



# Helping School Students Deal with Peer Provocations and Avoid Hostile Attribution Bias with the CATZ Cross-Age Teaching Zone Intervention

Michael J. Boulton<sup>1</sup> · Peter J. R. Macaulay<sup>2</sup>

Accepted: 5 March 2024  
© The Author(s) 2024

## Abstract

Cross-age tutoring and co-operative group work have been shown to help student tutors and tutees acquire academic and non-academic skills and knowledge. A novel intervention (Cross-Age Teaching Zone, CATZ) that combined them was tested for its effects on student tutors' thinking skills associated with (i) dealing pro-socially with peer provocations and (ii) avoiding hostile attribution bias. Small co-operative groups of 11- and 15-year-old students ( $N=228$ ) designed a CATZ lesson on these themes and delivered it to younger students. The CATZ tutors, but not matched controls ( $N=189$ ), showed significant improvements on both outcome measures. Participants aged 9 to 15 years ( $N=469$ ) were also asked about: (1) their willingness to act as CATZ tutors/tutees, (2) how effective they think such CATZ activities would be, (3) how much they valued autonomy in how they might deliver CATZ, and (4) their relative preference for being taught by older students versus teachers. Overall, participants expressed positive views of CATZ, which also helped students learn patterns of thinking that can help them avoid aggressive and conflict behavior. This initial evidence on the effectiveness of CATZ calls for further research to use CATZ across a range of social, emotional, and behavioral domains to support its wider uptake in schools.

**Keywords** CATZ · Hostile attribution bias · Peer provocations · Intervention · Social validity

## Introduction

Student to student aggression and conflict is widespread (Wang et al., 2020) and compromises their educational, social, and emotional development (Briesch & Chafouleas, 2009; Pellegrini & Blatchford, 2013). It may manifest in diverse ways including bullying/peer victimization (Monks et al., 2021), and aggression and fighting that is often provoked by peers (Boulton, 1993; Wang et al., 2020). Student aggression and conflict is a serious issue in schools because it disturbs teaching and learning (Boulton et al., 2008; Sullivan et al., 2014).

Many approaches to tackling aggression and conflict aim to teach students self-regulation skills (Eisenberg et al., 2010; Zins et al., 2007). They are predicated on a group of theories that emphasize social cognitive processes (e.g.,

thoughts, beliefs) and the notion that 'thinking errors' precipitate aggression and conflict (Dodge, 2006; Lochman & Wells, 2002). An exemplar is hostile attributional bias (hostile bias). Here, a person decides that they have been treated badly and with hostile intent even though the actual 'provocation' was ambiguous, and hence feels that a hostile 'retaliation' is justified (Dodge, 2006). As predicted by these theories, teaching positive thinking skills, especially in relation to peer provocations, has been found to reduce aggression and conflict (Sukhodolsky et al., 2004).

While most aggression and conflict are low intensity, even this can lead to more serious anti-social behavior (Goldstein, 1999). Moreover, many students often engage in low-level ambiguously provocative behavior that is ill-received by recipients and precipitates hostile bias, aggression and conflict (Boulton, 1993; Goldstein, 1999). Hence, there is value to be had in assisting *all* students to become aware of helpful thinking skills and the need to avoid hostile bias, especially in relation to the everyday provocations from peers that are a common part of school life. Such a view is consistent with a tiered approach to supporting good behavior within schools, in which tier 1 is recommended to facilitate social-emotional competence

✉ Peter J. R. Macaulay  
p.macaulay@derby.ac.uk

<sup>1</sup> School of Psychology, University of Chester, Parkgate Road, Chester CH1 4BJ, UK

<sup>2</sup> School of Psychology, University of Derby, Kedleston Road, Derby DE22 1GB, UK

across the entire population of students (Mayworm & Sharkey, 2014). While several meta-analyses attest to the benefits of a ‘thinking skills’ approach (Beck & Fernandez, 1998; Bennett & Gibbons, 2000; Sukhodolsky et al., 2004; Wilson et al., 2003), effect sizes are often relatively modest. Hence, there remains a need for alternative forms of tier 1 thought-based interventions to be evaluated.

### Role of Peers in Intervention Delivery

Considerable evidence and theory support the use of students to assist teachers in the delivery of academic and pastoral curricula and interventions (Baines et al., 2007; Karcher, 2008; Slavin, 1983; Topping et al., 2011). Pastoral curriculum and interventions focus on supporting the well-being of students in school from physical, social, emotional, and psychological care. Such interventions may vary in terms of the number of students involved (e.g., dyads versus groups), their age relationship (e.g., same versus cross-age) and the key aims (e.g., to learn a specific skill versus bolstering self-esteem). It is now clear that their unique features mean that the processes at work within them, and the kinds of issues they are best suited to addressing, are likely to vary (Karcher, 2007). Hence, there is value to be had in studying the effects of specific *formats* of student teaching in specific *learning contexts*.

*Co-operative group work* has been shown to assist students’ learning in academic (Baines et al., 2007; Slavin, 1983, 2010; Veldman et al., 2020) and non-academic (e.g., social/behavioral) domains (Blatchford et al., 2006; Cowie et al., 1994), including anti-bullying learning (Ttofi & Farrington, 2011). Similarly, *cross-age* approaches have been shown to benefit tutors’ literacy and numeracy development (Karcher, 2007, 2008; Robinson et al., 2005; Topping et al., 2011), and their self-esteem and social competence (Robinson et al., 2005; Watts et al., 2019). Given these positive *but separate* results for co-operative group work and cross-age teaching across such a wide variety of domains and variables, we developed an approach that combined them. The present study tested if this Cross-Age Teaching Zone intervention (CATZ) could be used to teach student tutors helpful aggression and conflict-related thinking skills, and the social validity of the CATZ intervention.

### The Cross-Age Teaching Zone Intervention

Several theoretical and empirical considerations led us to focus on the effects of CATZ on tutors rather than tutees. Acting as a CATZ tutor provides multiple opportunities for cognitive restructuring or elaboration, as students work with the lesson material, make links with what they already know and hence develop more advanced cognitive structures and schemas (Slavin, 1996; Thurston et al., 2007; Topping &

Ehly, 1998). In terms of Vygotsky’s sociocultural theory (1978), the fact that tutors are required to re-work the learning material that is provided to them into their own viable lesson, means tutors are likely to be working in the zone of proximal development, that is just outside what they can ‘do’ unaided. In our case, that means tutors will likely be ‘thinking about thinking’ in novel ways. Such notions fit well with the concept of metacognition. Although this has been defined in many ways (Dinsmore et al., 2008), most scholars have utilized Flavell’s (1979) seminal work that sees metacognition literally defined as “thinking about thinking” (see Dinsmore et al., 2008, p. 393). It consists in large part of a person’s ability to reflect on and monitor what they are learning. Importantly, Dinsmore et al. (2008) noted that metacognition implicates “a marriage between self-awareness and intention to act” (p. 404), and this supports our notion that activities that promote thinking about thinking in response to provocations will have practical value to CATZ tutors. Indeed, there is evidence that interventions that build metacognition lead to positive behavior change among school students (Holder et al., 2008; Whetstone et al., 2015). Thus, the notion that CATZ can promote metacognition within the domain of social and emotional learning and behavior seems reasonable on both theoretical and empirical grounds.

The fact that in CATZ, tutors are working *co-operatively* to develop and deliver their lesson may further optimize the likelihood that they will learn the lesson material (Slavin, 1996). Slavin (1996) argued that such co-operative activities provide ‘implicit’ reward and incentive structures; again, in our case, tutors are likely to see that they have a responsibility to their group that can be met if they themselves master the lesson material. Role theory also suggests that acting as a teacher promotes that sense of responsibility even more because it engenders a sense of care towards tutees (Biddle, 1986; Robinson et al., 2005). Finally, and with specific reference to helping students learn useful new ways of thinking, it is now apparent that approaches that do not seek to *directly* ‘challenge and change’ existing thought patterns can be effective (Longmore & Worrell, 2007). Thus, having CATZ tutors work on material about aggression and conflict-related thoughts in a general sense, in the absence of direct attempts to change them, could be sufficient for them to change the way they think about responding to provocations and to avoid hostile bias.

