Research Papers in Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rred20

Testing the social validity of the CATZ cross-age teaching zone anti-bullying intervention among school students

Michael J. Boulton & Peter J.R. Macaulay

To cite this article: Michael J. Boulton & Peter J.R. Macaulay (2024) Testing the social validity of the CATZ cross-age teaching zone anti-bullying intervention among school students, Research Papers in Education, 39:4, 629-649, DOI: 10.1080/02671522.2023.2189289

To link to this article: https://doi.org/10.1080/02671522.2023.2189289

Published online: 13 Mar 2023.

RESEARCH PAPERS IN EDUCATION
2024, VOL. 39, NO. 4, 629–649
https://doi.org/10.1080/02671522.2023.2189289

Routledge
Taylor & Francis Group

Testing the social validity of the CATZ cross-age teaching zone anti-bullying intervention among school students

Michael J. Boulton[a] and Peter J.R. Macaulay[b]

[a]School of Psychology, University of Chester, Chester, UK; [b]School of Psychology, University of Derby, Derby, UK

ABSTRACT

Bullying is a considerable problem among school students, and school-wide positive behaviour support interventions are regarded as helpful in addressing it. One approach is the CATZ Cross-age Teaching Zone anti-bullying intervention. The present study assessed the social validity of the CATZ anti-bullying intervention among a sample of 9–15-year-olds in a pre-post experimental design ($N = 817$, of which 546 experienced CATZ). Overall, participants expressed moderately positive views of the CATZ anti-bullying intervention, and these became significantly stronger following direct experience of it. Among participants who experienced the CATZ anti-bullying intervention, social validity ratings predicted a greater willingness to engage in it in the future. Alongside the extant data for its effectiveness, our findings support the wider use of the CATZ anti-bullying intervention in schools, and suggestions for how this might be brought about are discussed.

ARTICLE HISTORY
Received 12 July 2022
Accepted 14 February 2023

KEYWORDS
Anti-bullying; intervention; social validity; cross-age teaching; CATZ

Introduction

Bullying is a sub-class of aggressive behaviour that involves intentional and repeated attempts to cause distress or harm to a less powerful victim, and it may take many different forms (Olweus 1993, 2013). A substantial proportion of school students across the world are involved as victims, perpetrators, or onlookers Salmivalli 2010, 2014). Being bullied is associated with a host of negative outcomes including anxiety, depression, and loneliness (Hawker and Boulton 2000; Kowalski and Limber 2013), disrupted classroom concentration (Boulton 2008) and academic under achievement (Polanin, et al., 2021). While progress has been made in developing and evaluating anti-bullying interventions (Gaffney, Ttofi, and Farrington 2021; Samara and Smith 2008), and while reductions have been reported (Gaffney, Farrington, and Ttofi 2019), meta-analyses reveal that rates of bullying remain unacceptably high (Merrell et al. 2008; Ttofi and Farrington 2011).

CONTACT Peter J.R. Macaulay p.macaulay@derby.ac.uk School of Psychology, University of Derby, Kedleston Road, Derby DE22 1GB, UK

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

There are many reasons why anti-bullying interventions have had limited success, but of special relevance to the current study is the possibility that some, perhaps many, students are not receptive to the anti-bullying efforts and initiatives delivered by teachers and other adults. 2003) reported that 40% of a sample of nearly 8,000 14-year-olds believed that teachers were not usually interested in tackling bullying, and a similar proportion expressed little or no desire to collaborate with them in this regard. Similarly, Boulton and Boulton (2012) found that about 80% of 8 and 11-year-olds reported not attending to teachers' anti-bullying lessons, mainly because they said that they did not themselves engage in bullying and so it was not seen as personally relevant, they believed they knew enough about bullying already, that such lessons were boring or repeated things already addressed, and similar to Rigby and Bagshaw (2003), that teachers were not genuinely interested in stopping bullying. These kinds of notions are also reported elsewhere, with students perceiving that anti-bullying interventions did not engage them, and teachers delivered them in a repetitive manner (Cunningham et al. 2016).

For these reasons, student- led anti-bullying interventions have been implemented and evaluated. These often take the form of peer support service or buddy system (Boulton 2005; Cowie 2011; Tzani-Pepelasi et al. 2019). While helpful in some respects, notably supporting victims deal with negative emotions, these do not come close to being a 'proven' anti-bullying strategy either for those directly involved in bullying or as a primary prevention tool more generally (Gaffney, Ttofi, and Farrington 2021; Thompson and Smith 2011). Nevertheless, because students are now regarded as central to anti-bullying work, efforts to find alternative ways of involving them are warranted (Cowie 2011; Salmivalli 2010). This is especially true given that anti-bullying work needs to address a range of different issues beyond supporting victims, such as discouraging students from engaging in bullying in the first place and encouraging students to take a more active anti-bullying stance (Boulton et al. 2021; Macaulay, Boulton, and Betts 2019; Salmivalli 2014).

Recently, reports of a novel student-led anti-bullying intervention that combined two largely hitherto separate approaches, co-operative group work and cross-age teaching, have been published. While these approaches have rarely been used to address the problem of bullying (Gaffney, Ttofi, and Farrington 2021; Ttofi and Farrington 2011) and tend to have been used separately to tackle other issues, desirable outcomes for each have been reported. Co-operative group work has been shown to assist learning in academic (Slavin 2010; Veldman et al. 2020) and social/behavioural domains (Blatchford et al. 2006; Cowie et al. 1994). Similarly, cross-age teaching approaches have been found to benefit academic (Robinson, Schofield, and Steers-Wentzell 2005; Topping et al. 2011), and social/behavioural (Medway and Baron 1977; Robinson, Schofield, and Steers-Wentzell 2005) development. Given these positive *but separate* results for co-operative group work and cross-age teaching across such a wide variety of domains and variables, the novel intervention mentioned above sought to utilise their different characteristics and benefits. In this approach, called the Cross-Age Teaching Zone (CATZ henceforward), small co-operative groups of older students (tutors) are shown by adult facilitators how to develop and deliver a lesson on anti-bullying themes to younger schoolmates (tutees). No 'pull-out' groups are targeted in CATZ, rather all students can be invited to take part as tutors or tutees. A rationale is that *all* pupils and not just those directly involved in bullying can benefit from exposure to anti-bullying

learning opportunities. Indeed, it is well-known that 'onlookers' outnumber bullies and victims, and all of these sub-groups are regarded as having a significant role to play tackling bullying and its negative effects (Macaulay, Boulton, and Betts 2019; Salmivalli 2010, 2014). CATZ qualifies as school-wide positive behaviour support because it meets the criteria of promoting: (i) evidence-based practice (see below), (ii) change at the wider school community (not just individual) level, and (iii) schools' capacities to sustain good practice because one trained staff member can 'pyramid' expertise in CATZ to colleagues (Coyne, Simonsen, and Faggella-Luby 2008; Sugai and Horner 2006).

In terms of its anti-bullying evidence base, over a number of randomised control trials, CATZ has been shown to improve a range of anti-bullying beliefs (Boulton et al. 2021), help victims develop better adjustment (Boulton and Boulton 2017), and help students deal with peer provocations and avoid hostile attribution bias, factors known to precipitate bullying and aggression (Boulton & Macaulay, 2023). A further study found that CATZ also increased children's knowledge about how to stay safe online and that awareness could reasonably be seen as a way to help them avoid cyberbullying problems (Boulton et al. 2016). Effect sizes were mostly large, attesting to the practical value of CATZ for students themselves. Around 1000 students have already experienced CATZ as tutors in this ongoing program of work, and so the evidence-base for its effectiveness is not inconsiderable.

As well as assessing effectiveness of novel interventions, it is also important that researchers scrutinise their social validity. Most operational and conceptual definitions of social validity have the degree to which stake holders regard an intervention as 'acceptable' and 'useful' at their core (Cowan and Sheridan 2003; Kazdin 1980; Witt and Elliott 1985). Most studies have operationalised social validity with multi-item standardised measures, such as the Children's Intervention Rating Profile (Carter 2010). A virtue of this approach is that an 'overall acceptability' score for each stakeholder can be derived from those multiple items and hence statistical internal reliability and construct validity can be assessed. However, a focus on such psychometric indices may be at the expense of practical value because overall scores may reveal little (if anything) about consumers views of specific aspects of an intervention and the different ways in can be delivered. Carter (2010) sees this latter type of information as a key way to extract maximum value from social validity research, particularly in terms of guiding how practitioners may best implement an intervention. This principle has led some researchers to eschew 'standard though general' measures of social validity in favour of issue-specific measures that each can be operationalised with a single item (Cheon and Reeve 2015). For instance, Boyle et al. (2011) used focus groups of students to derive such single item measures. We would also suggest that using several of these more specific and student-derived measures facilitates a broad coverage of diverse social validity issues that can, in our case, guide how CATZ can best be delivered to students to maximise its benefits.

