

This is an Accepted Manuscript for **Psychological Medicine**. Subject to change during the editing and production process.

DOI: 10.1017/S0033291724002666

The effects and costs of an anti-bullying programme (KiVa) in UK primary schools: a multicentre cluster randomised controlled trial

Lucy Bowes^{1,2*}

Malavika Babu⁴

Julia R. Badger^{1,2,7}

Matthew R. Broome^{2,5}

Rebecca Cannings-John⁴

Suzy Clarkson³

Elinor Coulman⁴

Rhiannon Tudor Edwards³

Tamsin Ford⁶

Richard P. Hastings⁷

Rachel Hayes⁸

Fiona Lugg-Widger⁴

Eleri Owen-Jones⁴

Paul Patterson⁹

Jeremy Segrott⁴

Mia Sydenham⁴

Julia Townson⁴

Richard C. Watkins^{3,10}

Holly Whiteley³

Margiad E. Williams³

the Stand Together Team^{**}

Judy Hutchings³

****The Stand Together Team:**

Sadja Butt⁵

Chris Chiswell⁵

Mackenzie Fong⁴

Hayley Gains⁹

Helin Gosalia⁴

Anwen Jones³

Bryony Longdon⁹

Siobhan B. Mitchell⁹

Caitlin Murray⁹

Naomi Rose¹

Melanie Varley⁴

*Corresponding author: lucy.bowes@psy.ox.ac.uk

1. Department of Experimental Psychology, Oxford University, OX2 6GG
2. National Institute for Health and Care Research (NIHR) Oxford Health Biomedical Research Centre, Oxford University, OX3 7JX
3. Bangor University, Bangor, Gwynedd LL57 2DG
4. Centre for Trials Research, Cardiff University, Neuadd Meirionnydd, Heath Park, Cardiff, CF14 4YS, UK
5. Institute for Mental Health, University of Birmingham, Edgbaston, Birmingham B15 2TT and Birmingham Women's and Children's NHS Foundation Trust, Steelhouse Lane, Birmingham, B4 6NH
6. Department of Psychiatry, University of Cambridge, Hershel Smith Building, Cambridge Biomedical Campus, Robinson Way Cambridge CB2 0SZ
7. School of Educational Learning and Communication Sciences, University of Warwick, CV4 8UW
8. Faculty of Health and Life Sciences, Department for Public Health and Sport Sciences, South Cloisters, St Luke's Campus, University of Exeter, Exeter EX1 2LU
9. Birmingham Women's and Children's NHS Foundation Trust
10. GwE (North Wales Regional School Improvement Service), Bryn Eirias, Ffordd Abergele, Bae Colwyn, Conwy LL29 8BY

Word count: 5258

Background

Childhood bullying is a public health priority. We evaluated the effectiveness and costs of KiVa, a whole-school anti-bullying programme that targets the peer context.

Methods

A two-arm pragmatic multicentre cluster randomised controlled trial with embedded economic evaluation. Schools were randomised to KiVa-intervention or usual practice (UP), stratified on school size and Free School Meals eligibility. KiVa was delivered by trained teachers across one school year. Follow-up was at 12 months post randomisation. Primary outcome: student-reported bullying-victimization; secondary outcomes: self-reported bullying-perpetration, participant roles in bullying, empathy and teacher-reported Strengths and Difficulties Questionnaire. Outcomes were analysed using multilevel linear and logistic regression models.

Findings

Between 8/11/2019 – 12/02/2021, 118 primary schools were recruited in four trial sites, 11,111 students in primary analysis (KiVa-intervention: n=5,944; 49.6% female; UP: n=5,167, 49.0% female). At baseline, 21.6% of students reported being bullied in the UP group and 20.3% in the KiVa-intervention group, reducing to 20.7% in the UP group and 17.7% in the KiVa-intervention group at follow-up (odds ratio 0.87; 95% confidence interval 0.78 to 0.97, p-value=0.009). Students in the KiVa group had significantly higher empathy and reduced peer problems. We found no differences in bullying perpetration, school wellbeing, emotional or behavioural problems. A priori subgroup analyses revealed no differences in effectiveness by socioeconomic gradient, or by gender. KiVa costs £20.78 more per pupil than usual practice in the first year, and £1.65 more per pupil in subsequent years.

Interpretation

The KiVa anti-bullying programme is effective at reducing bullying victimisation with small-moderate effects of public health importance.

Funding

The study was funded by the UK National Institute for Health and Care Research (NIHR) Public Health Research programme (17-92-11). Intervention costs were funded by the Rayne