### Social Validity of Interventions

Evaluations of interventions to address aggression and conflict often neglect social validity or the extent to which *students* regard them as acceptable (Carter, 2010; Daunic et al., 2006). Most studies have assessed social validity with multi-item standardized measures (such as the Children’s

Intervention Rating Profile), largely because 'overall acceptability' scores can be derived from those multiple items on the basis of statistical internal reliability and construct validity assessments (Carter, 2010). However, a focus on psychometrics may be at the expense of practical value because overall scores may reveal little about consumers' views of *specific* aspects of an intervention and the different ways it can be delivered (Carter, 2010). This has led some researchers to eschew 'standard though general' measures of social validity in favor of several issue-specific measures that each can be operationalized with a single item (Boyle et al., 2011; Cheon & Reeve, 2015; Nickerson et al., 2014).

Studies have reported generally positive student views of peer *or* cross-age teaching (Boyle et al., 2011; Cunningham et al., 2016; Topping & Bryce, 2004; Willis et al., 2012). However, these investigations did not ask direct questions about social validity, and moreover, they were not focused on CATZ per se with its co-operative *and* cross-age teaching characteristics. One study that did so found substantial student support (Boulton & Macaulay, 2023). As Boulton and Macaulay (2023) noted, the specific 'topic' that is addressed via CATZ is important because it could affect acceptability ratings and limit generalizability; low acceptability ratings could be because students don't want to learn about the particular topic, or because they don't want to engage in CATZ, or any combination of the two. The present study investigated the social validity of CATZ as it was used to teach students about helpful aggression and conflict-related thoughts.

Another key issue for interventions targeting aggression and conflict is *who* delivers them. Anti-bullying lessons delivered by teachers are not always well-received by students (Boulton & Boulton, 2012; Rigby & Bradshaw, 2003), and when asked who they would prefer to teach them about avoiding engaging in bullying, around 90% of a sample of 817 9–15-year-olds chose older students over teachers (Boulton & Macaulay, 2023). With regards to CATZ as an anti-bullying intervention, Boulton and Macaulay (2023) found that students wanted autonomy to choose the content of the material to be delivered in their lesson. Other studies have also found that providing more autonomy was associated with higher acceptability of group based and peer-led interventions (Blatchford et al., 2003; Stukas et al., 1999). Similarly, Boulton and Macaulay (2023) reported that students rated freedom to choose who they worked with as highly important, a finding echoed in other work on students' views of peer-led interventions (Boulton, 2005; Cowie et al., 1994). These findings are consistent with other studies that have shown how allowing participants in diverse intervention programs freedom to choose some of the parameters of that intervention added to its effectiveness (Shogren et al., 2004). Moreover, allowing such choices is consistent with promoting self-regulation in the general population of school students (Eisenberg et al., 2010; Mayworm & Sharkey, 2014).

Boulton and Macaulay (2023) and Boulton et al. (2023) reported no significant differences in students' social validity views of CATZ as an anti-bullying intervention as a function of gender. However, girls tend to be more open to acting as peer supporters than boys (Boulton, 2005; Cowie, 2000; Cowie et al., 2002; Naylor & Cowie, 1999) and this suggests some/many of them may go on to hold more favorable beliefs about CATZ than boys. In terms of age, DePaulo et al. (1989) reported that the age of tutors appeared to be of greater concern to 10-year-olds than to 8-year-olds, but Boulton and Macaulay (2023) found no significant differences in children aged 9–15 years. These inconsistent findings suggest it is worthwhile to further examine age and gender differences.

## The Current Study

In designing this study, we took account of teachers' desire for: (i) brief interventions (Boulton, 2014; Chafouleas et al., 2009; Witt, 1986) and (ii) evidence that potential interventions they might utilize will be more effective than their current 'standard' practice (Boulton, 2014). Our own discussions with teachers confirmed how important these two factors were. Hence, we designed CATZ to be as short as possible while still allowing tutors sufficient time to develop a good lesson for their tutees, *and* we employed a control group whose experiences reflected the way teachers said they would normally deliver information to help students regulate their own anti-social behavior (i.e., via disseminating information to them in class to stimulate discussion and personal reflection). The primary aim of the current study was to test the CATZ intervention on the effects on student tutors' thinking skills associated with (i) dealing pro-socially with peer provocations and (ii) avoiding hostile attribution bias. Thus, we compared the effects of CATZ with that of a control condition that comprised 'business as usual' practice rather than with an alternative intervention that was of similar duration to CATZ and that also went beyond standard practice. The control condition also allowed us to investigate if any positive effects of CATZ were due simply to repeated testing effects. The implications of this aspect of our design are discussed later.

In the present study, we hypothesized that through a CATZ intervention in which student tutors develop and deliver a lesson about helpful thoughts and hostile attribution bias to younger tutees, their own capacity to generate helpful thoughts in the face of typical peer provocations and avoid hostile attribution bias would be facilitated. As a secondary aim, we also tested the social validity of the CATZ intervention. We asked students about their willingness to act as CATZ tutors/tutees, their beliefs about the likely effectiveness of CATZ, their importance ratings of three dimensions of autonomy (choosing who they worked with, the

lesson content, and the mode of delivery of the lesson) and their relative preference for being taught by older students versus teachers. Finally, we also tested for age and gender differences for all the above effects/variables.

## Method

### The CATZ Intervention

The CATZ intervention was developed by the lead author on the basis of considerable pilot work, other evaluation studies (Boulton et al., 2023) and cognitive and social development considerations outlined elsewhere (Blatchford et al., 2006; Cowie et al., 1994). For instance, we took account of the size of groups that children and young people are typically able to work effectively within (around five people) for the tutors, and the length of the CATZ lesson that tutees could be expected to attend to (around 30 min with multimedia stimulation). We implemented the intervention in the schools during Personal and Social Education lessons. Teachers were present only as observers. We encouraged 'buy-in' by explaining to potential CATZ tutors that taking part was voluntary, they could stop at any time (and re-join) without giving a reason, they were being invited to work in small groups of about five students to design a (roughly) 30-min lesson about how people could get on better with each other, deal with provocations and avoid aggression and serious conflict, and to deliver it to a small group of younger students. We stressed that this was an important task because the lesson could help the younger students learn important things, and that they might actually enjoy taking part. Because students have perceived adult-implemented initiatives to tackle aggression and conflict as 'boring' (Boulton & Boulton, 2012), we tried to engender a sense of fun and ownership of the lesson that complemented tutors' sense of responsibility. Tutors were informed that we would provide them with the lesson content, and offer suggestions about how to plan, test and deliver a lesson, but that the details would be left to them. We aimed to strike a balance between being suitably supportive on one hand and leaving tutors to take ownership of their lesson on the other. While the final say on the lesson itself was left to each group of tutors, we ensured that as a minimum, they all designed a *poster* that contained the lesson material, and prepared a *script* of what was to be said and done by each group member during their lesson.

A summary sheet of the lesson material was provided to each participant and discussed with them as a whole class and on a group-by-group basis as CATZ progressed. This drew a distinction between 'helpful thoughts' that help keep people calm in the face of provocations and de-escalate conflict, and 'unhelpful thoughts' that tend to precipitate

anger and escalate conflict. The nature of hostile bias was described along with the notion that decisions about 'hostile intent' often have to be inferred, and that those inferences could be incorrect. Points were illustrated with scenarios and real-life video examples from the internet. As we monitored the development of each CATZ tutor groups' lesson, we reminded all of them of the need to address these key issues. Prior to the delivery of the lessons, we confirmed that all of them had done so by ensuring that all of the material on the summary sheet that was provided to the tutors was included in their lesson.

Importantly, at no point did we state or even imply that we 'wanted' the tutors to learn this information, or that tutors themselves needed to change. Rather, tutors were reminded that this was the information they would help the younger tutees learn. Tutors had four roughly 60-min sessions to prepare their lesson, spread over about three weeks. Then, within a few days of their final 'dress rehearsal', they delivered their lesson. Each tutor group consisted of about five people, and they delivered their lesson to a similar sized group of tutees. Some activities included in the CATZ lessons including videos, handouts, quizzes, and puzzles.

