationship between systems is described as "nested, reciprocal, and dynamically interacting" (Pennell et al., 2020, p. 286). Bronfenbrenner (1977) described systems as *micro-, meso-, exo-,* and *macrosystems*, depending on how individuals are impacted by each one. All levels of the socio-ecological framework are implemented with anti-bullying programs at a school level (Pennell et al., 2020). The *macrosystem* is described as societal and cultural aspects that impact bullying, the *exosystem* as school practice influenced by communities, the *mesosystem* as school actions of prevention and intervention, and the *microsystem* as the prevention and intervention activities of stakeholders inside schools. These circles are arranged from the outermost (macrosystem) to the innermost student bullying behaviour, placing issues of bullying at the centre of schools' interrelated systems.

Initial Development of the SSAT

To develop the SSAT, a review was conducted of the Australian and international literature on system-level approaches to preparing for, preventing, and responding to bullying and cyberbullying and other challenges or disruptions to the social environment. This review revealed the lack of school level, self-assessment tools (Pennell et al., 2020). In addition, a review was conducted of Australian government policies and guidelines for schools to address bullying and cyberbullying. Based on these reviews, five Focus Areas were identified, which then informed the development of a 40-item instrument. The five Focus Areas were as follows: Data, Gateway Behaviours, Reporting, Response, and School social connectedness/climate.

In taking this approach, we acknowledge the synergies between the five Focus Areas and Gaffney et al.'s (2021) intervention components for effective school-based programs: school level, classroom level, teacher level, parent/guardian level, peer level, individual level, and intervention specific. Our SSAT and Gaffney et al.'s (2021) meta-analysis stress the importance of including "all actors" (2021) and across "multiple levels" (2019) within the school environment. Our five Focus Areas included all actors in the school environment, alongside the different levels of the school environment, from school to classroom level, the teacher, parent/guardian level, to the peer and individual level (Gaffney et al., 2021). The difference between the five Focus Areas of the SSAT and Gaffney et al.'s (2021) intervention components, we argue, was our focus on systems level preparedness to deal with bullying and related disruptive activities in contrast to Gaffney et al.'s evaluation of programs.

In the SSAT, the Data Focus Area included issues around the collection and analysis of data relevant to traditional bullying, cyberbullying, and social relationships; the priority placed by schools on the collection and analysis of such data; and the conduct of regular and integrated data reviews. Gateway Behaviours were defined as follows: "lower-risk behaviours, online or offline, which are used to show contempt and dominance. When left unchecked, these behaviours can escalate into conflict and bullying. Examples include posting embarrassing photos online, ignoring, name calling, whispering about people in front of them, eye rolling" (modified

from Englander, 2013). The Reporting Focus Area included the ease and clarity for students, staff, and parents around the processes for reporting incidents or concerns, as well as levels of trust that incidents and concerns are reported, acted upon, and resolved. The Response Focus Area included implementation of the school's response plan, meetings with the victim(s) and perpetrator(s) and their parents or carers, and follow-up. The School Social Connectedness/Climate Focus Area included issues around the collection of data on social relationships within the school and level of agreement by school staff on the importance of creating and maintaining a positive school environment.

The instrument also included two "global rating questions" asked at both the beginning and end of the SSAT. These questions asked: (a) Overall, how well placed would you say your school is to *prevent* a cyber bullying or other type of bullying incident? (b) Overall, how well placed would you say your school is to *respond to* a cyber bullying or other type of bullying incident? Following the first set of global rating questions, participants listened to a short audio vignette depicting a school bullying situation that participants were to refer to when responding to the 40 items. After the second set of global rating questions, three questions asked participants to evaluate the vignette on level of engagement, plausibility, and helpfulness in responding to the self-assessment questions. These questions asked: (a) How engaging did you find the audio scenario? (b) How plausible did you find the audio scenario? (c) To what extent did the audio scenario help you answer the questions above? Along with the SSAT, schools also completed a Gap Analysis tool. The Gap Analysis tool focused on identifying gaps at a system level and measuring the extent to which schools addressed these gaps over time. The Gap Analysis tool consisted of 21 yes/no items.

The aim of this paper was to describe the development, and reliability and validity testing, of a school self-assessment tool designed to help schools identify how prepared their school system is to prevent and respond to bullying and other disruptive social challenges, and where there is room for system-level improvements.

Methods

Ethics

The design of the SSAT, and its administration for the purposes of evaluating the effectiveness of the Framework, was approved by the Victoria University Human Research Ethics Committee [HRE20-004]. T; the Strategic Research Unit, Department of Education and Training Victoria (RISEC application), and the Catholic Dioceses of Melbourne, Bendigo, Sale, and Sandhurst all consented to the SSAT being administered in their schools. Independent (i.e. non-government) schools were approached individually.

Participants and Procedures

Participants consisted of Principals and/or eSmart Coordinators at Victorian schools that were part of the eSmart Schools Expansion project—a project funded by the Victorian Department of Education and Training (DET) designed to provide the Framework to all interested Victorian schools. All 609 expansion schools were invited by email to participate in the evaluation. A total of 36 schools participated. Participants provided verbal informed consent. Given the impact of COVID on schools' operations, capacity, and resources, it was not surprising that participation was lower than originally projected.

Participating schools included all three Australian school sectors (government; Catholic; independent) and all three school types (primary; secondary; combined). Table 1 shows participation by type, sector, socio-economic status, Australian Digital Inclusion Index (ADII), and location (metropolitan vs non-metropolitan and DET region). As can be seen, compared with all expansion schools, participating schools were more likely to be government schools, primary schools, higher SES, above average ADII, and located in metropolitan regions.