Social validity is also important is because it may influence treatment integrity and ultimately treatment outcome (Hieneman and Dunlap 2001; Witt and Elliott 1985). Indeed, interventions have even been known to fail *because* they are not well-received by students (Cowie et al. 1994). While assessments of social validity are usually from the perspective of those groups who will be in charge of implementing an intervention – often teachers and adult counsellors in school settings for school-wide positive behaviour

support – we agree with others who have called for students' views to be studied (Boyle et al. 2011; Cunningham et al. 2011; Doumas, Midgett, and Watts 2019). Relatedly, several authors have noted that social validity research is enhanced when it collects the views of multiple stake holders (Cowan and Sheridan 2003; Kazdin 1998). This is very relevant to CATZ given that there are two groups of 'direct' consumers; the older CATZ tutors who design and deliver anti-bullying lessons, and the younger CATZ tutees who receive them. Moreover, social validity is not static, and stakeholders' views of an intervention are known to change with experience (Frey et al. 2010). For these reasons, we assessed social validity among CATZ tutors and tutees both before and after they had experienced it.

Evaluations of co-operative group work and cross-age teaching that together 'make up' core aspects of CATZ have rarely reported social validity. Among those that have, Vilardo et al. (2013) found that all eight of their participants rated a cross-age teaching intervention (by third/fourth graders to help first graders with ADHD) very highly. However, such a small unrepresentative sample, half of whom had ADHD, is noteworthy. Another study found that school students *were* willing to work in co-operative groups to discuss bullying related issues and many reported enjoying doing so (Cowie and Berdondini 2001). This indirectly suggests that students may also have positive views of CATZ, but more direct tests are warranted given that the two interventions provide them with very different experiences.

Theories, too, provide a rationale for studying students' view of CATZ. Outcome-Expectancy Theory (Williams, Anderson, and Winett 2005) and the Theory of Planned Behavior (Ajzen 2002) both propose that people's intentions to engage in a behaviour are affected by outcome expectancies, that is, beliefs that particular outcomes will follow that behaviour. Importantly, this belief-behaviour link is even more likely when the beliefs and actions are *specific rather than general*. Thus, individuals who have positive (negative) views of CATZ as an anti-bullying tool, and high (low) self-efficacy for using CATZ in this way, should be more (less) likely to agree to engage in CATZ, or do so with greater (lesser) enthusiasm. In the domain of anti-bullying, but not in relation to CATZ *per se*, evidence has supported these predictions about beliefs-behaviour links for teachers (Boulton 2014) and for students (Boulton et al. 2013).

Studying stakeholders' subjective experiences of an intervention can help us understand *why* it is/is not acceptable to them, and ultimately why it might be (un)successful in terms of desired outcomes (Kazdin 1998). This approach has been used elsewhere to reveal process-outcome linkages for other non-CATZ peer-led anti-bullying initiatives among school students (Boulton, Trueman, and Rotenberg 2007; Cowie and Berdondini 2001). Students' subjective enjoyment and perceptions of success of group work have been shown to predict their future willingness to engage in it (Blatchford et al. 2003; Cowie and Berdondini 2001; Cowie et al. 1994; Slavin 2010). Based on this work, there is clear value to be had in assessing students' subjective experiences of CATZ and testing if these could predict willingness to engage in future CATZ activities.

When assessing social validity, it is important to consider what *specific* features or elements of an intervention students regard as more or less important (Boulton, Trueman, and Rotenberg 2007; Oliver and Candappa 2003). We now consider these in relation to some key parameters of CATZ that adult facilitators may benefit from manipulating. Firstly, Stukas, Clary, and Snyder (1999) found that *autonomy*, allowing students a significant say in *how* they helped their peers,

was rated by them as very desirable. Blatchford et al. (2003) similarly noted that students liked group work more if they had higher autonomy over group activities. Secondly, Robinson, Schofield, and Steers-Wentzell (2005) called for researchers to examine how much students valued the *training* provided by adults for new ways of working, but it would seem this has not yet been addressed in the literature. Thirdly, *who* students are allowed to work with may also be important. Boulton (2005) reported that just over a quarter of a sample of 11–16-year-olds preferred same-gender peer counsellors to help them with bullying and the rest did not express a preference. Cowie et al. (1994) found that many students preferred to work in friendship groups. Fourthly, *fear of public performance* may also influence the acceptability of CATZ because it is a common feature of more general social anxiety that often starts during the middle and later school years (Stein, Walker, and Forde 1996) and is a 'requirement' in CATZ for tutors. Clearly, a range of factors may influence how well CATZ is likely to be received by students (and hence its likely effectiveness) and therefore we included the aforementioned four parameters as part of our comprehensive assessment of its social validity.

Assessing social validity among students can also alert us to which sub-groups are more/less open to engaging positively with an intervention. There are different roles in bullying, and students adopting them may react differently to CATZ or have different views about it (Boulton 2013; Olweus 1993). Boulton (2013) found that victims avoided forming social relationships with other known victims partly because they believed doing so would increase their risk of becoming a target. Hence, victims may express more reservations about engaging in CATZ than other bullying sub-groups if they (mistakenly) believed that they would be 'teamed up' with other known victims. But if they could be enabled to realise that there is no such teaming up in CATZ, we might expect them to have social validity views that are similar to other sub-groups. Given that changing the beliefs and behaviour of perpetrators of bullying remains our ultimate goal, it is also important to assess how acceptable this sub-group regard CATZ to be in both an absolute sense, and relative to other sub-groups. Given that they *do* engage in bullying, it is reasonable to propose that they would hold pro-bullying beliefs. The latter may make bullies more resistant than the other sub-groups to CATZ, with its explicit aim to foster anti-bullying beliefs, given that people often strive for consistency according to Cognitive Dissonance Theory (Festinger 1957).

Girls are known to be more open to engaging in peer-led interventions (Boulton 2005; Cowie 2000), perhaps because acting in a somewhat formal helping capacity may compromise boys' sense of masculinity or macho self-image (Cowie 2000). McClellan and Kinsay (1999) found that mixed age classes can exert a positive effect on students relative to same age classes in terms of reduced social rejection and aggression and increased prosocial behaviour. DePaulo et al. (1989) found that 10-year-olds were more sensitive to the age of the students providing tutoring than were 8-year-olds. Boulton (2005) reported that boys were more concerned about the relative age of peer counsellors than girls, and he suggested that it was possible that some individuals, especially boys, may refuse to use peer-led services if available counsellors did not match their preferred age profile. Albeit indirectly, this evidence

suggests there may be gender and age differences in students' social validity views of CATZ.

Based on the rationales presented above, the following research questions were tested:

(1) How do student's evaluative attitudes towards CATZ, and their importance ratings of specific CATZ features (choosing who they worked with, choosing the content, choosing the mode of delivery, and the training provided by adults) prior to direct experience (called 'initial views')?
(2) What is the effect of direct experience of CATZ (as a tutor or tutee) on those social validity variables?
(3) What are CATZ tutors' subjective views of their CATZ experiences, and how do these predict an increased willingness to give a future CATZ lesson?
(4) Are there differences for bullying participant role, gender and age in participants' views of CATZ referred to in the first research questions above.