### The Control Condition

Corresponding with the final week of the CATZ intervention, we went into the Control classes for a one-hour session during a Personal and Social Education lesson. After being introduced by the class teacher who did not take an active role beyond this, a summary sheet of the lesson material was provided to each student and we engaged them in a class discussion about the issues in the same way as happened with the CATZ participants, using the same scenarios and video clips. This was done to try to make their exposure to the lesson material about helpful thoughts and hostile bias as similar as possible to that of the CATZ participants, with the exception that CATZ participants then incorporated that material into their own lesson for the tutees. Other than this, Control participants carried on with their normal school activities during the time CATZ tutors were preparing/delivering their lesson.

### Participants and Procedure

Participants were drawn from three public junior schools and three public high schools, selected on a convenience basis. They served a small city in the north of the UK. The ethnic composition of the sample was 92% white and 8% black and minority students, and this mirrors the composition in the local area of 94% white and 6% black and minority citizens. The schools all had a mixed catchment area in terms of socio-economic status. Consent was solicited from all of the students in the selected year/class groups for whom prior

parental or head teacher permission in their *loco parentis* role had been obtained. The response rate was 93.3%. It is helpful to think of two partially overlapping sub-sets of participants based on the data they provided for (i) social validity (Years 4, 6, 7 & 10) and (ii) evaluation of CATZ (Years 6 & 10 only). This information is summarized in Table 1. Social validity data were collected from 249 Year 4 and Year 6 ( $M_{\text{age}} = 9.5$  and 11.5 years, respectively) participants from the junior schools (53.4% girls), and 220 Year 7 and Year 10 ( $M_{\text{age}} = 12.5$  and 15.5 years, respectively) participants from the high schools (51.4% girls) ( $N = 469$ ). Those students in Years 6 and 10 formed part of the pool of participants who were recruited to act as CATZ or Control participants, along with an additional 85 Year 6 junior school students and 118 Year 10 high school students who had not provided the social validity data. Hence, measures of the two thoughts variables used for the assessment of CATZ were collected from 417 participants in Years 6 and 10. Randomization was at the class level. Of this pool of 417 participants, 228 were in classes randomly allocated to the CATZ condition, and of these, 108 were from Year 6 at junior school (51 girls) and 120 were from Year 10 at high schools (45 girls). Of the 189 who were in classes randomly allocated to the Control condition, 92 were from Year 6 at junior school (41 girls) and 97 were from Year 10 at high schools (51 girls). We confirmed the randomization process had produced groups that were similar by showing that the CATZ and Control participants did not differ on any of our variables at the first time of testing, all  $p$ 's > 0.05.

Baseline data were collected prior to CATZ/Control experiences at Time 1 (T1) via the administration of a single self-report questionnaire containing all of our measures (plus others not reported here that measured state self-esteem, attitudes to older/younger students, and self-efficacy for group work and communication skills). It took about 20 min to complete. Then Year 6 and 10 CATZ tutors experienced CATZ over a period of about three weeks (see below). In the final week of this block, the Control participants received their input from the researchers (see below). Time 2 and

time 3 (T2 and T3) data were collected about four weeks and eight weeks, respectively, after T1 (i.e., about one week and four weeks, respectively, after the CATZ/Control intervention had ended). The same questionnaire that was used at T1 was also employed here, with the exception that the social validity variables were not included. Hence, social validity data were collected at T1 only.

At each time of testing, data were collected on a whole class basis during Personal, Social and Health Education classes. Participants were informed that they were not being tested and as there were no right or wrong answers, and there was no need to copy from other students. We stressed that we were interested in their personal views so that they could be combined and allow us to get an idea of what many students think. Each participant was given a questionnaire, asked to listen as a researcher read out the instructions followed by each question in turn and then given time to respond. This ensured that all students, including those who were not good readers, could participate. Class teachers were present at all times because head teachers insisted on it for safeguarding reasons. In practice, all teachers played no active role in any of our research-related activities and most used the time we were in their class to prepare their teaching and/or mark student work, etc.

At T1, to ensure a shared understanding and enable considered responses to the social validity items, we read out a standard definition of CATZ: "CATZ stands for cross-age teaching zone. It is where older students act as teachers to younger students. The older students work together to design a lesson in a small group about something that they would not normally do with their teacher in a lesson. That is, not things like maths, history, science or English but things about people and how they can behave better with each other. Things like how to avoid having fights and serious arguments and being nasty to one another and how to cooperate more with each other. Adults help them do this, but it is very much the older student's lesson. They are in charge. They then go into a class of younger students to give their lesson to a small group. The adult teachers are always there to keep an eye on things, but they are not giving the lesson".

**Table 1** Participant Details in Each Sub-group/Condition

| Variable measured                 | Year                          |   |   |   |
|-----------------------------------|-------------------------------|---|---|---|
|                                   | 4                             | 6   | 7                                       | 10  |
| Social validity                   | 134 (69 girls)<br>(5 classes) | 115 (64 girls)<br>(4 classes)               | 121 (65 girls)<br>(5 classes)           | 99 (48 girls)<br>(4 classes)                |
| Helpful Thoughts/<br>Hostile Bias |                               | 200 (above)<br>115 + 85 from<br>(3 classes) |   | 217 (above)<br>99 + 118 from<br>(4 classes) |
|                                   |                               | CATZ<br>108 (51 girls)<br>(4 classes)       | Control<br>92 (41 girls)<br>(3 classes) | CATZ<br>120 (45 girls)<br>(4 classes)       |
|                                   |                               |   |   | Control<br>97 (51 girls)<br>(4 classes)     |



## Measures and Reliability

**Reliability** To assess test–retest reliability, we administered the questionnaire to a sub-set of our non-CATZ participants one week after initial data collection ( $n$ 's = 40, 37, 27 and 24 from Years 4, 6, 7 and 10, respectively,  $n = 128$ ). Schools were reluctant to release students for anything other than data collection but agreed to do so with this sub-group because they were available for this extra testing during a convenient registration period outside of lesson time. Other participants were involved in other school-related activities at this time. T-tests confirmed that the students who did provide reliability data did not differ from those who did not on any measure at T1, all  $p$ 's > 0.05. Test–retest correlations for helpful thoughts and hostile bias, and percentage agreement scores for each social validity item, are presented in brackets below. They all show high levels of test–retest reliability.

**Helpful thoughts in response to provocation (Helpful Thoughts)** This was measured with eight open questions about how the participant would react to some kind of typical peer provocation. Some questions were general and did not prompt thoughts per se (e.g., “If another student does something to you that you do not like, what would you do?”) and others did so (e.g., “Imagine that you are waiting in line for something good and another student pushes in front of you, what would be your first thought?”). A coding scheme was developed to classify responses as either showing or not showing helpful thoughts (i.e., thoughts that help keep a person calm and avoid an aggressive response or conflict). Examples of helpful thoughts we identified (in response to the questions given above) are “I would think they are having a bad day and so I would let it go” and “I would think they wanted to be with their friend and that I might do that too sometimes”. Two researchers independently coded all of the responses to all of the questions, and we obtained 94% agreement. Disagreements were discussed until coders agreed. Helpful Thoughts scores were summed across the items and could range from 0–8 (test–retest  $r = 0.91$ ,  $p < 0.001$ ), with the highest score reflecting that the participant had offered a helpful thought on all eight open questions and the lowest score reflecting that the participant had offered a helpful thought on none of them.

**Hostile Bias** As in Cillessen et al. (2014), this was assessed with three vignettes. An example is, “Imagine that you leave your desk for a short while and when you come back another student had spilt a drink over your work. Would they have done this on purpose to be nasty to you?” Response options were “No, Can't be sure, and Yes”, coded 0, 1 and 2, respectively. Responses were summed and Hostile Bias scores could range from 0–6 (test–retest  $r = 0.89$ ,  $p < 0.001$ ), with high scores reflecting more hostile bias.

**Acceptability of CATZ** This was assessed with three items, (i) “How much would you like to give a CATZ lesson to younger students, I mean to be the teacher?” (test–retest 93.3%), (ii) “How much would you like to have a CATZ lesson, I mean for older students to teach you?” (test–retest 95.2%), and (iii) “How much do you think CATZ would help young students learn new things about people and how they can behave better with each other?” (test–retest 97.1%). Response options were “A lot, A bit, or Not at all”. As noted above, given that these all tap different aspects of acceptability, responses from each question were analyzed separately.