Table 1 Demographics of participating schools

	Sample (n)	% of expansion schools
School type		
Primary	25	4.1%
Secondary	5	0.8%
Combined	6	1.0%
School sector		
Government	22	3.6%
Catholic	11	1.8%
Independent	3	0.5%
Socio-economic status		
High	23	3.7%
Medium	8	1.3%
Low	5	0.8%
ADII		
Above avg	24	3.9%
Below avg	9	1.4%
(Unavailable)	3	0.5%
Location		
Metro	20	3.2%
Non-metro	16	2.6%
Location–DET regions		
North-eastern	10	1.6%
North-western	8	1.3%
South-eastern	10	1.6%
South-western	8	1.3%

Measures

Participants were asked to indicate their endorsement of each of the 40 items on a 4-point scale (0: not true; 1: a little bit true, 2: mostly true; 3: completely true). Table 2 shows the five Focus Areas, number of items within each, and indicative questions.

Analysis

Reliability and validity analyses were conducted using IBM SPSS Statistics (Version 26). Reliability, which was examined in terms of internal consistency, was determined by running Cronbach alpha tests for each of the five Focus Areas and for all 40 items together. A sample size of 30 is considered adequate when the minimum effect size is 0.7 (Bujang et al., 2018). In addition, both convergent and discriminant validity (also called divergent validity) were examined. Validity estimates have been found to be stable for 75% of scales in samples of $n = 40$ (Hobart et al., 2012). To test for discriminant validity, correlations between Focus Areas and Vignette questions were examined; it was hypothesised that these correlations would be small and not reach statistical significance. To test for convergent validity (Anastasi & Urbina, 1997), correlations between each of the Focus Areas and: (a) the post-evaluation global rating questions, and (b) the Gap Analysis, were examined. Hypotheses were that both correlations would be positive and reach statistical significance. As a result of these analyses (see "Results" section), in combination with input from the Alannah & Madeline Foundation, the 40-item SSAT was reduced to 24 items. Internal reliability and construct validity analyses were conducted with the revised, 24-item instrument.

Following this, Rasch analysis was conducted using Bond and Fox Steps to examine the measurement properties of the SSAT-24, including unidimensionality, item difficulty, and category functioning, using the steps laid out by Bond and Fox (2007). Rasch analysis was deemed appropriate as the sample size was larger than 30 (Linacre, 1994). Undimensionality was investigated using the infit and outfit statistics, point measure correlations, and the principal component analysis (PCA) of the residuals (Bond & Fox, 2007). Although some analyses examined the five Focus Areas, our primary interest was in the instrument as a whole, which we contend measures the underlying construct, system-level preparedness to prevent and respond to bullying. According to McCreary et al. (2013), the criteria for unidimensionality in PCA are: (i) the variance explained by measures is 40% or more, (ii) the variance explained by the first principal component of the residuals is 15% or less, and (iii) the ratio of variance in measures to variance in the first principal component of the residuals is 3:1 or more.

To determine whether the tool items targeted the ability level of the participants, the Wright map was used (Bond & Fox, 2007). In this study, "higher ability" equates to being better prepared to prevent and respond to bullying. To establish whether participants meaningfully differentiated between the categories in the rating scale, category frequencies, monotonicity of average measures, and step calibrations were examined (Bond & Fox, 2007). As a result of these analyses, two additional items were dropped and the

Table 2 Indicative items in Initial tool (SSAT-40)

Focus area	Number of items	Indicative items
Data	18	*In order to understand the social relationships within the school, the following types of high-quality data are collected and analysed: feedback from focus groups and surveys with **students**, which includes student perspectives on what is considered most helpful in reducing cyber bullying and bullying.* *The collection of data to understand the social relationships within the school is viewed as a **crucial activity** by the Principal, school staff, students, and parents*
Gateway behaviours	4	*The identification and reporting of gateway behaviours is viewed as a **crucial** activity for the school by the Principal, all school staff, students, and parents.* *School staff are trained in, and consistently practice, identifying both overt and covert gateway behaviours, including online behaviours.*
Reporting	7	*Students, staff and parents know which staff member(s) within the school have been assigned responsibility to receive reports of cyber bullying and bullying.* *Students believe that their reports will be acted on and positively resolved.*
Response	7	*The school's response plan, which includes a response team with designated roles and responsibilities, is immediately enacted when cyber bullying or bullying occur.* *A designated staff member(s) meets with the **victim(s)** of the cyber bullying or bullying and their parents/carers as soon as feasible to provide: comfort, support such as counselling, and referral to additional assistance as needed such as mental health support*
School climate	4	*All school staff agree on the imperative for creating and maintaining a positive school environment.* *Data gathered on social relationships within the school includes measurement of the following themes: data on relationships between students and teachers/staff, including levels of trust, support, empathy, and kindness.*

4-item response scale was reduced to a 3-item scale, with "not true" and "a little bit true" combined. Finally, internal reliability and construct validity analyses were performed with the final, 22-item instrument.

Results

Internal Reliability Testing of the Initial, 40-item Instrument

Internal reliability testing of each of the five Focus Areas and the SSAT-40 as a whole using Cronbach's alpha demonstrated good to excellent internal reliability (DeVellis, 2012; George & Mallery, 2003) (SSAT-40: $n = 40$; $\alpha = 0.96$; Data: $n = 18$; $\alpha = 0.96$; Gateway Behaviours: $n = 4$; $\alpha = 0.84$; Response: $n = 7$; $\alpha = 0.86$; Reporting: $n = 7$; $\alpha = 0.85$; Social Culture: $n = 4$; $\alpha = 0.83$). An examination of individual items for very high mean values (> 3.5 on the 4-point scale), coupled with small standard deviations (< 0.70), low corrected–total item correlation scores (< 0.4), and increase in Cronbach's alpha if item deleted, as well as a primary components analysis, suggested that 16 items be dropped.

Internal reliability testing and validity testing was then conducted with the revised, SSAT-24 tool.