Method

Participants and procedure

The Department of Psychology ethics committee at the authors' university approved this study. Participants were drawn from three junior schools and two high schools in the UK selected on a convenience basis (being in the same City as the researchers' university to facilitate data collection). Each school identified classes of students that were available at the time of data collection. After discussions about the scope of the study, the nature of the intervention and the need to have a two-year age difference between CATZ *tutors* and *tutees* within each school, we focused on the oldest students (Year 6, $M_{age} = 11.5$ years) who acted as CATZ *tutors* or age-matched controls, and those in Year 4 ($M_{age} = 9.5$ years) who acted as CATZ *tutees* or age-matched controls at junior school, and on Year 10 ($M_{age} = 15.5$ years) students who acted as CATZ *tutors* or age-matched controls, and those in Year 7 ($M_{age} = 12.5$ years) who acted as CATZ *tutees* or age-matched controls at high school because teachers wanted these age groups to experience CATZ. Thus, Years 10 and 6 acted as CATZ *tutors* or age-matched controls, and Years 7 and 4 acted as CATZ *tutees* or age-matched controls

Consent was solicited from all of the students in the selected year/class groups, and their parents or head teacher in their *loco parentis* role, and we obtained a 96% response rate. All students who were subsequently offered it, took up the opportunity to participate in this study. This cohort made up what we refer to as the overall sample ($N = 817$). Our study had three phases, summarised as (i) Time 1 data collection of participants' initial views of CATZ and of nominations of bullying participant roles, (ii) participants experience the intervention/control condition, and (iii) Time 2 data collection of follow-up views of CATZ. We now describe each in detail.

At Phase 1, Time 1 data collection involved all participants responding to a questionnaire with nine items to provide what we call their initial views of CATZ. At this point, no student had any direct experience of CATZ and their only knowledge of it was based solely on a description provided by the researchers. Three separate classes of analyses were conducted on these Time 1 data. One set of analyses focused on testing

Table 1. Mean (standard deviation in brackets) of participants' responses to questions about CATZ at T1.

	Year 4			Year 6			Year7			Year10			
Dependent variable *n*	Girls (101)	Boys (101)	Both (202)	Girls (105)	Boys (89)	Both (194)	Girls (105)	Boys (109)	Both (214)	Girls (109)	Boys (98)	Both (207)	All (817)
Like to give CATZ	6.15 (1.39)	6.00 (1.39)	6.07 (1.39)	6.04 (1.87)	6.61 (1.66)	6.30 (1.79)	6.28 (1.39)	6.32 (1.34)	6.30 (1.36)	6.42 (2.02)	6.05 (2.00)	6.25 (2.02)	6.23 (1.66)
CATZ help younger pupils learn about bullying	7.04 (1.09)	7.31 (1.16)	7.17 (1.13)	6.86 (1.38)	6.93 (1.44)	6.89 (1.40)	6.88 (1.21)	6.98 (1.06)	6.93 (1.13)	6.92 (1.71)	6.81 (1.69)	6.86 (1.70)	6.96 (1.36)
CATZ help you learn about bullying	7.08 (1.26)	7.27 (1.16)	7.17 (1.21)	7.32 (1.23)	7.04 (1.30)	7.20 (1.26)	7.03 (1.07)	7.15 (1.08)	7.09 (1.07)	7.01 (1.23)	7.01 (1.43)	7.01 (1.32)	7.12 (1.22)
Importance of choosing group	8.12 (1.16)	7.89 (1.15)	8.01 (1.16)	8.02 (1.39)	8.10 (1.35)	8.06 (1.37)	7.88 (1.13)	7.89 (1.04)	7.88 (1.08)	8.11 (1.47)	7.99 (1.35)	8.05 (1.41)	8.00 (1.26)
Importance of choosing content	6.36 (1.22)	6.25 (1.15)	6.30 (1.19)	6.50 (1.23)	6.00 (1.35)	6.27 (1.31)	6.35 (1.12)	6.27 (1.15)	6.31 (1.13)	6.19 (1.44)	6.06 (1.31)	6.13 (1.38)	6.25 (1.25)
Importance of choosing way CATZ delivered	5.15 (1.19)	5.24 (1.18)	5.19 (1.18)	4.97 (1.60)	5.04 (1.24)	5.10 (1.44)	5.15 (0.99)	5.13 (1.16)	5.14 (1.07)	5.39 (1.75)	5.52 (1.53)	5.45 (1.65)	5.20 (1.36)
Importance of adult training	8.27 (1.23)	8.28 (0.97)	8.27 (1.11)	8.08 (1.25)	8.30 (1.15)	8.18 (1.21)	7.88 (1.09)	7.88 (1.01)	7.88 (1.05)	7.99 (1.51)	7.91 (1.59)	7.95 (1.55)	8.07 (1.25)
How good at CATZ	6.13 (1.57)	5.85 (1.42)	5.99 (1.50)	5.74 (2.08)	5.51 (1.95)	5.63 (2.02)	6.02 (1.65)	5.98 (1.66)	6.00 (1.65)	5.54 (1.96)	5.63 (1.93)	5.58 (1.94)	5.81 (1.79)
Worried about public speaking	8.30 (0.93)	8.19 (0.97)	8.24 (0.95)	8.50 (1.47)	8.45 (1.66)	8.48 (1.55)	7.84 (0.95)	8.15 (1.00)	8.00 (0.99)	8.03 (2.03)	8.03 (2.03)	8.04 (2.05)	8.18 (1.46)

Table 2. Mean (standard deviation in brackets) ratings given by pupils in CATZ tutor versus control groups at T1 and T2, tests of group x time interaction effects, and effect sizes.

	CATZ (n = 239)		Control (n = 162)		Group x Time interaction	
Dependent variable	Time 1	Time 2	Time 1	Time 2	F[a]	Partial η2
Like to give CATZ?	6.22 (1.89)	8.77 (1.33)	6.35 (1.94)	6.34 (1.73)	169.99***	.302
CATZ help younger pupils learn about bullying	6.90 (1.46)	8.97 (0.99)	6.85 (1.70)	6.51 (1.37)	213.48***	.352
CATZ help you learn about bullying	7.17 (1.32)	9.20 (1.17)	6.99 (1.26)	7.07 (1.12)	162.88***	.293
Importance of choosing group	8.08 (1.47)	8.50 (1.37)	8.02 (1.26)	7.73 (1.30)	60.69***	.134
Importance of choosing content	6.10 (1.34)	7.73 (1.35)	6.35 (1.26)	6.30 (1.26)	124.91***	.241
Importance of choosing way CATZ delivered	5.27 (1.54)	7.49 (1.56)	5.19 (1.61)	5.37 (1.37)	129.61***	.248
Importance of training	8.06 (1.44)	8.11 (1.54)	8.07 (1.34)	8.02 (1.58)	0.13	<.001
How good at CATZ?	5.67 (2.11)	8.20 (1.24)	5.52 (1.77)	5.33 (1.55)	199.32***	.337
Worried about public speaking	8.15 (1.98)	5.12 (2.12)	8.40 (1.59)	8.24 (1.64)	176.28***	.310

Note. [a]Degrees of freedom = (1,393).
***$p < .001$.

for age and sex differences in these initial views–see Table 1. A second set of analyses focused on testing for bullying participant role differences in these initial views. The third set of analyses compared these initial responses with those collected after participants had taken part in CATZ as (i) a tutor/age-matched control or (ii) a tutee/age-matched control (i.e. solicited at Time 2 data collection (see below) – see Tables 2–4). Because different numbers of participants were involved in these different analyses, our tables show sub-group numbers. The number of males/females at each Year who provided their initial views of CATZ are shown in Table 1. Hence, all participants in the overall sample responded at T1 to measures related to 'evaluative views of CATZ' (three items), 'importance of CATZ elements' (four items), 'public speaking anxiety' (one item) and 'self-efficacy for CATZ (one item) – see below for details about these measures.

Phase 2 of the study involved participants being assigned to and experiencing CATZ as a tutor (or age-matched control) or as a tutee (or age-matched control). This phase lasted about 2–3 weeks. Allocation to be in either a CATZ or an age-matched control condition was done on a class-by-class basis, with at least one CATZ tutor class and one age-matched control class, and one CATZ tutee class and one age-matched control, being present in each school. Allocation of classes within each school to CATZ versus control conditions was done using random allocation. Age matching was done by using classes within the same year group in each school.

As noted above, the CATZ intervention requires CATZ *tutors* (those who design and deliver a lesson) to be at least two years older than CATZ *tutees* (those who are recipients of a lesson). Students in Year 6 (63 girls and 57 boys) in junior schools and in Year 10 (59 girls and 60 boys) in high schools served as CATZ *tutors* or age-matched controls (Year 6 = 42 girls and 32 boys, and Year 10 = 50 girls and 38 boys). Students in Year 4 (47 girls and 51 boys) and Year 7 (50 girls and 55 boys) served as CATZ *tutees* or age-matched controls (Year 4 = 54 girls and 50 boys, and Year 7 = 55 girls and 54 boys).