**Importance of autonomy** This was measured with three questions that began with the stem, “Thinking about you giving a CATZ lesson, how important is it that you choose . . . .” followed by, (i) “who you work with?” (test–retest 95.2%), (ii) “what goes in to your CATZ lesson?” (test–retest 88.6%), and (iii) “how to give the CATZ lesson?” (test–retest 96.2%). Response options were “A lot, A bit, or Not at all”. Again, responses from each question were analyzed separately.

**Relative preference for CATZ versus teachers** This was assessed with, “Who would you prefer to teach you new things about people and how they can behave better with each other?” (test–retest 95.2%). Response options were “Older students with CATZ, Don't mind, and Teachers”.

## Plan of Analysis

There was some movement, considerable for some pupils, between members of the different CATZ groups. Hence group data were not independent and so multi-level analyses were not appropriate. As in some previous studies to evaluate programs (Boulton et al., 2023; Pereira et al., 2014), the effects of CATZ, and tests of gender and age as moderators, were analyzed in 2 (Condition)  $\times$  3 (Time, repeated measures)  $\times$  2 (Gender)  $\times$  2 (Year) mixed analysis of variance (ANOVA) tests, one each for Helpful Thoughts and Hostile Bias. Partial eta squared ( $\eta^2$ ) was used as an index of effect size. Post-hoc one-way ANOVA and t-tests were used to identify sub-group differences, with Bonferroni corrections to control for family-wise inflation of Type I errors. Only significant effects are reported.

On Helpful Thoughts and Hostile Bias, Jacobson and Truax's (1991) Reliable Change Index identified participants deemed to have improved (T1 to T2, and T1 to T3) because of the intervention rather than as a result of chance or repeated testing effects. Using a 95% confidence interval, Reliable Change Index scores above 1.96 indicate reliable change, corresponding to an improvement of 3 or more on each of our DVs. Chi-square ( $\chi^2$ ) tests of association compared the proportion of CATZ versus Control participants that did/did not show reliable change. Log-linear analysis

tested if gender or age moderated this effect (Stevens, 1992; Tabachnick & Fidell, 2014). Only significant effects are reported.

For the social validity variables, key data are the numbers/percentages of students who chose the different response options for each question, summarized in Table 2. For each question, a  $\chi^2$  goodness of fit test determined if there were significant departures from chance in the proportion of students overall that selected each of the three response options. Then, age and gender differences on each question were tested initially by means of a log-linear analysis. On no question did the likelihood ratio  $G^2$  or the Pearson model selection criteria indicate that the best fitting model contained the three-way interaction. Hence, gender differences, and separately, year differences, were tested in two-way  $\chi^2$  tests of association. A Bonferroni correction was applied to the alpha level to control for family-wise type 1 errors. Only significant results are reported.

## Results

### Effects of CATZ and Tests of Gender and Age as Moderators

For interpreting the Results, T1 is pre-test before CATZ intervention was administered, T2 is 1-week post CATZ intervention, and T3 is 4-week post CATZ intervention. Mean (and standard deviation) Helpful Thoughts and Hostile Bias scores are presented in Table 3. There was a significant Time x Condition interaction on both Helpful Thoughts and Hostile Bias, Wilk's Lambda = 0.62,  $F(2, 402) = 123.39$ ,  $\eta^2 = 0.38$ ,  $p < 0.001$ , and Wilk's Lambda = 0.88,  $F(2, 402) = 26.35$ ,  $\eta^2 = 0.12$ ,  $p < 0.001$ , respectively. In both cases, whereas the Controls showed no significant change on Helpful Thoughts or Hostile Bias across time, all  $p > 0.10$ , the CATZ participants did, for Helpful Thoughts, Wilk's Lambda = 0.34,  $F(2, 226) = 219.93$ ,  $\eta^2 = 0.66$ ,  $p < 0.001$ , and for Hostile Bias, Wilk's Lambda = 0.74,  $F(2, 226) = 40.30$ ,  $\eta^2 = 0.26$ ,  $p < 0.001$ . In exploring the pattern of this change, CATZ participants showed a significant *increase* in the number of Helpful Thoughts from T1 (mean = 0.44) to T2 (mean = 3.55), and from T1 to T3 (mean = 2.60), both  $p < 0.001$ , but a significant *decline* from T2 to T3 ( $p < 0.001$ ). Among CATZ participants, there was a significant *reduction* in Hostile Bias from T1 (mean = 1.72) to T2 (mean = 0.90), and from T1 to T3 (mean = 1.00), both  $p < 0.001$ .

In testing if the effect of CATZ was moderated, for Helpful Thoughts, the Time x Condition x Year interaction was significant, Wilk's Lambda = 0.87,  $F(2, 402) = 30.23$ ,  $\eta^2 = 0.13$ ,  $p < 0.001$ . While there was a significant *reduction* in Helpful Thoughts from T2 to T3 among both CATZ

age groups (both  $p < 0.001$ ), the reduction was larger among the Year 10 s (means = 3.28 and 1.93, respectively), than among the Year 6 s (means = 3.81 and 3.32, respectively). However, it is important to note that in both age groups of CATZ tutors, T3 Helpful Thoughts scores were still significantly *higher* than at T1 (both  $p < 0.001$ ).

Although not related to the assessment of the effects of CATZ, we report that for Hostile Bias there was no significant gender difference,  $F(1, 403) = 2.71$ , and that Year 6 students (mean = 1.76) scored significantly higher than Year 10 students (mean = 1.31),  $F(1, 403) = 18.40$ ,  $p < 0.001$ .

### Reliable Change

On Helpful Thoughts, 150 out of 228 (66%) CATZ tutors evidenced a reliable improved from T1 to T2 compared to only 7 out of 189 (4%) Control participants, a significant difference,  $\chi^2(1) = 144.1$ ,  $p < 0.001$ . From T1 to T3, 96 out of 228 (42%) CATZ tutors improved compared to only 6 out of 189 (3%) Control participants, a significant difference,  $\chi^2(1) = 129.9$ ,  $p < 0.001$ .

On Hostile Bias, 48 out of 228 (21%) CATZ tutors improved from T1 to T2 compared to 5 out of 189 (3%) Control participants, a significant difference,  $\chi^2(1) = 29.92$ ,  $p < 0.001$ . From T1 to T3, 37 out of 228 (16%) CATZ tutors improved compared to only 7 out of 189 (4%) Control participants, a significant difference,  $\chi^2(1) = 15.87$ ,  $p < 0.001$ .

### Acceptability of CATZ

Across the three items, the vast majority of participants as a whole indicated high acceptability of CATZ. A significantly higher proportion than chance selected 'A lot' when asked how much they (i) would like to act as a CATZ tutee (93.6%),  $\chi^2(2) = 767.9$ ,  $p < 0.001$ , (ii) would like to act as a CATZ tutor (53.9%),  $\chi^2 = 120.4$ ,  $p < 0.001$ , and (iii) thought CATZ would help younger students learn new things, (91.3%),  $\chi^2(2) = 712.0$ ,  $p < 0.001$ .

### Importance of Autonomy

Across the three items, the vast majority of participants as a whole rated autonomy as highly important. A significantly higher proportion than chance selected 'A lot' when asked how important was choosing (i) who they worked with, (93.3%),  $\chi^2(2) = 762.9$ ,  $p < 0.001$ , (ii) what goes into their lesson, (68.7%),  $\chi^2(2) = 291.0$ ,  $p < 0.001$ , and (iii) how to give it, (68.7%),  $\chi^2(2) = 278.6$ ,  $p < 0.001$ . Additionally, on the latter two questions, proportionally more Year 7 (73.6% and 76.8%, respectively) and Year 10 (92.9% and 98.9%, respectively) students than expected, and fewer Year 4 (50.7% and 44.8%, respectively) and Year 6 (63.5% and 61.7%, respectively) students than expected, chose the 'A lot'