Reliability and Validity Testing of the Revised, 24-item Instrument

Internal reliability testing of each of the five Focus Areas and the SSAT-24 as a whole using Cronbach's alpha demonstrated good to excellent internal reliability (DeVellis, 2012; George & Mallery, 2003) (SSAT-24: $n = 24$; $\alpha = 0.96$; Data: $n = 9$; $\alpha = 0.95$; Gateway Behaviours: $n = 4$; $\alpha = 0.83$; Reporting: $n = 4$; $\alpha = 0.85$; Response: $n = 4$; $\alpha = 0.87$; Social Culture: $n = 3$; $\alpha = 0.91$). Discriminant validity was demonstrated by the small to moderate correlations between each Focus Area and the vignette questions (Cohen, 1988), none of which reached statistical significance (Data: $r(7) = -0.151$, $p = 0.410$; Gateway Behaviours: $r(2) = -0.005$, $p = 0.980$; Reporting: $r(2) = 0.221$, $p = 0.224$; Response: $r(3) = 0.043$, $p = 0.816$; School Climate: $r(1) = -0.209$, $p = 0.250$). Convergent validity was demonstrated by moderate to large (Cohen, 1988) positive correlations between each Focus Area and the second global evaluation rating (Data: $r(7) = 0.421$, $p = 0.010$; Gateway Behaviours: $r(2) = 0.546$, $p = 0.001$; Reporting: $r(2) = 0.475$, $p = 0.003$; Response: $r(2) = 0.663$, $p < 0.001$; School Climate: $r(1) = 0.472$, $p = 0.004$), as well as small to moderate positive correlations between each Focus Area and the Gap Analysis scale (Data: $r(7) = 0.251$, $p = 0.140$; Gateway Behaviours: $r(2) = 0.317$, $p = 0.060$; Reporting: $r(1) = 0.319$, $p = 0.058$; Response: $r(3) = 0.323$, $p = 0.055$; School Climate: $r(1) = 0.462$, $p = 0.005$).

Rasch Analysis of the SSAT-24

To test the unidimensionality of the SSAT-24 scale items, fit statistics were examined. The infit and outfit statistics of all items were within the suggested range of 0.6 to 1.4 except item Data 4.5. The infit and outfit statistics are shown in Table 3. All point measure correlations were positive. Unidimensionality was further examined using PCA of the residuals that remained after the linear Rasch measure had been extracted. The variance explained by measures was 73.4% and the variance explained by the first principal component of the residuals was 6.7%. The ratio of variance in measures to variance in the first principal component of the residuals was 10.96:1. The PCA results therefore supported the unidimensionality of the items.

We next tested whether the categories of the 4-point Likert scale measures were performing as expected. The number of responses met the minimum requirement of 10 per category (Bond & Fox, 2007). The average measures were ordered and increased monotonically from -0.93 logits for category 0 (*Not true*) to 2.16 for category 3 (*Completely true*). The infit and outfit mean squares of all categories were within the range of 0.6 to 1.4.

Table 3 Infit and outfit statistics of 24- and 22-item versions

	24-item version (4 categories)		22-item version (3 categories)	
	Infit	Outfit	Infit	Outfit
Data 2.1	1.11	1.01	1.34	1.35
Data 2.5.1	1.03	0.96	1.00	0.98
Data 2.5.3	0.84	0.81	1.00	0.98
Data 2.6	1.05	1.07	1.02	0.97
Data 4.1	1.39	1.30	1.33	1.31
Data 4.3	1.15	1.11	1.24	1.23
Data 4.5	0.44	0.44	Deleted	
Data 4.6	1.02	0.95	0.85	0.82
GB 1	1.21	1.07	1.03	0.95
GB 2	0.95	1.19	0.82	0.99
GB 3	0.86	0.91	0.97	0.95
GB 4	0.74	0.71	0.74	0.74
Report 2	0.60	0.63	Deleted	
Report 3	0.65	0.78	0.41	0.39
Report 5	0.76	0.75	0.79	0.76
Report 7	1.22	1.16	1.27	1.31
Resp 2.1	1.15	1.14	1.13	1.18
Resp 3	1.01	1.01	0.83	0.78
Resp 4	0.94	0.92	1.02	1.03
Resp 5	1.07	0.95	1.07	0.95
Resp 6	1.17	1.13	0.86	0.83
SC 2.1	1.37	1.27	1.31	1.25
SC 2.2	1.24	1.10	1.05	0.97
SC 2.3	0.85	0.79	0.96	0.93

However, for step calibrations, the thresholds of categories 1 and 2 were less than 1.4 logits apart, suggesting that categories needed revisions. The threshold probability curve is shown in Fig. 1.

To test the targeting of the items, the Wright map was examined. As can be seen in Fig. 2, there were insufficient items to target respondents with higher abilities (i.e., those schools better prepared to prevent and respond to bullying). The items ranged from "very easy" (item 20 – Resp_5_F: *As soon as possible, the school follows up with parents/carers of the victim to let them know what actions they are taking now and their plan for the future.*) to "medium" (items 10 - GB2: *School staff are trained in, and consistently practice, identifying both overt and covert gateway behaviours;* and item 12: *Students immediately and consistently report any gateway behaviours observed*). The average infit was 0.97 and the average outfit was 0.96. The person reliability and separation were 0.93 and 3.52. The item reliability and separation were 0.86 and 2.44.