Table 3. Mean (standard deviation in brackets) ratings given by pupils in CATZ tutee versus control groups at times 1 and 2, tests of group x time interaction effects, and effect sizes.

	CATZ ($n = 203$)		Control ($n = 213$)		Group x Time interaction	
Dependent variable	Time 1	Time 2	Time 1	Time 2	F[a]	Partial η2
Like to give CATZ?	6.35 (1.43)	7.88 (1.35)	6.03 (1.31)	6.10 (1.38)	84.11***	.171
CATZ help younger pupils learn about bullying	7.05 (1.21)	8.09 (0.99)	7.05 (1.07)	7.05 (0.86)	92.18***	.184
CATZ help you learn about bullying	7.05 (1.09)	7.77 (1.05)	7.20 (1.19)	7.19 (1.16)	104.68***	.204

Note. [a]Degrees of freedom = (1,408).
***$p < .001$.

Finally, Phase 3 involved Time 2 data collection in which, depending on their role as either CATZ tutor (or age-matched control) or tutee (or age-matched control), they were invited to complete a follow-up questionnaire that contained a sub-set of the questions asked in the original questionnaire. Time 2 data collection was 4–5 weeks after Time 1 data collection. For CATZ *tutors* and age-matched controls, Table 2 show a list of the follow-up questions. For the CATZ *tutor*s only (i.e. not the age-matched controls), this follow-up questionnaire also contained three items that asked them to reflect on their CATZ experiences (see Table 4). Likewise at Time 2, after providing their initial views of CATZ at Time 1, CATZ *tutees* and age-matched controls were then presented with their version of the follow-up questionnaire that repeated some of the items that were in in the initial views' questionnaire – see Table 3 for a list of these questions

As a reminder for readers, the number of students in sub-groups that were compared in analyses are given in Tables 2–4.

For participants assigned to experience CATZ, the procedure was the same as that employed in previous studies that have evaluated CATZ (Boulton et al. 2016; Boulton and Boulton 2017; Boulton et al. 2021, Boulton & Macaulay, 2023). After being invited to take part, groups of about five tutors were formed either on the basis of existing within-class table groups in junior schools or within-registration class friendship groups in high schools. We ensured that all students were happy working with fellow group members by giving them initial and ongoing opportunities to change groups at their own discretion. Over four or five circa one-hour sessions, we assisted each group to develop a circa 40-minute lesson for a small group of younger tutees. We provided them with information about bullying (such as the different forms it may take, why they are all unacceptable, and how students can safely take a stand against them) but groups were encouraged to 'take ownership' of this content and to incorporate this into their own lesson. As a minimum, we encouraged each group to develop a poster and associated script that could be used to structure and deliver their content to the tutees. Control participants did not experience CATZ but rather continued with 'business as usual' in their classes. When they were invited to give their 'follow-up' views of CATZ at Time 2, we prefaced this by stating that we were interested in seeing if their views had stayed the same or had changed.

Time 1 and Time 2 data were collected on a whole class basis, with each participant being given a questionnaire. A researcher read out the instructions followed by each question. So that participants knew what CATZ is, they were informed that, "CATZ stands for the cross-age teaching zone. It is where older students act as teachers to

Table 4. Mean (standard deviation in brackets) of tutors' subjective views about their CATZ experience.

	Year 6			Year 10			All		
Dependent variable	Girls ($n = 63$)	Boys ($n = 57$)	Both ($n = 120$)	Girls ($n = 59$)	Boys ($n = 60$)	Both ($n = 119$)	Girls ($n = 122$)	Boys ($n = 117$)	Both ($n = 239$)
How well CATZ went	7.32 (1.92)	7.32 (1.98)	7.32 (1.94)	7.15 (2.13)	6.67 (2.36)	6.91 (2.25)	7.24 (2.02)	6.98 (2.20)	7.11 (2.11)
How well group worked	7.21 (1.95)	7.02 (1.99)	7.12 (1.96)	6.64 (2.20)	7.07 (2.34)	6.86 (2.27)	6.93 (2.09)	7.04 (2.17)	7.00 (2.12)
How much enjoyed CATZ	8.46 (1.74)	7.81 (2.14)	8.15 (1.96)	7.63 (2.27)	7.95 (2.13)	7.79 (2.20)	8.06 (2.05)	7.88 (2.13)	7.97 (2.09)

younger students. The older students design a lesson in a small group about bullying covering things like what can help people understand it, the different types of bullying, how students can help to stop it happening, and how they can help people who have been bullied. Adults help them do this, but it is very much the older students' lesson. They are in charge. They then go into a class of younger students to give their lesson about bullying to a small group. The adult teachers are always there to keep an eye on things, but they are not giving the lesson. This questionnaire is not a test so there are no right or wrong answers. We just want to know what each of you thinks about CATZ as a possible way to learn about bullying, and so there is no need to copy what somebody else has put. Is that, OK?" Bullying was also defined in a standard way, 'Bullying is where somebody who is stronger or more powerful does something nasty to someone who is weaker or less powerful. They don't just do this once but at least several times and the person being bullied finds it hard to defend themselves. Bullying can be different things like hitting and kicking, calling people nasty names, leaving them out on purpose, spreading nasty stories about them, getting other people not to like them or be their friend, sending nasty texts and online messages and things like that'.

Due to their sensitive nature, peer nominations to determine bullying participant roles (see below) were collected in individual interviews.

Measures and test-retest reliability

In a pilot test, and in line with the approach of Boyle et al. (2011), we invited a separate group of students who had experienced CATZ to identify any specific issues that they thought affected their views of CATZ. Several of these corresponded with issues highlighted in our literature review in the Introduction, and we focused on them. These were each operationalised into a question that, in combination, provided a relatively comprehensive assessment of social validity (see below). Such single item measures preclude tests of internal consistency and so we assessed test-retest reliability over a one-week period with a sub-set of the sample ($N = 92$). Coefficients were high and significant for all the questions (r's .70 to .92, all $p < .001$). All but one question (see below for exception) had a 1 to 10 response scale, anchored by 'not at all' and 'a lot', respectively, with higher scores representing a more positive view of CATZ or a higher value placed on the relevant feature.

Evaluative views of CATZ

These were measured with three items; (i) 'How much would you like to design and give a CATZ lesson about bullying to younger students, I mean be the teacher?', (ii) 'How much do you think CATZ would *help younger students* learn useful things about bullying and how to stop it happening, if they were taught about it by older students?', and (iii) 'How much do you think CATZ would *help you* learn useful things about bullying and how to stop it happening, I mean if you could teach younger students about it?'

Importance of CATZ elements

This was measured with four questions that began with the stem, 'Thinking about you giving a CATZ lesson about bullying, how important . . . .' followed by, (i) 'is it that you choose *who you work with*?', (ii) 'is it that you choose *what goes in to your CATZ lesson*?', (iii) 'is it that you choose *how to give the lesson*?', and (iv) 'is the *training you get from an adult* so you know how to do it?'

Public speaking anxiety

This was measured with, 'Thinking about you giving a CATZ lesson about bullying, how worried would you be about speaking in front of other people?'

Self-efficacy for CATZ

This was measured with, 'Thinking about you giving a CATZ lesson about bullying, how good at this do you think you would be?'

Subjective views of CATZ experiences

Among CATZ tutors only (controls did not experience CATZ), this was assessed at T2 with three items; (i) 'How successful did you think CATZ was for your group, how well did it go?', (ii) 'How well did your group work together?' and (iii) 'How much did you enjoy doing CATZ?'

Relative preference for students versus teachers

This was assessed with, 'Who would you prefer to teach you new and important things about bullying and how to stop it happening?' Response options were 'Older students with CATZ, Don't mind, and Teachers'.

Bullying participant roles

These were measured with peer nominations using existing protocols (Boulton and Smith 1994). Each participant was presented with a list of classmates and asked to identify those who were victims and bullies. To control for different class sizes, raw counts were converted to percentages. Using similar criteria to Boulton and Smith (1994), Victims were those who received 25+% victim nominations and less than 15% bully nominations, Bullies were those who received 25+% bully nominations and less than 15% victim nominations, Bully-Victims were those who received 25+% victim nominations and 25+% bully nominations, and Not-Involved were those who received less than 15% victim and 15% bully nominations.