**Table 2** Frequency (Percentage in Brackets) of Participants' Responses to Questions about CATZ

|  | Year 4    |            |           | Year 6    |            |           | Year 7    |            |           | Year 10   |           |            | All       |
|--|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|-----------|------------|-----------|
|  | Girls     | Boys       | Both      | Girls     | Boys       | Both      | Girls     | Boys       | Both      | Girls     | Boys      | Both       |           |
|  | (n = 69)  | (n = 65)   | (n = 134) | (n = 64)  | (n = 51)   | (n = 115) | (n = 65)  | (n = 56)   | (n = 121) | (n = 48)  | (n = 51)  | (n = 99)   | (n = 469) |
| How much would you like to give a CATZ lesson to younger pupils, I mean to be the teacher?                                       |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 31 (44.9)  | 28 (43.1) | 58 (43.3)  | 41 (64.1) | 29 (56.9) | 70 (60.9)  | 40 (61.5) | 31 (55.4) | 71 (58.7)  | 24 (50.0) | 30 (58.8) | 54 (54.5) | 253 (53.9) |           |
| 29 (42.0)  | 20 (30.8) | 50 (37.3)  | 19 (29.7) | 16 (31.4) | 35 (30.4)  | 20 (30.8) | 15 (26.8) | 35 (28.9)  | 21 (43.8) | 16 (31.4) | 37 (37.3) | 157 (33.5) |           |
| 9 (13.0)   | 17 (26.2) | 26 (19.4)  | 4 (6.3)   | 6 (11.8)  | 10 (8.7)   | 5 (7.7)   | 10 (17.9) | 15 (12.4)  | 3 (6.3)   | 5 (9.8)   | 8 (8.1)   | 59 (12.6)  |           |
| How much would you like to have a CATZ lesson, I mean for older pupils to teach you?   |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 69 (100)   | 65 (100)  | 134 (100)  | 60 (93.8) | 51 (100)  | 111 (96.5) | 61 (93.8) | 46 (82.1) | 107 (88.4) | 41 (85.4) | 46 (90.2) | 87 (87.9) | 439 (93.6) |           |
| 0 (0)  | 0 (0)     | 0 (0)      | 4 (6.2)   | 0 (0)     | 4 (3.5)    | 4 (6.2)   | 7 (12.5)  | 11 (9.1)   | 7 (14.6)  | 3 (5.9)   | 10 (10.1) | 25 (5.3)   |           |
| 0 (0)  | 0 (0)     | 0 (0)      | 0 (0)     | 0 (0)     | 0 (0)      | 0 (0)     | 3 (5.4)   | 3 (2.5)    | 0 (0)     | 2 (3.9)   | 2 (2.0)   | 5 (1.1)    |           |
| How much do you think CATZ would help young pupils learn new things about people and how they can behave better with each other? |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 69 (100)   | 65 (100)  | 134 (100)  | 52 (81.3) | 39 (76.5) | 91 (79.1)  | 62 (95.4) | 49 (87.5) | 111 (91.7) | 48 (100)  | 44 (86.3) | 92 (92.9) | 428 (91.3) |           |
| 0 (0)  | 0 (0)     | 0 (0)      | 11 (17.2) | 10 (19.6) | 21 (18.3)  | 3 (4.6)   | 7 (12.5)  | 10 (8.3)   | 0 (0)     | 7 (13.7)  | 7 (7.1)   | 38 (8.1)   |           |
| 0 (0)  | 0 (0)     | 0 (0)      | 1 (1.6)   | 2 (3.9)   | 3 (2.6)    | 0 (0)     | 0 (0)     | 0 (0)      | 0 (0)     | 0 (0)     | 0 (0)     | 3 (0.6)    |           |
| How important is it that you choose who you work with?   |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 60 (86.9)  | 65 (100)  | 125 (93.3) | 58 (90.6) | 49 (96.1) | 107 (93.0) | 60 (92.3) | 53 (94.6) | 113 (93.3) | 44 (91.7) | 49 (96.1) | 93 (93.9) | 438 (93.3) |           |
| 9 (13.0)   | 0 (0)     | 9 (6.7)    | 6 (9.4)   | 2 (3.9)   | 8 (7.0)    | 4 (6.2)   | 3 (5.4)   | 7 (5.8)    | 2 (4.2)   | 1 (1.9)   | 3 (3.00)  | 27 (5.8)   |           |
| 0 (0)  | 0 (0)     | 0 (0)      | 0 (0)     | 0 (0)     | 0 (0)      | 1 (1.5)   | 0 (0)     | 1 (0.8)    | 2 (4.2)   | 1 (1.9)   | 3 (3.0)   | 4 (0.9)    |           |
| How important is it that you choose what goes in to your CATZ lesson?  |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 31 (44.9)  | 37 (56.9) | 68 (50.7)  | 40 (62.5) | 33 (64.7) | 73 (63.5)  | 48 (73.8) | 41 (73.2) | 89 (73.6)  | 44 (91.7) | 48 (94.1) | 92 (92.9) | 322 (68.7) |           |
| 31 (44.9)  | 23 (35.4) | 54 (40.3)  | 19 (29.7) | 15 (29.4) | 34 (29.6)  | 15 (23.1) | 10 (17.9) | 25 (20.7)  | 4 (8.3)   | 3 (5.9)   | 7 (7.1)   | 120 (25.6) |           |
| 7 (10.1)   | 5 (7.7)   | 12 (9.0)   | 5 (7.8)   | 3 (5.9)   | 8 (6.9)    | 2 (3.0)   | 5 (8.9)   | 7 (5.8)    | 0 (0)     | 0 (0)     | 0 (0)     | 27 (5.7)   |           |
| How important is it that you choose how to give the CATZ lesson?   |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 28 (40.6)  | 32 (49.2) | 60 (44.8)  | 38 (59.4) | 33 (64.7) | 71 (61.7)  | 49 (75.4) | 44 (78.6) | 93 (76.8)  | 47 (97.9) | 51 (100)  | 98 (98.9) | 322 (68.7) |           |
| 21 (30.4)  | 28 (43.1) | 49 (36.6)  | 20 (31.2) | 16 (31.4) | 36 (31.3)  | 14 (21.5) | 8 (14.3)  | 22 (18.2)  | 1 (2.1)   | 0 (0)     | 1 (1.1)   | 108 (23.0) |           |
| 20 (29.0)  | 5 (7.7)   | 25 (18.6)  | 6 (9.4)   | 2 (3.9)   | 8 (6.9)    | 2 (3.0)   | 4 (7.1)   | 6 (5.0)    | 0 (0)     | 0 (0)     | 0 (0)     | 39 (8.3)   |           |
| Who would you prefer to teach you new things about people and how they can behave better with each other?                        |           |            |           |           |            |           |           |            |           |           |           |            |           |
| 62 (89.9)  | 64 (98.5) | 126 (94.0) | 55 (85.9) | 49 (96.1) | 104 (90.4) | 60 (92.3) | 50 (89.3) | 110 (90.9) | 40 (83.3) | 43 (84.3) | 83 (83.8) | 423 (90.2) |           |
| 0 (0)  | 0 (0)     | 0 (0)      | 7 (10.9)  | 0 (0)     | 7 (6.1)    | 1 (1.5)   | 2 (3.6)   | 3 (2.5)    | 0 (0)     | 1 (1.9)   | 1 (1.0)   | 11 (2.3)   |           |
| 7 (10.1)   | 1 (1.5)   | 8 (6.0)    | 2 (3.1)   | 2 (3.9)   | 4 (3.5)    | 4 (6.2)   | 4 (7.2)   | 8 (6.6)    | 8 (16.7)  | 7 (13.7)  | 15 (15.2) | 35 (7.5)   |           |

Data were collected at Time 1 only



**Table 3** Mean (and Standard Deviation) Scores of CATZ and Controls at Each Assessment

|                  | Time 1                  | Time 2                | Time 3                  |
|------------------|-------------------------|-----------------------|-------------------------|
| Measure          | CATS Control            | CATS Control          | CATS Control            |
| Helpful Thoughts | 44 (.86) .43 (.93)      | 3.55 (2.11) 63 (.88)  | 2.60 (1.53) 71 (.78)    |
| Hostile Bias     | 1.72 (1.60) 1.97 (1.57) | 90 (1.19) 2.04 (1.48) | 1.00 (1.22) 2.15 (1.66) |

option,  $\chi^2(6) = 50.5, p < 0.001$  and  $\chi^2(6) = 89.9, p < 0.001$ , respectively.

### Relative Preference for CATZ Versus Teachers

Significantly more participants than chance, 90.2%, expressed a preference for being taught new things by CATZ tutors, and less than expected preferred teachers (7.5%) or had no preference (2.3%),  $\chi^2(2) = 684.1, p < 0.001$ .