To revise the scale, the middle two categories of the 4-point scale were collapsed to form a 3-point scale: *Not true, A little bit true/Mostly true, Completely true*. With the 3-point scale, the infit and outfit statistics of items Data 4.5, report 2 and report 3 were outside the recommended range. Nevertheless, due to the limited number of "Report" items, it was decided to drop only Report 2 and Data 4.5 and keep Report 3 (refer to Table 3 for infit/outfit statistics). After dropping these two items, all point measure correlations were positive but the infit/outfit statistics of Report 3 were still outside the recommended range. PCA results indicated that the variance explained by measures was 59.7% and the variance explained by the first principal component of the residuals was 9.7%. The ratio of variance in measures to variance in the first principal component of the residuals was 6.15:1. In terms of category functioning of the 22-item scale, the number of responses for each category was above the minimum requirement of 10. The average measures were ordered and increased monotonically from -1.04 logits for category 0 ("Not true") to 2.38 for category 2 ("Completely true"). The infit and outfit mean squares of all categories were within the range of 0.6 to 1.4. For step calibration, the thresholds of categories were more than 1.4 logits apart (Fig. 3).

The average infit was 0.99 and the average outfit was 0.98. The person reliability and separation were 0.88 and 2.68. The item reliability and separation were 0.82 and 2.13. In terms of targeting of items, there continued to be insufficient items to target respondents with high ability (Fig. 4).

Reliability and validity testing was then conducted on the final, 22-item instrument.

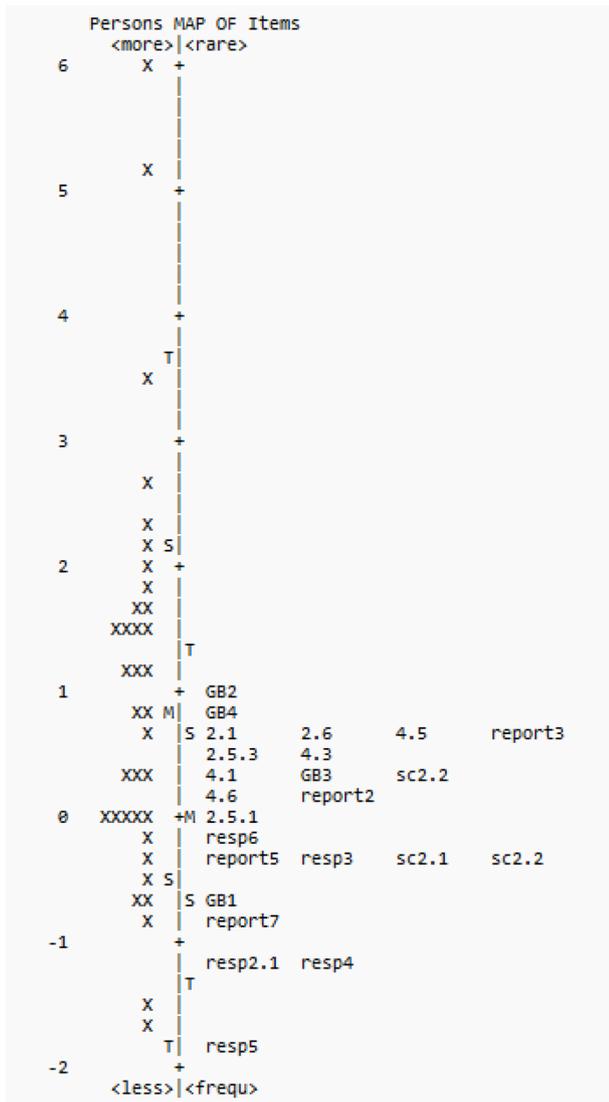**Fig. 1** Category probabilities (4 categories)
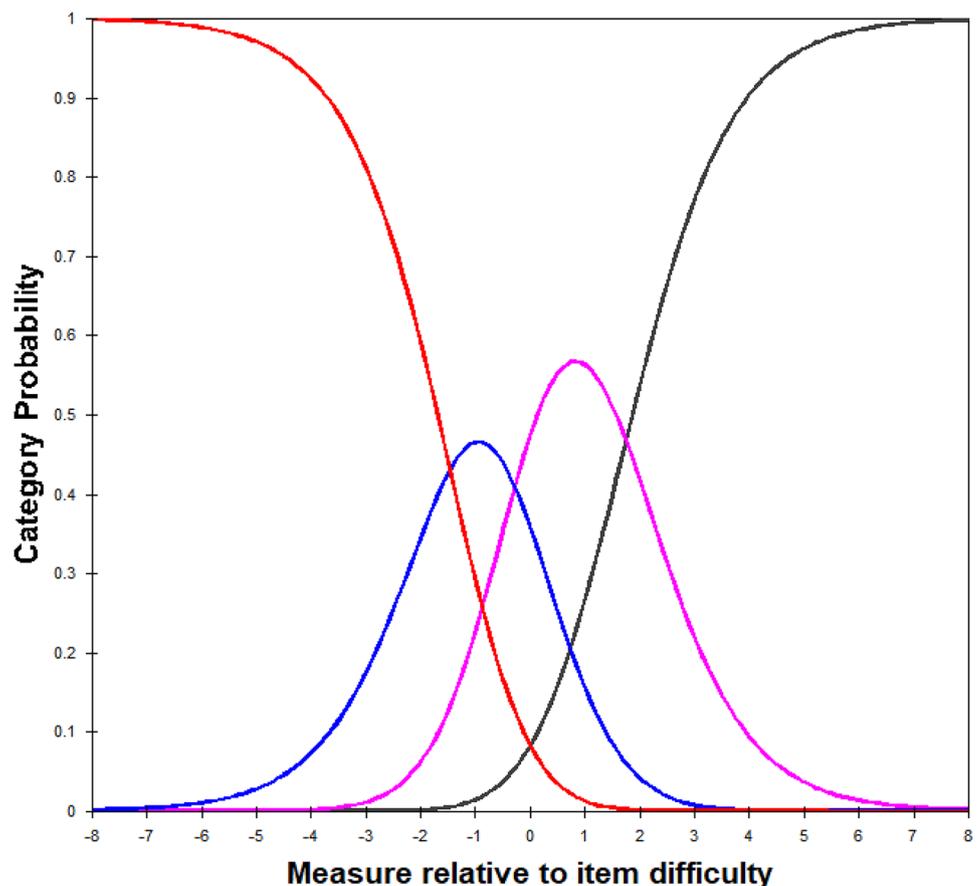
Fig. 2 Wright map (24 items, 4 categories)