Results

Plan of analyses and tests of CATZ versus control group equivalence on initial views

To examine if there were significant year or gender differences across initial views for all items (except relative preference for CATZ tutors/teachers, see below), responses from the entire sample at T1 were analysed in an omnibus 2 (Year) x 2 (Gender) analysis of variance (ANOVA) tests, with the 1–10 ratings as dependent variables (DVs) (descriptive data are in Table 1). Year and gender difference on the item concerning relative preference for CATZ tutors/teachers at T1, were tested with loglinear analysis (Stevens 1992) since responses were categorical.

To determine if there were effects of experience as a CATZ *tutor* on ratings, a series of 2 (Time) x 2 (Group: CATZ tutor versus control) x 2 (Year) x 2 (Gender) mixed ANOVA tests were carried out, with a separate item serving as DV in each test. Similar tests examined the effects of experience as a CATZ *tutee* on ratings concerning evaluative views of CATZ. No Group x Time x Gender, Group x Time x Year, or Group x Time x Gender x Year interaction was significant, showing that the effects of CATZ experience were not qualified (moderated) by year and/or gender. Hence, the findings from simplified 2 (Group: CATZ versus Control) x 2 (Time) ANOVAs, and their associated descriptive data are presented (Tables 2 and 3). Prior to carrying out these tests of the effects of CATZ experiences, independent samples t-tests confirmed that the CATZ and control participants within each Year group did not differ significantly on any of the items (all $p > .10$).

A one-way ANOVA tested for differences between bullying participant roles, again with ratings as DVs. On the item to solicit relative preference for CATZ tutors/teachers, a χ^2 test was employed since responses were categorical.

Given that multiple tests were carried out, a more conservative alpha level of <.01 was adopted to help attenuate the possibility of a type 1 error. Partial eta squared (η2) was used as a measure of effect size for ANOVAs.

Initial evaluative attitudes towards CATZ, and effects of CATZ Experience

Descriptive data concerning initial evaluative attitudes in the whole sample (i.e. CATZ tutors and tutees and all age-matched controls, see Table 1) indicated moderately high ratings on all three items (means ranged from 6.23 to 7.12), with no significant year or gender differences. On all three items, there was a significant effect of CATZ tutor experience (Table 2), and separately of CATZ tutee experience (Table 3). In all cases, CATZ experience among both tutors and tutees led to a significantly more positive evaluation of CATZ from T1 to T2. Effect sizes were stronger for the effect of CATZ tutor experience than for CATZ tutee experience. Scores for age-matched control students, both tutors and tutees, did not differ significantly between T1 and T2 on any of these variables.

Initial importance ratings of CATZ elements, and effects of CATZ experience among tutors

Descriptive data concerning initial importance ratings of CATZ elements in the whole sample (Table 1) indicated moderate to high scores (means ranged from 5.20 to 8.07),

with no significant year or gender differences. The *relative* importance of the four CATZ elements was assessed with a repeated measures ANOVA. This was significant, F (3, 2448) = 974.27, $p < .001$. Follow-up tests, with a Bonferroni correction for multiple comparisons, indicated that ratings of adult training ($M = 8.07$) and choosing group members ($M = 8.00$) did not differ significantly, but were each significantly higher than for the other two elements. Choosing content was rated next highest ($M = 6.25$) and significantly higher than choosing the way CATZ is delivered ($M = 5.20$). All significant differences were $p < .001$.

On three elements, there was a significant and positive effect of CATZ tutor experience as indicated by a significant Time x Group interaction (Table 2). Effect sizes were higher for choosing the way CATZ is delivered and choosing the content than for choosing group members, although the latter was rated high initially at T1 (see above) and was highest of all at T2, and so there was less room for change in a positive direction on this variable. Scores for age-matched control students did not differ significantly between T1 and T2 on these three variables. There was no significant effect of experience on importance ratings of adult training; it remained high across time for both CATZ and control groups.

Initial self-efficacy for CATZ and fear of public speaking, and effects of CATZ experience among tutors

Overall, participants in the whole sample initially reported a moderate level of self-efficacy for CATZ ($M = 5.81$) and a high fear of public speaking ($M = 8.18$) (Table 1). CATZ tutor experience led to a significant increase in self-efficacy, and a significant decrease in fear of public speaking, and in both cases the effect size was large (Table 2).

Subjective views of CATZ tutor experience and using these to predict changes in willingness to give a CATZ lesson

Descriptive data relating to subjective CATZ tutor experience (Table 4) indicated moderately high ratings (means ranged from 7.00 to 7.97), with no significant year or gender differences. To determine if subjective views could predict changes in willingness to take part in CATZ, we could not use multilevel (ML) analyses to account for the nesting of participants into groups because there had been some changing of group membership that would have violated the assumption of independence across the different groups (Tabachnick and Fidell 2014). Instead, a hierarchical OLS multiple regression model was employed. This allowed us to test our research question 3 (i.e. of CATZ tutors' subjective views of their CATZ experiences could predict an increased willingness to give a future CATZ lesson. T2 willingness to give CATZ was the dependent variable and in order to obtain a measure of change in it, the corresponding T1 willingness score was entered at the first step (as in Boulton, Smith, and Cowie 2010). Year and gender were also entered at step 1 both as control variables and so their product terms with the experience variables could be entered at step 3. The mean centred (Tabachnick & Fidell, 20014) experience rating variables were then entered together at step 2, allowing us to determine how much variance in change in willingness to give a CATZ lesson they could account for collectively, *and* if each was a unique predictor. Finally, the product terms of each of the three experience rating variables with year and with gender were entered at step 3,

allowing us to determine if the latter two variables moderated the influence of the experience rating variables. Step 1 was significant, F (3, 235) = 4.79, p = .003, accounting for 5.8% of the variance. Of greater interest are the results for step 2, which was also significant, F (3,232) = 31.83, p < .001. Collectively, the three experience variables accounted for 33.2% of the variance, and moreover, each one was a significant unique predictor (how well CATZ went β = .33, p < .001, how well the group worked β = .19, p < .01, and how much CATZ was enjoyed β = .52, p < .001). None of the product terms at step 3 were significant when entered as a group, or individually to avoid multicollinearity.

Relative preference for CATZ tutors versus teachers

Overall, at T1, 764 (93.2%) participants said they would prefer to learn new things about bullying from CATZ tutors, 31 (3.8%) didn't mind, and only 22 (2.7%) preferred teachers.

Bully participant role differences in initial ratings

Victims (M = 9.11) rated the importance of choosing group members significantly higher than Bullies (M = 7.95), Bully-Victims (M = 7.44) and Not Involved (M = 7.81), and the latter three groups did not differ, F (3, 813) = 42.49, p < .001. Bullies (M = 8.03) had significantly higher CATZ self-efficacy scores than Victims (M = 5.43), Bully-Victims (M = 5.94) and Not involved (M = 5.52), and the latter three groups did not differ, F (3, 813) = 66.08, p < .001. On a χ^2 test, there was no significant difference in the proportion of students in the four groups that indicated a preference for being taught new things about anti-bullying by CATZ tutors versus teachers (p > .10). As noted above, there was a strong preference for CATZ tutors among all sub-groups.

Discussion

The general aim of this study was to solicit students' beliefs about a relatively new school-wide positive behaviour support intervention, CATZ, and to examine the effect of direct experience of CATZ on those views. Overall, participants expressed moderately positive views of CATZ as an anti-bullying intervention initially, and CATZ experiences among both tutors and tutees led to significantly better ratings of it. Some potential barriers to participation were identified at T1 in terms of quite low self-efficacy and high fear of public speaking scores. Once again, experience as a CATZ tutor led to significant improvements on these variables (effects of tutee experience on these variables was not assessed because they did not deliver a CATZ lesson). This is important given theories which highlight such variables, especially self-efficacy, as contributing to willingness to engage in actual behaviour (Ajzen 2002; Williams, Anderson, and Winett 2005). We also found an overwhelming preference among our participants as a whole for learning new and helpful things about bullying from CATZ tutors over teachers. This is consistent with reports that many students are rather unreceptive to being taught about anti-bullying issues by their teachers (Boulton and Boulton 2012; Cunningham et al. 2016; Rigby and Bagshaw 2003).