### Discussion

A key aim of this study was to test the effects of the novel CATZ intervention on tutors' ability to generate helpful thoughts and avoid hostile bias in the face of typical peer provocations. Results indicated significant improvements on both variables among CATZ, but not Control, participants. These benefits were evident shortly after taking part in CATZ and also after one month. While this suggests the effects may persist, some caution is called for because we also found a significant reduction in Helpful Thoughts from T2 to T3 (but T3 Helpful Thoughts scores were still significantly higher than at T1). The latter effect was also found to be moderated by age in that the reduction in Helpful Thoughts 'gains' was stronger among Year 10 than Year 6 tutors. No other moderating effect of age or gender was found on these variables, suggesting that CATZ may be as helpful to girls and boys at both ages, with the possible exception for Year 10 students on Helpful Thoughts. It seems reasonable to suggest that assisting students to develop helpful thoughts is best started well before they reach their mid-teens.

### Effectiveness and Social Validity of CATZ

Some of our other findings also attest to the effectiveness and practical utility of CATZ. The effect size for 'improvements' in Hostile Bias scores (0.12) approached, and for Helpful Thoughts scores (0.38) exceeded, that deemed to represent a large effect (0.138). Moreover, substantial numbers of individual CATZ tutors, and significantly more CATZ than Control participants, showed reliable change on both of our outcomes. For example, two thirds of CATZ participants reported at least three more helpful thoughts immediately after CATZ than before it. Although the

number showing reliable change in Hostile Bias was lower than for Helpful Thoughts, it was still the case that some 21% of CATZ tutors had Hostile Bias scores at least three points lower after CATZ than before it. This is an important finding because Hostile Bias is known to be quite resistant to change (Cillessen et al., 2014). As with the results of the ANOVAs (with the one exception discussed above), these reliable change results did not indicate moderating effects of age or gender.

In terms of social validity, responses to questions about acceptability all indicated substantial positive support for CATZ, and the overwhelming majority of participants opted for older CATZ tutors over teachers to teach them about avoiding aggression and conflict. This might reflect students' lack of enthusiasm for being instructed by teachers about these kinds of issues that has been observed elsewhere (Boulton & Boulton, 2012) and so implicates CATZ as an alternative worth considering. There were no significant gender or year group differences on these variables. These findings indicating high social validity are consistent with those for other cross-age teaching interventions that have focused on literacy (Willis et al., 2012), leadership skills (Besnoy et al., 2016) and critical thinking skills (Topping & Bryce, 2004). They attest to the value students see in cross-age teaching, both as deliverers of learning to younger students and as consumers of teaching delivered by older students. At the same time, they raise other important questions for future studies to address. Why do students have such positive views? What is it about cross-age teaching that seems so attractive? Are there 'optimum' age gaps? In terms of the latter, we have demonstrated that a 2-year age difference works well at both primary and secondary school level, but others are worth exploring.

Our findings that suggest CATZ may have a similarly strong appeal across girls and boys aged 9 to 15 years parallels results of a study of acceptability of CATZ to teach students about anti-bullying issues and e-safety (Boulton et al., 2016, 2023). This seems important given that boys tend to be less open to delivering other diverse peer-led interventions than girls (Boulton, 2005; Cowie, 2000; Cowie et al., 2002; Naylor & Cowie, 1999). It might be because CATZ is a less formal, or 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 may be that giving a lesson to younger students allows boys (and girls) to recognize

their superior knowledge, status or power and to feel good about using this with possibly admiring younger tutees. The reasons why boys (and some girls) may be more open to CATZ than to other forms of peer support is a worthy topic of study in the future. So is the influence of the gender composition of the CATZ tutor groups and the tutees they work with, not least because gender has been found to influence how well students work in groups (Blatchford et al., 2006; Cowie et al., 1994). Future studies could compare and contrast same- and mixed-gender groupings, and also if the latter promote more co-operation and better social relationships between girls and boys that are often lacking (Cowie et al., 1994). Further, future research may also consider the popularity status of CATZ tutors and their involvement in bullying and victimisation. Such endeavours could be achieved via a combination of self-report measures and peer-nominations.

Being able to choose what goes into a CATZ lesson, how it would be delivered, and who to work with, especially the latter, were all rated as very important by most participants. These findings are again consistent with those of Boulton and Macaulay (2023) with anti-bullying CATZ. They suggest CATZ tutors should be offered such choices by adults who oversee the delivery of the intervention. Indeed, our own experiences implementing CATZ over several years give us the very strong impression that the more tutors have a sense of ‘ownership’ of their lesson, the more they tend to engage productively with CATZ activities and the more they seem to enjoy and benefit from them. Value placed on these different aspects of autonomy are unlikely to be fixed but rather may change as a function of experience with CATZ and the specific issue it is being used to address. Indeed, Boulton and Macaulay (2023) found that freedom to choose lesson content and mode of delivery both increased significantly after direct experience of (anti-bullying) CATZ. Research into changes (or stability) in what students ‘want’ when engaging in peer-led interventions is almost entirely absent from the literature, and so is clearly warranted in future studies.

Students’ views of the importance of being in control of CATZ features did indicate some age differences. Proportionally more (older) year 7 and 10 students than (younger) year 4 and 6 students wanted autonomy over choosing the content of a CATZ lesson and how to deliver it. This age trend is consistent with a more general desire for increasing autonomy that has been observed over the adolescent years (Noom et al., 2001). These findings suggest that offering such autonomy to older students may lead them to engage more willingly with CATZ activities that are offered to them, and in turn this may motivate them to ‘do a better job’ and so increase associated learning. Though plausible, these notions could be tested directly in future studies.

## Strengths and Limitations

Limitations of our study merit discussion. Important amongst these was the fact that CATZ and Control participants differed in the amount of time they had to think about and elaborate on the lesson material. It is entirely possible that allowing Control students the same time as our CATZ participants had to engage with the lesson material could enable them to learn as much as the latter evidently did but without the need for CATZ per se. We were aware of this methodological issue as we designed our study, but teachers refused to allow students acting as Controls this amount of time to work on such a discrete topic. They reasoned that students would become bored and learn little after the first one hour session (and our data showed they learned little during this session). Importantly, they said that the one-hour Control condition experience we did employ was a realistic way they or their colleagues would deliver this kind of teaching in school during normal Personal and Social Education lessons. On the other hand, teachers and school Principals *were* happy to allow their students what we recommended was sufficient time to act as CATZ tutors. The upshot of this is that while our findings show that CATZ had desirable effects that were not evident among students who experienced a shorter but realistic alternative way that teachers normally try to engender those benefits, we cannot claim that CATZ would be more effective than other interventions of comparable duration. To some extent, though, this may matter less to teachers than the evidence that CATZ *is* beneficial compared to what they normally do.

A related limitation of our work is that we did not set out to identify which elements of CATZ, such as preparing versus delivering the lesson or working co-operatively or cross-age interactions, contributed to its positive effects. Our view is that it is the *combination* of these elements that lies behind the success of CATZ, but we acknowledge that is speculation. Future studies could manipulate these elements, as well as ask students directly, to help identify those that are particularly important. Again, though, teachers may be more interested to know that CATZ as it was employed in this study was effective than they would be to know which aspects of it were responsible. Regarding the social validity, while participants reported positive perceptions of CATZ, it is possible social desirability bias may have also influenced how tutors responded to these items given how the researcher explained the CATZ intervention, and asked participants to rate different aspects of the intervention. However, measures were taken to reduce this influence. For instance, we reiterated to all participants that their responses would remain anonymous, and that we were more interested in the collective responses of participants, and not individual answers.

Our sample, while not small, was drawn from relatively few schools in the UK and so there is a need to see if our findings are replicable. Tests in different national/cultural contexts would be especially welcome, especially given that our sample was largely made up of white students and so did not reflect the cultural/ethnic diversity of the UK as a whole and other countries. Our social validity measures were single items rather than scales per se, although they did tap different facets of students' views of CATZ, and all had high test–retest reliability. The latter suggests students *were* giving their considered opinions. Moreover, other researchers have defended the use of single item measures of specific features of social validity (Boyle et al., 2011; Cheon & Reeve, 2015; Nickerson et al., 2014).