Reliability and Validity Analysis of the Final, 22-item Instrument

Internal reliability testing of each of the five Focus Areas and the SSAT-22 as a whole demonstrated good to excellent internal reliability (DeVellis, 2012; George & Mallery, 2003) (SSAT-22: $n=22$; $\alpha=0.94$; Data: $n=8$; $\alpha=0.93$; Gateway Behaviours: $n=4$; $\alpha=0.84$; Reporting: $n=3$; $\alpha=0.73$; Response: $n=4$; $\alpha=0.86$; School Climate: $n=3$; $\alpha=0.90$). Discriminant validity was demonstrated by small to moderate correlations between each Focus Area and the vignette questions (Cohen, 1988), none of which reached statistical significance (Data: $r(6)=-0.100$, $p=0.586$; Gateway Behaviours: $r(2)=0.011$, $p=0.951$; Reporting: $r(1)=0.168$, $p=0.359$; Response: $r(2)=0.065$, $p=0.724$; School Climate:

$r(1)=-0.120$, $p=0.512$). Convergent validity was demonstrated by moderate to large (Cohen, 1988) positive correlations between each Focus Area and the second global evaluation rating (Data: $r(6)=0.428$, $p=0.009$; Gateway Behaviours: $r(2)=0.575$, $p<0.001$; Reporting: $r(1)=0.500$, $p=0.002$; Response: $r(2)=0.636$, $p<0.001$; School Climate: $r(1)=0.501$, $p=0.002$), as well as small to moderate positive correlations between each Focus Area and the Gap Analysis scale (Data: $r(6)=0.250$, $p=0.142$; Gateway Behaviours: $r(2)=0.337$, $p=0.045$; Reporting: $r(1)=0.265$, $p=0.119$; Response: $r(2)=0.317$, $p=0.059$; School Climate: $r(1)=0.426$, $p=0.010$).

Discussion

Given the prevalence of bullying in schools and its impact on both individual students and the broader school environment, it is surprising that no tool exists that allows schools to measure and monitor their level of system-level preparedness to prevent and respond to socially disruptive events. While individual interventions have been trialled and evaluated, these evaluations were conducted by outside academic researchers, with no mechanism by which the schools themselves could self-evaluate or self-monitor over time. Moreover, the emphasis of the evaluations was on individual bullying and victimisation behaviour, and specific programs and experiences, rather than system-level functioning (Gaffney et al., 2019, 2021).

Tanrikulu (2018), for example, in a systematic review of school-based cyberbullying prevention and intervention programs, identified 17 studies that presented empirical evaluations of effectiveness. The studies identified overwhelmingly used assessment tools that focused on individual behaviour rather than system-level preparedness (e.g., the Student Bullying Survey (Tangen & Campbell, 2010); a modified Olweus Bully/Victim questionnaire (Williford et al., 2013); the Cyber bullying Experience Survey (Pennell, 2013)). Moreover, none of the evaluations included in the review covered both prevention of and response to cyberbullying, with 15 focused on prevention and 2 on response. In addition, none explicitly emphasised preparedness.

In a more recent study, the Olweus Bullying Questionnaire (OBQ) is described as a school level measure (Olweus et al., 2020). In this study, Norwegian primary schools that did and did not regularly use the OBQ were compared two to eight years after the original implementation of the Olweus Bullying Prevent Program (between 2001 and 2005). It found that the former had fewer bullying problems, including significantly lower odds of students being bullied. However, while the authors described the OBQ as a school level measure, the instrument assesses "students' experiences with being bullied and bullying other students, their reactions as bystanders and their attitudes to bullying" (p.109). The article goes on to say that the results helped "the school leadership and staff, parents and, to some extent, students, to get a detailed

Fig. 3 Category probabilities
(22 items, 3 categories)
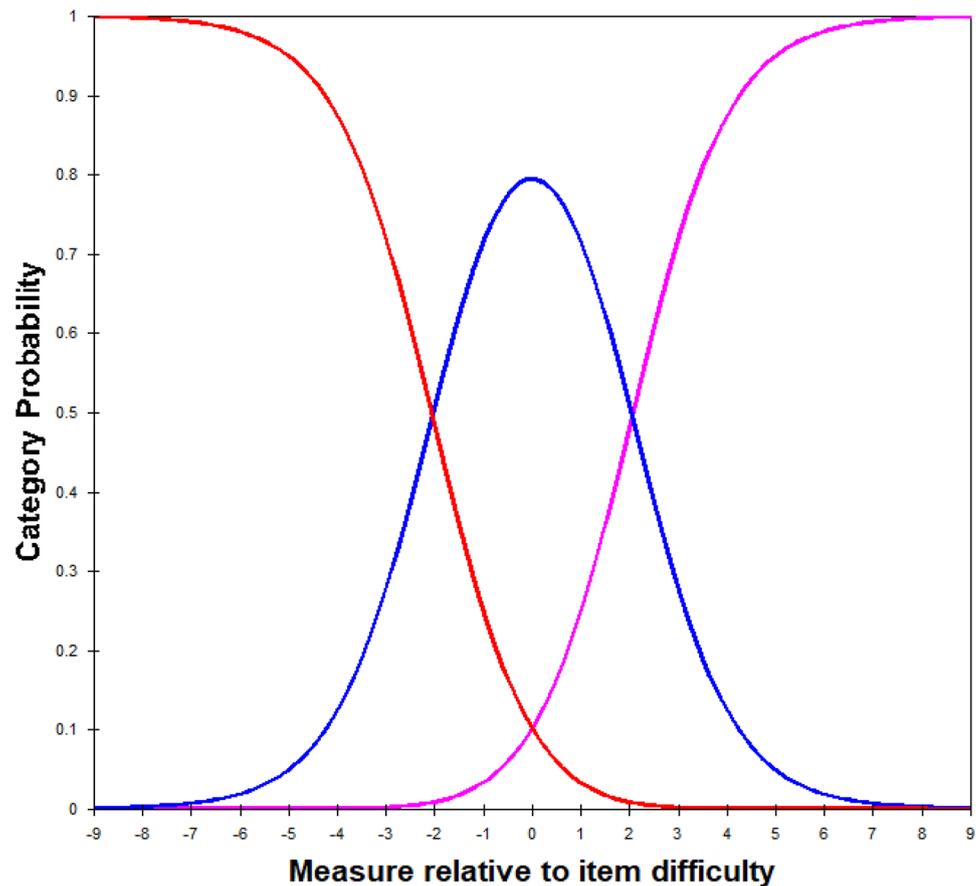
and reasonably realistic picture of the 'bully/victim situation' at their school'" (p.109). In other words, the focus was on individual behaviour rather than system-level preparedness.