Our study has also revealed what elements of CATZ students regard as more or less important. Across the sample as a whole, the training provided by adults and being able to choose group members were seen as significantly more important than being able to choose the content of the lesson, that in turn was rated significantly higher than being able to choose the way CATZ is delivered. The relatively low initial ratings of being able to decide what and how to teach is perhaps not surprising since acting as a teacher would be outside the experience of most students. Such a notion is supported by our finding that direct experience as a CATZ tutor also led students to rate three of these four CATZ elements as more important than they initially did. While experience did not affect the value placed on training provided by adults, it is noteworthy that these ratings remained high (about 8 on our 10-point rating scale) in both CATZ and control participants across T1 and T2. Hence, adults who offer students the chance to engage in CATZ more than once would do well to offer them 'enough' training initially, and increasingly offer greater freedom to choose what they do and who they do it with. This notion is consistent with our finding that CATZ tutor experience led to gains in CATZ self-efficacy as this result suggest that direct experience may have left students feeling that they were now ready to receive that greater autonomy. Moreover, after taking part as a CATZ tutor, participants reported overall relatively high levels of success, enjoyment and satisfaction with their group (around 7–8 on our 10-point scale). Scores on each one of these subjective ratings of their CATZ experience uniquely predicted (i.e. after their common variance was controlled) a positive change from T1 to T2 in willingness to engage in further CATZ activities, and collectively they accounted for a very substantial amount of the variance (33.2%) in that change.

Taken together, our findings indicate that engaging in CATZ led students to view it and their own competence even more highly. This provides a *prima facia* case for testing the possibility that more than one CATZ experience may lead to even more benefits than the considerable ones that have so far been found (Boulton et al. 2016; Boulton and Boulton 2017; Boulton et al. 2021, Boulton & Macaulay, 2023) and for testing the effect of different 'doses'. Providing opportunities for students to take part in CATZ, especially more than once, is entirely consistent with calls from influential policy makers to encourage greater student engagement with all kinds of school issues and problems, including bullying (e.g. Australian Government Department of Education and Early Childhood Development, Student Engagement Policy Guidelines 2013) and with school-wide positive behaviour support principles.

We did not find any significant differences between girls' and boy's views of CATZ, either initially or following CATZ experience, including willingness to act as a CATZ tutor. In some ways, this was unexpected given that boys have been found to be less open to delivering other peer-led interventions than girls (Boulton 2005; Cowie 2000). It might be because CATZ is a less formal, or even less overtly 'explicit', kind of helping that doesn't challenge (some) boys' sense of masculinity or macho-self-image as may happen with other forms of peer support. It is also possible that giving a lesson to younger students' during CATZ allows boys (and girls) to recognise their superior knowledge or 'status', gains them respect, and so is something they value. The reasons why boys, and some girls, appear to be more open to CATZ than to other forms of peer support are worthy of study in the future.

In terms of participant roles, victims rated choosing who they worked with as more important than any of the other sub-groups. This may reflect victims' reluctance to engage in activities that could involve contact with individuals they know to be bullies or victims, or unknown others that may be. Encouraging victims to take part in CATZ is important given that they often do not receive appropriate social support, and that they have been shown to benefit in important ways through acting as CATZ tutors (Boulton and Boulton 2017). Hence, adults who run CATZ program should be sensitive to these kinds of concerns and ensure there is adequate scope for *all* potential participants, especially victims, to have a say in who they work with or at least feel comfortable working with them. In our study, we encourage students to work in existing table or friendship groups and gave them autonomy over changing groups if they wanted to.

We also found significantly higher CATZ self-efficacy ratings among Bullies than any of the other sub-groups. We deem this to be noteworthy when considered alongside the principle contained within some social psychological theories (Ajzen 2002; Williams, Anderson, and Winett 2005) that self-efficacy and other outcome expectancy beliefs will likely influence behaviour. It seems reasonable to suggest, albeit speculatively, that bullies with high CATZ self-efficacy may embrace CATZ anti-bullying activities and so develop more anti-bullying beliefs and in turn engage in less bullying. Future studies could test this idea because while a rationale for the universal application of CATZ is that *all* pupils can benefit from exposure to anti-bullying learning opportunities, a principle consistent with school-wide positive behaviour support (Coyne, Simonsen, and Faggella-Luby 2008; Sugai and Horner 2006), it is still the case that interventions that promote positive change in bullies themselves is especially desirable.

Our results, taken as a whole, attest to the potential value of CATZ as an anti-bullying intervention. It was initially viewed moderately positively by the majority of girls and boys aged between 9 and 15 years and who occupied various bullying participant roles in our sample. Our findings that CATZ experience led to even more positive views among *both* tutors and tutees is noteworthy because it suggests that CATZ has the potential to provide 'twice the bang for the buck' because for the same amount of staff effort to run CATZ activities, we yield benefits in two rather than one group of students. This further supports its fit within a school-wide positive behaviour support framework.

An important assumption that guided the present study was that treatment acceptability is the first step in a sequence that impacts on treatment integrity and ultimately outcome (Witt and Elliott 1985). It would be useful if researchers could test this more directly; *is* it the case that students with a greater willingness to engage in CATZ invest more of themselves in the process and end up learning more (and imparting more knowledge to tutees), relative to those who are less enthusiastic? If so, then adults could invest time 'selling' the benefits of CATZ prior to actually implementing it. In our own work, we have found that (older) tutors are often competent 'ambassadors' willing to encourage other (younger) students to take part.

A strength of our study is that while the views expressed at T1 were hypothetical in the sense that participants had no experience of CATZ, our design allowed us to chart the influence of actual CATZ experience on their ratings. As Cowan and Sheridan pointed out (Cowan and Sheridan 2003, 17), 'Obtaining treatment acceptability ratings prior to and following interventions may allow researchers to better determine

whether actual exposure to intervention influenced acceptability ratings'. That our sample consisted of a diverse age range that encompassed the likely target groups for CATZ tutor and tutee experiences across the school years is also a strength. However, while fairly large, the sample was restricted to three junior schools and two high schools in the UK and so the results may not generalise and replication attempts are warranted.

In conclusion, while our participants may have had some initial reservations, they also expressed moderately positive views about CATZ and were fairly willing to participate in it. Direct experience led to improvements on all of our measures of social validity. This is encouraging considered in the light of the growing evidence that CATZ can be an effective way to help students develop anti-bullying beliefs. Hence, we believe the time is right to suggest that CATZ be utilised more widely as a school-wide positive behaviour support intervention to help students themselves address the widespread and on-going problem of bullying. Our findings also suggest that offering potential CATZ participants 'what they want' (choosing working partners and the details of their lesson, adequate training, etc.) and allaying their fears (of public speaking) at the time they are initially invited to take part may encourage them to engage willingly and productively in that process.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Michael J. Boulton is an Emeritus Professor of Psychology at the University of Chester, UK. He has research interests in bullying related issues, is recognised as an international expert on bullying among school pupils and has published numerous papers on the topic over 25 years.

Peter J. R. Macaulay is a Lecturer in Psychology at the University of Derby, UK. His main research interests are in social development, focusing specifically on teacher's perceptions towards cyberbullying, children's bystander behaviour in the online/offline domain, and children's knowledge of online risks and safety.

ORCID

Peter J.R. Macaulay http://orcid.org/0000-0003-4891-9940

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Ajzen, I. 2002. "Residual Effects of Past on Later Behavior: Habituation and Reasoned Action Perspectives." *Personality and Social Psychology Review* 6 (2): 107–122. doi:10.1207/S15327957PSPR0602_02.

Australian Government Department of Education and Early Childhood Development, Student Engagement Policy Guidelines. (2013). http://www.education.vic.gov.au/school/principals/participation/Pages/studengage.aspx,20.3.14.

Blatchford, P., E. Baines, C. Rubie-Davies, P. Bassett, and A. Chowne. 2006. "The Effect of a New Approach to Group Work on Pupil-Pupil and Teacher-Pupil Interactions." *Journal of Educational Psychology* 98 (4): 750. doi:10.1037/0022-0663.98.4.750.

Blatchford, P., P. Kutnick, E. Baines, and M. Galton. 2003. "Toward a Social Pedagogy of Classroom Group Work." *International Journal of Educational Research* 39 (1–2): 153–172. doi:10.1016/S0883-0355(03)00078-8.