A further limitation was that our measures of Helpful Thoughts and Hostile Bias were in response to hypothetical scenarios and so the improvements we documented may not be evident in real life situations that are likely to arouse stronger emotions and impulsive reactions. Nevertheless, our findings offer a *prima facie* case for conducting a more challenging study that measures the effects of CATZ on thinking responses to real-life provocations, and on their actual engagement in aggression and conflict. Future studies that assess these variables using ratings from peers, teachers and parents alongside self-reports would be especially helpful. Moreover, a key rationale of 'thinking skills training' is that practice in analog settings such as ours is helpful in aiding the transfer to the real world (Goldstein et al., 1998). Such a view is also consistent with a tiered approach to helping all students develop good social-emotional skills, especially before behavioral problems arise (Mayworm & Sharkey, 2014).

## Summary

Overall, the results of the current study open some important areas of research that can probe further into the suitability and effectiveness of CATZ to help students learn important things about avoiding aggression and conflict. Here, we used it with all students in participating classes and other schools could do likewise as part of their tier 1 universal prevention efforts. It is noteworthy that participants as a whole reported positive perceptions of it in a general sense, and not a single student declined to take part or withdrew early. Another reason why CATZ may be suitable for wider take-up by schools is its fairly short duration. Moreover, it is our belief, based on talking with teachers who have witnessed ourselves facilitating CATZ, that it can be delivered by teachers themselves with minimal training, although a focused study of its acceptability among this group would provide better information.

Despite these limitations, and although many key questions remain, our findings suggest CATZ has the potential to be both an effective *and* acceptable way for student tutors to learn pro-social thinking skills that may lessen their risk of becoming involved in aggression and conflict. Tackling these behaviors remains a priority because they are common, impact negatively on young people and often seem resistant to adult-led interventions.

**Author Contributions** Michael J. Boulton: Conceptualization; data curation; formal analysis; investigation; methodology; writing – original draft; writing – review and editing. Peter J. R. Macaulay: Formal analysis; investigation; methodology; project administration; writing – original draft; writing – review and editing.

**Funding** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. No funds, grants, or other support was received.

**Data Availability** The data used in the research are available. The data can be obtained by emailing p.macaulay@derby.ac.uk. The materials used in the research are available. The materials can be obtained by emailing p.macaulay@derby.ac.uk.

## Declarations

**Conflict Interest** There are no competing interests to declare.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Baines, E., Blatchford, P., & Chowne, A. (2007). Improving the effectiveness of collaborative group work in primary schools: Effects on science attainment. *British Educational Research Journal*, 33(5), 663–680. <https://doi.org/10.1080/01411920701582231>
- Beck, R., & Fernandez, E. (1998). Cognitive-behavioral therapy in the treatment of anger: A meta-analysis. *Cognitive Therapy and Research*, 22(1), 63–74. <https://doi.org/10.1023/A:1018763902991>
- Bennett, D. S., & Gibbons, T. A. (2000). Efficacy of child cognitive-behavioral interventions for antisocial behavior: A meta-analysis. *Child & Family Behavior Therapy*, 22(1), 1–15. [https://doi.org/10.1300/J019v22n01\\_01](https://doi.org/10.1300/J019v22n01_01)
- Besnoy, K. D., & McDaniel, S. C. (2016). Going up in dreams and esteem: Cross-age mentoring to promote leadership skills in high school–age gifted students. *Gifted Child Today*, 39(1), 18–30. <https://doi.org/10.1177/1076217515613386>
- Biddle, B. J. (1986). Recent developments in role theory. *Annual Review of Sociology*, 12(1), 67–92. <https://doi.org/10.1146/annurev.so.12.080186.000435>

- Blatchford, P., Baines, E., Rubie-Davies, C., Bassett, P., & Chowne, A. (2006). The effect of a new approach to group work on pupil-pupil and teacher-pupil interactions. *Journal of Educational Psychology*, 98(4), 750. <https://doi.org/10.1037/0022-0663.98.4.750>
- Blatchford, P., Kutnick, P., Baines, E., & Galton, M. (2003). Toward a social pedagogy of classroom group work. *International Journal of Educational Research*, 39(1–2), 153–172. [https://doi.org/10.1016/S0883-0355\(03\)00078-8](https://doi.org/10.1016/S0883-0355(03)00078-8)
- Boulton, M. J. (1993). Proximate causes of aggressive fighting in middle school children. *British Journal of Educational Psychology*, 63(2), 231–244. <https://doi.org/10.1111/j.2044-8279.1993.tb01054.x>
- 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. <https://doi.org/10.1080/03069880500327546>
- 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. <https://doi.org/10.1016/j.beth.2013.12.004>
- Boulton, M. J., & Boulton, R. (2012). Resistant to the message: Are pupils unreceptive to teachers' anti-bullying initiatives and if so why? *Educational Studies*, 38(5), 485–489. <https://doi.org/10.1080/03055698.2011.643112>
- Boulton, M. J., Boulton, L., Camerone, E., Down, J., Hughes, J., Kirkbride, C., & Sanders, J. (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. <https://doi.org/10.1089/cyber.2016.0046>
- Boulton, M. J., Macaulay, P. J., Atherton, S., Boulton, L., Colebourne, T., Davies, M., & Turner, C. (2023). Promoting Junior School Students' Anti-bullying Beliefs with the CATZ Cross-age Teaching Zone Intervention. *International Journal of Bullying Prevention*, 5, 38–51. <https://doi.org/10.1007/s42380-021-00111-9>
- Boulton, M. J., & Macaulay, P. J. (2023). Testing the social validity of the CATZ cross-age teaching zone anti-bullying intervention among school students. *Research Papers in Education*, 1–21. <https://doi.org/10.1080/02671522.2023.2189289>
- Boulton, M. J., Trueman, M., & Murray, L. (2008). Associations between peer victimization, fear of future victimization and disrupted concentration on class work among junior school pupils. *British Journal of Educational Psychology*, 78(3), 473–489. <https://doi.org/10.1348/000709908X320471>
- Boyle, C., Lynch, L., Lyon, A., & Williams, C. (2011). The use and feasibility of a CBT intervention. *Child and Adolescent Mental Health*, 16(3), 129–135. <https://doi.org/10.1111/j.1475-3588.2010.00586.x>
- Briesch, A. M., & Chafouleas, S. M. (2009). Review and analysis of literature on self-management interventions to promote appropriate classroom behaviors (1988–2008). *School Psychology Quarterly*, 24(2), 106. <https://doi.org/10.1037/a0016159>
- Carter, S. L. (2010). *The social validity manual: A guide to subjective evaluation of behavior interventions*. Academic Press.
- Chafouleas, S. M., Briesch, A. M., Riley-Tillman, T. C., & McCoach, D. B. (2009). Moving beyond assessment of treatment acceptability: An examination of the factor structure of the Usage Rating Profile—Intervention (URP-I). *School Psychology Quarterly*, 24(1), 36. <https://doi.org/10.1037/a0015146>
- Cheon, S. H., & Reeve, J. (2015). A classroom-based intervention to help teachers decrease students' amotivation. *Contemporary Educational Psychology*, 40, 99–111. <https://doi.org/10.1016/j.cedpsych.2014.06.004>
- Cillessen, A. H., Lansu, T. A., & Van Den Berg, Y. H. (2014). Aggression, hostile attributions, status, and gender: A continued quest. *Development and Psychopathology*, 26(3), 635–644.
- 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. [https://doi.org/10.1002/\(SICI\)1098-2337\(2000\)26:1%3C85::AID-AB7%3E3.0.CO;2-5](https://doi.org/10.1002/(SICI)1098-2337(2000)26:1%3C85::AID-AB7%3E3.0.CO;2-5)
- Cowie, H., Naylor, P., Talamelli, L., Chauhan, P., & SMITH, P. K. (2002). Knowledge, use of and attitudes towards peer support: A 2-year follow-up to the Prince's Trust survey. *Journal of Adolescence*, 25(5), 453–467. <https://doi.org/10.1006/jado.2002.0498>
- Cowie, H., Smith, P. K., Boulton, M. J., & Lava, R. (1994). *Cooperation in the multi-ethnic classroom: The impact of cooperative group work on social relationships in middle schools*. Routledge.
- Cunningham, C. E., Mapp, C., Rimas, H., Cunningham, L., Mielko, S., Vaillancourt, T., & Marcus, M. (2016). What limits the effectiveness of antibullying programs? A thematic analysis of the perspective of students. *Psychology of Violence*, 6(4), 596.
- DePaulo, B. M., Tang, J., Webb, W., Hoover, C., Marsh, K., & Litowitz, C. (1989). Age differences in reactions to help in a peer tutoring context. *Child Development*, 423–439. <https://doi.org/10.2307/1130987>
- Daunic, A. P., Smith, S. W., Brank, E. M., & Penfield, R. D. (2006). Classroom-based cognitive-behavioral intervention to prevent aggression: Efficacy and social validity. *Journal of School Psychology*, 44(2), 123–139. <https://doi.org/10.1016/j.jsp.2006.01.005>
- Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, 20, 391–409. <https://doi.org/10.1007/s10648-008-9083-6>
- Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology*, 18(3), 791–814. <https://doi.org/10.1017/S0954579406060391>
- Eisenberg, N., Valiente, C., & Eggum, N. D. (2010). Self-regulation and school readiness. *Early Education and Development*, 21(5), 681–698. <https://doi.org/10.1080/10409289.2010.497451>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Goldstein, A. P. (1999). *Low-level aggression: First steps on the ladder to violence*. Research Press.
- Goldstein, A. P., Glick, B., & Gibbs, J. C. (1998). *Aggression Replacement Training*. Research Press.
- Holder, C., Whetstone, P., & Sheinker, J. (2008). When research meets practice: Using metacognitive strategies to teach social skills. *International Journal of Learning*, 15, 205–212.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Karcher, M. J. (2007). Cross-age peer mentoring. *Youth Mentoring: Research in Action*, 1(7), 3–17.
- Karcher, M. J. (2008). The cross-age mentoring program: A developmental intervention for promoting students' connectedness across grade levels. *Professional School Counseling*, 12(2), 10494–10503. <https://doi.org/10.1177/2156759X080120020>
- Lochman, J. E., & Wells, K. C. (2002). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the Coping Power program. *Development and Psychopathology*, 14(4), 945–967. <https://doi.org/10.1017/S09545794020004157>
- Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review*, 27(2), 173–187. <https://doi.org/10.1016/j.cpr.2006.08.001>
- Mayworm, A. M., & Sharkey, J. D. (2014). Ethical considerations in a three-tiered approach to school discipline policy and practice. *Psychology in the Schools*, 51, 693–704. <https://doi.org/10.1002/pits.21782>