A study with middle schools in the United States explored the use and impact of the Wisconsin Bullying Prevent Program Assessment Tool, which included nine components that seem to focus on system-level topics: policy and procedures, program selection/implementation, staff training, parent education and communication, classroom instruction/training, and analysis and continuous quality improvement (Bowser et al., 2020). The tool, however, is described as a gap analysis based on a review of schools' bullying prevention programs. This contrasts with the aim of the SSAT, which is to provide a "test" of the systems that schools currently have in place to prepare for, and respond, to, challenges and changes that impact on the school social environment such as bullying.

In this paper, we presented the development and initial testing of such an instrument. The result is a short, evidence-based, reliable, and validated instrument—the SSAT-22. Situated within socio-ecological system theory (Bronfenbrenner, 1977), and drawing on theories of behaviourism, social learning theory, prevention science, and systems change in the school climate literature, the SSAT-22 emphasises system-level, whole-school actions or approaches to combatting

bullying. For example, most items reference all students (e.g., *Students find the process of reporting easy and clear*) rather than just the victims or perpetrators of bullying, recognising that bullying is a whole-school phenomenon, not isolated individual bad behaviour. The SSAT-22 also includes items that specify the involvement of multiple actors (e.g., *The identification and reporting of gateway behaviours is viewed as a crucial activity for the school by the Principal, all school staff, students, and parents*), again recognising a whole of community approach. In this way, our approach mirrors the use of Bronfenbrenner's theory in Australian schools (Pennell et al., 2020), as it acknowledges that the spheres of influence around cyberbullying exist beyond the school in wider systems and structures in "reciprocal and dynamic ways" (Pennell et al., 2022, p. 290). The SSAT-22 responds to the literature arguing that whole-school approaches should be augmented with individual components (Rapee et al., 2020) (e.g., *A designated staff member(s) meets with the victim(s) of the cyber bullying or bullying and their parents/carers as soon as feasible to provide: comfort, support, and referral*). At the same time, although in a different format, the SSAT acknowledges Gaffney et al.'s (2021) argument that whole-school approaches that include all actors and address all levels of the school environment
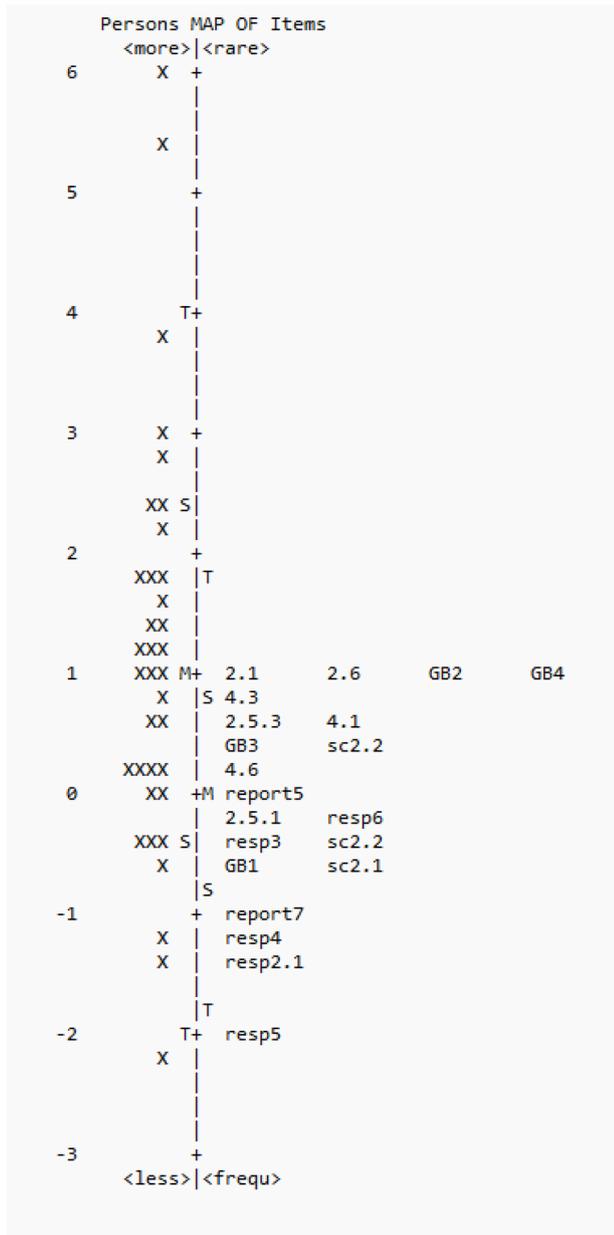
```
     Persons MAP OF Items
        <more>|<rare>
   6      X  +
             |
             |
          X  |
             |
   5         +
             |
             |
             |
             |
   4         T+
          X  |
             |
             |
   3      X  +
          X  |
             |
         XX S|
          X  |
   2         +
        XXX  |T
          X  |
         XX  |
        XXX  |
   1    XXX M+  2.1       2.6       GB2      GB4
          X  |S 4.3
         XX  |  2.5.3     4.1
             |  GB3       sc2.2
       XXXX  |  4.6
   0     XX  +M report5
             |  2.5.1     resp6
       XXX S|  resp3     sc2.2
          X  |  GB1       sc2.1
             |S
  -1         +  report7
          X  |  resp4
          X  |  resp2.1
             |
             |T
  -2      T+  resp5
          X  |
             |
             |
             |
  -3         +
        <less>|<frequ>

TABLE 12.5 C:\Cynthia Data\Cynthia 2009\Back to M ZOU
```