Boulton, M. J. 2005. "School Peer Counselling for Bullying Services as a Source of Social Support: A Study with Secondary School Pupils." *British Journal of Guidance & Counselling* 33 (4): 485–494. doi:10.1080/03069880500327546.

Boulton, M. 2008. "Pupils' Perceptions of Bullying and Disruptions to Concentration and Attention to School Work." *Pastoral Care in Education* 26 (2): 83–89. doi:10.1080/02643940802062592.

Boulton, M. J. 2013. "The Effects of Victim of Bullying Reputation on adolescents' Choice of Friends: Mediation by Fear of Becoming a Victim of Bullying, Moderation by Victim Status, and Implications for Befriending Interventions." *Journal of Experimental Child Psychology* 114 (1): 146–160. doi:10.1016/j.jecp.2012.05.001.

Boulton, M. J. 2014. "Teachers' Self-Efficacy, Perceived Effectiveness Beliefs, and Reported Use of Cognitive-Behavioral Approaches to Bullying Among Pupils: Effects of In-Service Training with the I DECIDE Program." *Behavior Therapy* 45 (3): 328–343. doi:10.1016/j.beth.2013.12.004.

Boulton, M. J., and R. Boulton. 2012. "Resistant to the Message: Are Pupils Unreceptive to teachers' Anti-Bullying Initiatives and if so Why?" *Educational Studies* 38 (5): 485–489. doi:10.1080/03055698.2011.643112.

Boulton, M. J., and L. Boulton. 2017. "Modifying Self-Blame, Self-Esteem, and Disclosure Through a Cooperative Cross-Age Teaching Intervention for Bullying Among Adolescents." *Violence and Victims* 32 (4): 609–626. doi:10.1891/0886-6708.VV-D-15-00075.

Boulton, M. J., L. Boulton, E. Camerone, J. Down, J. Hughes, C. Kirkbride, R. Kirkham, P. Macaulay, and J. Sanders. 2016. "Enhancing Primary School Children's Knowledge of Online Safety and Risks with the CATZ Cooperative Cross-Age Teaching Intervention: Results from a Pilot Study." *Cyberpsychology, Behavior and Social Networking* 19 (10): 609–614. doi:https://doi.org/10.1089/cyber.2016.0046.

Boulton, M. J., and P. J. Macaulay 2023. "Helping School Students Deal with Peer Provocations and Avoid Hostile Attribution Bias with the CATZ Cross-Age Teaching Zone Intervention: Evidence for Effectiveness and Social Validity." Manuscript submitted for publication.

Boulton, M. J., P. J. Macaulay, S. Atherton, L. Boulton, T. Colebourne, M. Davies, J. Down, I. Garner, B. Harriss, L. Kenton, B. Lomas. 2021. "Promoting Junior School Students' Anti-Bullying Beliefs with the CATZ Cross-Age Teaching Zone Intervention." *International Journal of Bullying Prevention* 1–14. doi:https://doi.org/10.1007/s42380-021-00111-9.

Boulton, M. J., D. Murphy, J. Lloyd, S. Besling, J. Coote, J. Lewis, R. Perrin, and L. Walsh. 2013. "Helping Counts: Predicting Children's Intentions to Disclose Being Bullied to Teachers from Prior Social Support Experiences." *British Educational Research Journal* 39 (2): 209–221. doi:10.1080/01411926.2011.627420.

Boulton, M. J., and P. K. Smith. 1994. "Bully/Victim Problems in Middle□school Children: Stability, Self□perceived Competence, Peer Perceptions and Peer Acceptance." *The British Journal of Developmental Psychology* 12 (3): 315–329. doi:10.1111/j.2044-835X.1994.tb00637.x.

Boulton, M. J., P. K. Smith, and H. Cowie. 2010. "Short-Term Longitudinal Relationships Between Children's Peer Victimization/Bullying Experiences and Self-Perceptions: Evidence for Reciprocity." *School Psychology International* 31 (3): 296–311. doi:10.1177/0143034310362329.

Boulton, M. J., M. Trueman, and K. Rotenberg. 2007. "User Perceptions of Process–Outcome Linkages in Pupil Peer Counselling for Bullying Services in the UK." *British Journal of Guidance & Counselling* 35 (2): 175–187. doi:10.1080/03069880701256585.

Boyle, C., L. Lynch, A. Lyon, and C. Williams. 2011. "The Use and Feasibility of a CBT Intervention." *Child and Adolescent Mental Health* 16 (3): 129–135. doi:10.1111/j.1475-3588.2010.00586.x.

Carter, S. L. 2010. *The Social Validity Manual: A Guide to Subjective Evaluation of Behavior Interventions*. London: Academic Press.

Cheon, S. H., and J. Reeve. 2015. "A Classroom-Based Intervention to Help Teachers Decrease students' Amotivation." *Contemporary Educational Psychology* 40: 99–111. doi:10.1016/j.cedpsych.2014.06.004.

Cowan, R. J., and S. M. Sheridan. 2003. "Investigating the Acceptability of Behavioral Interventions in Applied Conjoint Behavioral Consultation: Moving from Analog Conditions to Naturalistic Settings." *School Psychology Quarterly* 18 (1): 1. doi:10.1521/scpq.18.1.1.20877.

Cowie, H. 2000. "Bystanding or Standing By: Gender Issues in Coping with Bullying in English Schools." *Aggressive Behavior: Official Journal of the International Society for Research on Aggression* 26 (1): 85–97. doi:10.1002/(SICI)1098-2337(2000)26:1<85:AID-AB7>3.0.CO;2-5.

Cowie, H. 2011. "Peer Support as an Intervention to Counteract School Bullying: Listen to the Children." *Children & Society* 25 (4): 287–292. doi:10.1111/j.1099-0860.2011.00375.x.

Cowie, H., and L. Berdondini. 2001. "Children's Reactions to Cooperative Group Work: A Strategy for Enhancing Peer Relationships Among Bullies, Victims and Bystanders." *Learning and Instruction* 11 (6): 517–530. doi:10.1016/S0959-4752(01)00006-8.

Cowie, H., P. K. Smith, M. J. Boulton, and R. Lava. 1994. *Cooperation in the Multi-Ethnic Classroom: The Impact of Cooperative Group Work on Social Relationships in Middle Schools*. London: David Fulton.

Coyne, M. D., B. Simonsen, and M. Faggella-Luby. 2008. "Cooperating Initiatives: Supporting Behavioral and Academic Improvement Through a Systems Approach." *TEACHING Exceptional Children* 40 (6): 54–59. doi:10.1177/004005990804000606.

Cunningham, C. E., H. Rimas, S. Mielko, C. Mapp, L. Cunningham, D. Buchanan, T. Vaillancourt, Y. Chen, K. Deal, and M. Marcus. 2016. "What Limits the Effectiveness of Antibullying Programs? A Thematic Analysis of the Perspective of Teachers." *Journal of School Violence* 15 (4): 460–482. doi:https://doi.org/10.1080/15388220.2015.1095100.

Cunningham, C. E., T. Vaillancourt, L. J. Cunningham, Y. Chen, and J. Ratcliffe. 2011. "Modeling the Bullying Prevention Program Design Recommendations of Students from Grades Five to Eight: A Discrete Choice Conjoint Experiment." *Aggressive Behavior* 37 (6): 521–537. doi:10.1002/ab.20408.

DePaulo, B. M., J. Tang, W. Webb, C. Hoover, K. Marsh, and C. Litowitz. 1989. "Age Differences in Reactions to Help in a Peer Tutoring Context." *Child Development* 60 (2): 423–439. doi:https://doi.org/10.2307/1130987.

Doumas, D. M., A. Midgett, and A. D. Watts. 2019. "A Pilot Evaluation of the Social Validity of a Bullying Bystander Program Adapted for High School Students." *Psychology in the Schools* 56 (7): 1101–1116. doi:10.1002/pits.22249.

Festinger, L. 1957. *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Frey, A. J., K. Lee Park, T. Browne-Ferrigno, and T. L. Korfhage. 2010. "The Social Validity of Program-Wide Positive Behavior Support." *Journal of Positive Behavior Interventions* 12 (4): 222–235. doi:10.1177/1098300709343723.

Gaffney, H., D. P. Farrington, and M. M. Ttofi. 2019. "Examining the Effectiveness of School-Bullying Intervention Programs Globally: A Meta-Analysis." *International Journal of Bullying Prevention* 1 (1): 14–31. doi:10.1007/s42380-019-0007-4.