- Monks, C. P., Smith, P. K., & Kucaba, K. (2021). Peer victimisation in early childhood; observations of participant roles and sex differences. *International Journal of Environmental Research and Public Health*, 18(2), 415. <https://doi.org/10.3390/ijerph18020415>
- Naylor, P., & Cowie, H. (1999). The effectiveness of peer support systems in challenging school bullying: The perspectives and experiences of teachers and pupils. *Journal of Adolescence*, 22(4), 467–479. <https://doi.org/10.1006/jado.1999.0241>
- Nickerson, A. B., Serwacki, M. I., Brock, S. E., Savage, T. A., Woitaszewski, S. A., & Reeves, M. A. L. (2014). Program evaluation of the prepare school crisis prevention and intervention training curriculum. *Psychology in the Schools*, 51, 466–479. <https://doi.org/10.1002/pits.21757>
- Noom, M. J., Deković, M., & Meeus, W. (2001). Conceptual analysis and measurement of adolescent autonomy. *Journal of Youth and Adolescence*, 30(5), 577–595. <https://doi.org/10.1023/A:1010400721676>
- Pellegrini, A., & Blatchford, P. (2013). *The child at school: Interactions with peers and teachers*. Routledge.
- Pereira, A. I., Marques, T., Russo, V., Barros, L., & Barrett, P. (2014). Effectiveness of the friends for life program in portuguese schools: Study with a sample of highly anxious children. *Psychology in the Schools*, 51(6), 647–657. <https://doi.org/10.1002/pits.21767>
- Rigby, K., & Bagshaw, D. (2003). Prospects of adolescent students collaborating with teachers in addressing issues of bullying and conflict in schools. *Educational Psychology*, 23(5), 535–546. <https://doi.org/10.1080/0144341032000123787>
- Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review*, 17(4), 327–362.
- Shogren, K. A., Faggella-Luby, M. N., Bae, S. J., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior: A meta-analysis. *Journal of Positive Behavior Interventions*, 6(4), 228–237. <https://doi.org/10.1177/10983007040060040401>
- Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, 94(3), 429. <https://doi.org/10.1037/0033-2909.94.3.429>
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43–69.
- Slavin, R. E. (2010). Co-operative learning: What makes group-work work? In H. Dumont, D. Istance, & F. Benavides (Eds.), *The nature of learning: Using research to inspire practice* (pp. 161–178). OECD.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences*. Lawrence Erlbaum.
- Stukas, A. A., Jr., Clary, E. G., & Snyder, M. (1999). Service learning: Who benefits and why. *Social Policy Report*, 13(4), 1–23.
- Sukhodolsky, D. G., Kassinove, H., & Gorman, B. S. (2004). Cognitive-behavioral therapy for anger in children and adolescents: A meta-analysis. *Aggression and Violent Behavior*, 9(3), 247–269. <https://doi.org/10.1016/j.avb.2003.08.005>
- Sullivan, A. M., Johnson, B., Owens, L., & Conway, R. (2014). Punish them or engage them? Teachers' views of unproductive student behaviours in the classroom. *Australian Journal of Teacher Education*, 39(6), 43–56.
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics*. Harlow: Pearson Education.
- Thurston, A., Van de Keere, K., Topping, K. J., Kosack, W., Gatt, S., Marchal, J., et al. (2007). Peer learning in primary school science: Theoretical perspectives and implications for classroom practice. *Electronic Journal of Research in Educational Psychology*, 5, 477–496.
- Topping, K., & Ehly, S. (1998). *Peer-Assisted Learning*. Lawrence Erlbaum.
- Topping, K. J., & Bryce, A. (2004). Cross-age peer tutoring of reading and thinking: Influence on thinking skills. *Educational Psychology*, 24(5), 595–621. <https://doi.org/10.1080/0144341042000262935>
- Topping, K. J., Miller, D., Murray, P., Henderson, S., Fortuna, C., & Conlin, N. (2011). Outcomes in a randomised controlled trial of mathematics tutoring. *Educational Research*, 53(1), 51–63. <https://doi.org/10.1080/00131881.2011.552239>
- Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology*, 7(1), 27–56. <https://doi.org/10.1007/s11292-010-9109-1>
- Veldman, M. A., Doolaard, S., Bosker, R. J., & Snijders, T. A. B. (2020). Young children working together. Cooperative learning effects on group work of children in Grade 1 of primary education. *Learning and Instruction*, 67, 101308.
- Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. Harvard University Press.
- Wang, Z., Chen, X., Liu, J., Bullock, A., Li, D., Chen, X., & French, D. (2020). Moderating role of conflict resolution strategies in the links between peer victimization and psychological adjustment among youth. *Journal of Adolescence*, 79, 184–192. <https://doi.org/10.1016/j.adolescence.2020.01.002>
- Watts, G. W., Bryant, D. P., & Carroll, M. L. (2019). Students with emotional-behavioral disorders as cross-age tutors: A synthesis of the literature. *Behavioral Disorders*, 44(3), 131–147.
- Whetstone, P. J., Gillmor, S. C., & Schuster, J. G. (2015). Effects of a metacognitive social skill intervention in a rural setting with at-risk adolescents. *Rural Special Education Quarterly*, 34(2), 25–35. <https://doi.org/10.1177/875687051503400205>
- Willis, P., Bland, R., Manka, L., & Craft, C. (2012). The ABC of peer mentoring—what secondary students have to say about cross-age peer mentoring in a regional Australian school. *Educational Research and Evaluation*, 18(2), 173–185. <https://doi.org/10.1080/13803611.2011.650920>
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 71(1), 136.
- Witt, J. C. (1986). Teachers' resistance to the use of school-based interventions. *Journal of School Psychology*, 24(1), 37–44. [https://doi.org/10.1016/0022-4405\(86\)90040-3](https://doi.org/10.1016/0022-4405(86)90040-3)
- Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2007). The scientific base linking social and emotional learning to school success. *Journal of Educational and Psychological Consultation*, 17(2–3), 191–210. <https://doi.org/10.1080/10474410701413145>