Fig. 4 Wright map (22 items, 3 categories)

are more effective than those that do not include these intervention components.

Schools may use the SSAT-22 in a number of ways, depending on their needs. They may, for example, use it as a monitoring tool—to track progress over time—overall and/or in each of the five Focus Areas. The focus of the SSAT-22 is on assessing schools' systems level preparedness to respond, but it may be interesting for schools to consider any association with reduced, or low rates of, traditional bullying, cyberbullying, and victimisation. They may also wish to examine individual SSAT-22

items within Focus Areas, to identify at a more granular level both their strengths and challenges. Finally, as part of a broader school consortium, they may want to compare results with other schools, and collaboratively explore ways to improve.

Limitations

As with any study, this one has some limitations. First, the sample size was small—just 36 schools, all located in Victoria, Australia. While this sample size is adequate for conducting internal reliability testing and Rasch analysis (Bujang et al., 2018; Linacre, 1994), it is somewhat underpowered for conducting validity testing (Hobart et al., 2012). Recruiting schools to participate in a study in the midst of the COVID pandemic was extremely difficult. We would argue that the importance of developing and testing the SSAT outweighs this limitation. It is hoped that future studies can replicate these analyses with a larger sample of schools as well as schools outside of Victoria, Australia. Second, there were insufficient items in the SSAT-22 targeting schools highly prepared to prevent and respond to bullying. Future research may want to consider adding and testing one or two additional items that better target higher-scoring schools. Third, we have not been able to examine the measure's ability to reflect/identify changes between groups or across times. Future studies should examine these issues.

The instrument should be adaptable to many other education systems around the world and will, we believe, provide an invaluable resource for schools grappling with the scourge of bullying and other socially disruptive activities.

Author Contribution NV: conceptualization; methodology; formal analysis; data curation; writing—original draft; FM: conceptualization; writing—original draft; RB: investigation; resources; data curation; writing—review and editing; CN: Writing—original draft; visualisation; CL: methodology; formal analysis; writing—review and editing; visualization; SF: investigation; resources; data curation; writing—review and editing.

Data Availability The data that support the findings of this study are available upon reasonable request and with the permission of the Alannah & Madeline Foundation (AMF). The data were gathered from a small set of schools as part of an evaluation of AMF's eSmart Schools Framework, and thus are not publicly available.

Declarations

Conflict of Interest The authors declare no competing interest.

References

Allen, K. P. (2010). A bullying intervention system in high school: a two-year school-wide follow-up. *Studies in Educational Evaluation, 36*(3), 83–92. https://doi.org/10.1016/j.stueduc.2011.01.002

Anastasi, A. & Urbina, S. (1997). *Psychological testing* (4th ed.). Prentice Hall Inc.

Bond, T., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Lawrence Erlbaum Associates.

Bowser, J., Bellmore, A., & Larson, J. (2020). Use and impact of the wisconsin bullying prevention program assessment tool in addressing middle school bullying. *International Journal of Bullying Prevention, 2*(4), 280–291.

Boyd, S., & Lawes, E. (2018). Leading schools that make a difference to bullying behaviour. *Journal of Educational Leadership, Policy and Practice, 33*(2), 90–103. https://doi.org/10.3316/informit.340690951301515

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513–531. https://doi.org/10.1037/0003-066X.32.7.513

Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for Cronbach's alpha test: a simple guide for researchers. *The Malaysian Journal of Medical Sciences: MJMS, 25*(6), 85.

Chalmers, C., Campbell, M. A., Spears, B. A., Butler, D., Cross, D., Slee, P., & Kift, S. (2016). School policies on bullying and cyberbullying: perspectives across three Australian states. *Educational Research, 58*(1), 91–109. https://doi.org/10.1080/00131881.2015.1129114

Cohen, J. (1988). The effect size. *Statistical power analysis for the behavioral sciences*, 77–83.

Cross, D., Shaw, T., Epstein, M., Pearce, N., Barnes, A., Burns, S., Waters, S., Lester, L., & Runions, K. (2018). Impact of the Friendly Schools whole-school intervention on transition to secondary school and adolescent bullying behaviour. *European Journal of Education, 53*(4), 495–513. https://doi.org/10.1111/ejed.12307

del Smith, P., Barrio, C., & Tokunaga, R. (2012). Definitions of bullying and cyberbullying: how useful are the terms. In S. Bauman, D. Cross, & J. Walker (Eds.), *Principles of cyberbullying research: Definitions, measures, and methods* (pp. 26–40). Routledge.

DeVellis, R. F. (2012). *Scale development: theory and applications* (3rd ed.). Sage.