Gaffney, H., M. M. Ttofi, and D. P. Farrington. 2021. "What Works in Anti-Bullying Programs? Analysis of Effective Intervention Components." *Journal of School Psychology* 85: 37–56. doi:10.1016/j.jsp.2020.12.002.

Hawker, D. S., and M. J. Boulton. 2000. "Twenty years' Research on Peer Victimization and Psychosocial Maladjustment: A Meta-Analytic Review of Cross-Sectional Studies." *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 41 (4): 441–455. doi:10.1111/1469-7610.00629.

Hieneman, M., and G. Dunlap. 2001. "Factors Affecting the Outcomes of Community-Based Behavioral Support: II. Factor Category Importance." *Journal of Positive Behavior Interventions* 3 (2): 67–74. doi:10.1177/109830070100300202.

Kazdin, A. E. 1980. "Acceptability of Alternative Treatments for Deviant Child Behavior." *Journal of Applied Behavior Analysis* 13 (2): 259–273. doi:10.1901/jaba.1980.13-259.

Kazdin, A. 1998. *Research Design in Clinical Psychology*. Boston, MA: Allyn and Bacon.

Kowalski, R. M., and S. P. Limber. 2013. "Psychological, Physical, and Academic Correlates of Cyberbullying and Traditional Bullying." *Journal of Adolescent Health* 53 (1): S13–20. doi:10.1016/j.jadohealth.2012.09.018.

Macaulay, P. J., M. J. Boulton, and L. R. Betts. 2019. "Comparing Early adolescents' Positive Bystander Responses to Cyberbullying and Traditional Bullying: The Impact of Severity and Gender." *Journal of Technology in Behavioral Science* 4 (3): 253–261. doi:10.1007/s41347-018-0082-2.

McClellan, D. E., and S. J. Kinsay. 1999. "Children's Social Behavior in Relation to Participation in Mixed-Age or Same-Age Classrooms." *Early Childhood Research and Practice* 1: 1–26.

Medway, F. J., and R. M. Baron. 1977. "Locus of Control and Tutor's Instructional Style as Determinants of Cross-Age Tutoring Effectiveness." *Contemporary Educational Psychology* 2 (3): 298–310. doi:10.1016/0361-476X(77)90033-9.

Merrell, K. W., B. A. Gueldner, S. W. Ross, and D. M. Isava. 2008. "How Effective are School Bullying Intervention Programs? A Meta-Analysis of Intervention Research." *School Psychology Quarterly* 23 (1): 26. doi:10.1037/1045-3830.23.1.26.

Oliver, C., and M. Candappa. 2003. *Tackling Bullying: Listening to the Views of Children and Young People*. Research Report RR400 for UK Department for Education and Skills.

Olweus, D. 1993. *Bullying at School: What We Know and What We Can Do*. Oxford: Blackwell.

Olweus, D. 2013. "School Bullying: Development and Some Important Challenges." *Annual Review of Clinical Psychology* 9 (1): 751–780. doi:10.1146/annurev-clinpsy-050212-185516.

Polanin, J. R., D. L. Espelage, J. K. Grotpeter, E. Spinney, K. M. Ingram, A. Valido, A. El Sheikh, C. Torgal, and L. Robinson. 2021. "A Meta-Analysis of Longitudinal Partial Correlations Between School Violence and Mental Health, School Performance, and Criminal or Delinquent Acts." *Psychological Bulletin* 147 (2): 115. https://psycnet.apa.org/doi/10.1037/bul0000314.

Rigby, K., and D. Bagshaw. 2003. "Prospects of Adolescent Students Collaborating with Teachers in Addressing Issues of Bullying and Conflict in Schools." *Educational Psychology* 23 (5): 535–546. doi:10.1080/0144341032000123787.

Robinson, D. R., J. W. Schofield, and K. L. Steers-Wentzell. 2005. "Peer and Cross-Age Tutoring in Math: Outcomes and Their Design Implications." *Educational Psychology Review* 17 (4): 327–362. doi:10.1007/s10648-005-8137-2.

Salmivalli, C. 2010. "Bullying and the Peer Group: A Review." *Aggression and Violent Behavior* 15 (2): 112–120. doi:10.1016/j.avb.2009.08.007.

Salmivalli, C. 2014. "Participant Roles in Bullying: How Can Peer Bystanders Be Utilized in Interventions?" *Theory into Practice* 53 (4): 286–292. doi:10.1080/00405841.2014.947222.

Samara, M., and P. K. Smith. 2008. "How Schools Tackle Bullying, and the Use of Whole School Policies: Changes Over the Last Decade." *Educational Psychology* 28 (6): 663–676. doi:10.1080/01443410802191910.

Slavin, R. E. 2010. "Co-Operative Learning: What Makes Groupwork Work?" In *The Nature of Learning: Using Research to Inspire Practice*, edited by H. Dumont, D. Istance, and F. Benavides, 161–178. Paris, France: OECD.

Stein, M. B., J. R. Walker, and D. R. Forde. 1996. "Public-Speaking Fears in a Community Sample: Prevalence, Impact on Functioning, and Diagnostic Classification." *Archives of General Psychiatry* 53 (2): 169–174. doi:10.1001/archpsyc.1996.01830020087010.

Stevens, J. 1992. *Applied Multivariate Statistics for the Social Sciences*. Hillsdale, New Jersey: Lawrence Erlbaum.

Stukas, A. A., E. G. Clary, and M. Snyder. 1999. "Service Learning: Who Benefits and Why?" *Social Policy Report: Society for Research in Child Development* 8 (4): 1–19. doi:10.1002/j.2379-3988.1999.tb00039.x.

Sugai, G., and R. R. Horner. 2006. "A Promising Approach for Expanding and Sustaining School-Wide Positive Behavior Support." *School Psychology Review* 35 (2): 245–259. doi:10.1080/02796015.2006.12087989.

Tabachnick, B. G., and L. S. Fidell. 2014. *Using Multivariate Statistics*. Harlow: Pearson Education.

Thompson, F., and P. K. Smith, (2011). The Use and Effectiveness of Anti-Bullying Strategies in Schools. UK Government Research Report DFE-RR098. Accessed 19 March 2014. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182421/DFE-RR098.pdf

Topping, K. J., D. Miller, P. Murray, S. Henderson, C. Fortuna, and N. Conlin. 2011. "Outcomes in a Randomised Controlled Trial of Mathematics Tutoring." *Educational Research* 53 (1): 51–63. doi:10.1080/00131881.2011.552239.

Ttofi, M. M., and D. P. Farrington. 2011. "Effectiveness of School-Based Programs to Reduce Bullying: A Systematic and Meta-Analytic Review." *Journal of Experimental Criminology* 7 (1): 27–56. doi:10.1007/s11292-010-9109-1.

Tzani-Pepelasi, C., M. Ioannou, J. Synnott, and D. McDonnell. 2019. "Peer Support at Schools: The Buddy Approach as a Prevention and Intervention Strategy for School Bullying." *International Journal of Bullying Prevention* 1 (2): 111–123. doi:10.1007/s42380-019-00011-z.

Veldman, M. A., S. Doolaard, R. J. Bosker, and T. A. B. Snijders. 2020. "Young Children Working Together. Cooperative Learning Effects on Group Work of Children in Grade 1 of Primary Education." *Learning and Instruction* 67: 101308. doi:10.1016/j.learninstruc.2020.101308.

Vilardo, B. A., G. J. DuPaul, L. Kern, and R. L. Hojnoski. 2013. "Cross-Age Peer Coaching: Enhancing the Peer Interactions of Children Exhibiting Symptoms of ADHD." *Child & Family Behavior Therapy* 35 (1): 63–81. doi:10.1080/07317107.2013.761043.

Williams, D. M., E. S. Anderson, and R. A. Winett. 2005. "A Review of the Outcome Expectancy Construct in Physical Activity Research." *The Society of Behavioural Medicine* 29 (1): 70–79. doi:10.1207/s15324796abm2901_10.

Witt, J. C., and N. S. Elliott. 1985. "Effectiveness of Classroom Management Strategies." In *Advances in School Psychology*, edited by T. R. Kratochwill, 251–288. Vol. 4. Hillsdale, NJ: Lawrence Earlbaum.