Englander, E. (2013). *Bullying and cyberbullying: what every educator needs to know*. Harvard Education Press.

eSafety Commissioner. (2021). The digital lives of Aussie teens. https://www.esafety.gov.au/research/digital-lives-aussie-teens

Ford, R., King, T., Priest, N., & Kavanagh, A. (2017). Bulling and mental health and suicidal behaviour among 14- to 15- year-olds in a representative sample of Australian children. *Australian & New Zealand Journal of Psychiatry, 51*(9), 897–908. https://doi.org/10.1177/0004867417700275

Gaffney, H., Farrington, D., & Ttofi, M. (2019). Examining the effectiveness of school-bullying intervention programs globally: a meta-analysis. *International Journal of Bullying Prevention, 1*, 14–31. https://doi.org/10.1007/s42380-019-0007-4

Gaffney, H., Ttofi, M., & Farrington, D. (2021). What works in anti-bullying programs? Analysis of effective intervention components. *Journal of School Psychology, 85*, 37–56. https://doi.org/10.1016/j.jsp.2020.12.002

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.

Gupta, S., & Jawanda, M. K. (2020). The impacts of COVID-19 on children. *Acta Paediatrica, 109*(11), 2181–2183. https://doi.org/10.1111/apa.15484

Hawkins, L. G., Brown, C. C., Goad, C., Rhynehart, A., Hemphill, T., & Snyder, H. (2021). Bullying, family cohesion, and school connectedness: a moderated-mediation multigroup analysis of adolescents. *International Journal of Systemic Therapy, 32*(2), 93–114. https://doi.org/10.1080/2692398X.2021.1899738

Hobart, J. C., Cano, S. J., Warner, T. T., & Thompson, A. J. (2012). What sample sizes for reliability and validity studies in neurology? *Journal of Neurology, 259*(12), 2681–2694.

Hornby, G. (2016). Bullying: an ecological approach to intervention in schools. *Preventing School Failure: Alternative Education for Children and Youth, 60*(3), 222–230. https://doi.org/10.1080/1045988X.2015.1086969

Hu, Y., Bai, Y., Pan, Y., & Li, S. (2021). Cyberbullying victimization and depression among adolescents: a meta-analysis. *Psychiatry Research,* 305, 114198. https://doi.org/10.1016/j.psychres.2021.114198

Linacre, J. M. (1994). Sample size and item calibrations stability. *Rasch Measurement Transactions, 7*(4), 328.

Lindstrom Johnson, S., Waasdorp, T., Gaias, L., & Bradshaw, C. (2019). Parental responses to bullying: Understanding the role of school policies and practices. *Journal of Educational Psychology, 111*(3), 475–487. https://doi.org/10.1037/edu0000295

McCreary, L. L., Conrad, K. M., Conrad, K. J., Scott, C. K., Funk, R. R., & Dennis, M. L. (2013). Using the Rasch measurement model in psychometric analysis of the family effectiveness measure. *Nursing Research, 62*(3), 149–159.

Njelesani, J., Attard, K., Duimstra, A., & Zenderman, N. (2020). Addressing school bullying with a multi-tiered system of support approach. *Journal of Occupational Therapy, Schools, & Early Intervention*, 1–16. https://doi.org/10.1080/19411243.2020.1852466

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

Olweus, D., Solberg, M. E., & Breivik, K. (2020). Long-term school-level effects of the Olweus Bullying Prevention Program (OBPP). *Scandinavian Journal of Psychology, 61*(1), 108–116.

Pennell, B. D. (2013). *The effects of middle school single-sex bullying prevention classes on cyber bullying*. Trevecca Nazarene University.

Pennell, D., Campbell, M., & Tangen, D. (2020). What influences Australian secondary schools in their efforts to prevent and intervene in cyberbullying? *Educational Research, 62*(3), 284–303. https://doi.org/10.1080/00131881.2020.1795701

Pennell, D., Campbell, M., Tangen, D., & Knott, A. (2022). Should Australia have a law against cyberbullying? Problematising the murky legal environment of cyberbullying from perspectives within schools. *The Australian Educational Researcher, 49*, 827–844. https://doi.org/10.1007/s13384-021-00452-w

Rapee, R. M., Shaw, T., Hunt, C., Bussey, K., Hudson, J. L., Mihalopoulos, C., Roberts, C., Fitzpatrick, S., Radom, N., Cordin, T., Epstein, M., & Cross, D. (2020). Combining whole-school and targeted programs for the reduction of bullying victimization: a randomized, effectiveness trial. *Aggressive Behavior, 46*(3), 193–209. https://doi.org/10.1002/ab.21881

Sabia, J. J., & Bass, B. (2017). Do anti-bullying laws work? New evidence on school safety and youth violence. *Journal of*

Population Economics, 30(2), 473–502. https://doi.org/10.1007/s00148-016-0622-z

Sivaraman, B., Nye, E., & Bowes, L. (2019). School-based anti-bullying interventions for adolescents in low- and middle-income countries: a systematic review. *Aggression and Violent Behavior, 45*, 154–162. https://doi.org/10.1016/j.avb.2018.07.007

Tangen, D., & Campbell, M. (2010). Cyberbullying prevention: one primary school's approach. *Australian Journal of Guidance and Counselling, 20*, 225–234. https://doi.org/10.1375/ajgc.20.2.225

Tanrikulu, I. (2018). Cyberbullying prevention and intervention programs in schools: a systematic review. *School Psychology International,* 014303431774572. https://doi.org/10.1177/0143034317745721

United Nations. (2021). Safer Internet Day: prioritize and protect child well-being online, February 9th. https://news.un.org/en/story/2021/02/1084262

Williford, A., Elledge, L. C., Boulton, A. J., DePaolis, K. J., Little, T. D., & Salmivalli, C. (2013). Effects of the KiVa antibullying program on cyberbullying and cybervictimization frequency among Finnish youth. *Journal of Clinical Child and Adolescent Psychology, 42*, 820–833. https://doi.org/10.1080/15374416.2013.787623