Introduction

Bullying in childhood may be one of the most tractable risk factors for mental health problems in childhood and adolescence (Clarkson et al., 2022; Ford, 2014). Generally defined as a pattern of 'unwanted, aggressive behaviour ... that involves a real or perceived imbalance of power' (Olweus, 1994), the psychiatric morbidity that may arise from childhood bullying is substantial. Population studies suggest that 25–40% of mental health problems including depression, anxiety and self-harm in young adults may be attributable to childhood bullying (Bowes et al., 2015; Fisher et al., 2012). Bullied children access more school health, primary care and specialist child mental health services than their counterparts (Brimblecombe et al., 2018) and experience poorer mental and physical health into adulthood (Takizawa et al., 2014). Bullying is also associated with school absenteeism (Brown, 2011), which may impact on future educational attainment and employment prospects. Furthermore, children who bully are also at risk of harm and more likely to show later violent behaviour and illicit drug use (Ttofi et al., 2016). Universal whole-school interventions, promoting school-wide change are the most effective at reducing bullying (Gaffney, 2019) and likely to provide a non-stigmatising approach to prevention. One of the most widely used bullying prevention programmes across Europe is the Finnish *Kiusaamisen Vastaa* or "KiVa" programme. The programme was informed by a social architecture model of bullying, which highlights the significant roles of bystanders in supporting or standing against bullying (Salmivalli, 1996). In a large-scale RCT conducted in Finland between 2007–2009 across 234 schools (8,166 students), KiVa significantly reduced bullying and victimisation among 7 to 11-year-old students (Karna, Voeten, Little, Poskiparta, Kaljonen, et al., 2011) and reducing anxiety and depression (Williford et al., 2012). Since 2009, KiVa has been implemented in over 90% of Finnish public schools (approximately 2700 schools) and demonstrates year-on-year positive effects (Karna, Voeten, Little, Poskiparta, Alanen, et al., 2011). Whilst KiVa has been found to be successful in reducing victimization in trials in the Netherlands (Huitsing et al., 2020) and Italy (Nocentini & Menesini, 2016), other trials, including in Wales, found no effect (Axford et al., 2020). There are important differences

between the education systems in the UK and Finland that may influence whether KiVa intervention effects will transfer. The route to becoming a teacher in the UK is more varied and children are tested more to assess their academic capabilities. Schools are typically more diverse and have greater social inequalities (Hadjar, 2016). Whilst every school in the UK is required by law to have an anti-bullying policy, more guidance on effective, evidence-based anti-bullying programmes is needed. In this two-arm pragmatic multicentre cluster randomised controlled trial (cRCT) with embedded economic evaluation, our primary aim was to assess the effectiveness and costs of the KiVa anti-bullying programme in reducing student-reported bullying victimisation. Additionally, a range of secondary outcomes were explored including student-reported bullying perpetration, participant roles in bullying and school related well-being; teacher-reported student emotional and behavioural difficulties; and teachers' self-efficacy in dealing with bullying, mental well-being, and burnout.

Methods:

A full outline of the research methodology can be found in our [protocol paper](#) (Clarkson et al., 2022).

Study Design

We conducted a two-arm pragmatic, multicentre, cluster RCT with an embedded economic evaluation. Schools and students were recruited across four areas (North Wales, Birmingham, Southeast England, and Southwest England).

Ethical approval was granted by Bangor University Psychology Research Ethics and Governance Committee (2019-16592) on 13th November 2019. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Schools and participants

We included mainstream UK state-maintained primary schools with at least two KS (Key Stage) 2 classes for children aged 7-11 years. We excluded schools that: primarily delivered education through a language other than English or Welsh; were already implementing a recognised anti-bullying intervention programme or had previously implemented KiVa;

catered solely for children with Special Educational Needs; and schools without a leadership that could guarantee project participation for the year of data collection/ implementation. All students at participating schools in school years 3-5 (KS 2) were eligible for the trial. Head teachers provided opt-in consent for school-level data collection and the implementation of KiVa; parents/carers provided opt-out consent for data collection relating to their child; and teaching staff provided opt-in consent for data about themselves. Parents, carers or teachers could not withhold consent for KiVa to be delivered as part of the school's statutory personal, social education/personal, social, health and economic education (PSE/PSHE) curriculum. Children provided their written assent via e-tablets for self-reported outcomes at each data collection point. Prior to the 12-month follow-up, all parents received a letter detailing the plans to link their child's trial data to data held by the Department for Education, National Pupil Database (NPD) and in Wales, the Welsh Secure, Anonymised Information Linkage (SAIL) database. Parents had the opportunity to opt-out of this activity without it impacting on their child's continued participation in the trial.

Randomisation and masking

Schools were randomised to receive the KiVa intervention or continue with UP by the Centre for Trials Research at Cardiff University using random permuted blocks in a 1:1 ratio. Randomisation was stratified by the following: study area (North Wales, Birmingham, Southeast and Southwest England); number of KS2 students (UK school years 3-6, ages 7-11) (below or above site median); and percentage of students eligible for Free School Meals (FSM) (below or above site median). Randomisation of schools occurred after all schools had been recruited. Due to COVID-19 pandemic related school closures between January and March 2021 delaying the timing of child-reported baseline data collection and the need for intervention schools to be trained in KiVa before the end of academic year, 48% of schools were informed of their allocation before child-reported baseline data were collected. Most (94%) of teacher reported TSDQ data were collected before schools were informed of their allocation. Researchers collecting baseline school data were blind to the randomisation process, as were students (who were the source of the primary outcome data). It was not possible to blind researchers and field workers at follow-up due to the requirement that KiVa schools make the intervention public (e.g. through posters and staff tabards). The statistician analysing the primary and secondary outcome data and the health economists undertaking the economic analysis remained blinded.

Patient and public involvement

Two advisory groups were established in North Wales, one with teachers and head teachers, and one with six children from three different primary schools. Groups met twice at the start of the project and reviewed and advised on data collection methods, participant information and consent forms. The groups reconvened after the COVID-19 pandemic in 2023 to advise on the output and dissemination activities.

Procedures

Baseline survey data were collected from students in school years 3-5 between 20th April – 30th June 2021. Child questionnaires were administered via e-tablets with researchers reading the questions aloud to the whole class. All questions included the option “I’d rather not say”. Teachers completed questionnaires either online or with paper and pen. The KiVa anti-bullying intervention was implemented in the academic year which ran from September 2021 to July 2022, and survey outcomes were measured from 26th April – 15th July 2022.

Intervention

KiVa is based on research demonstrating that ‘bystanders’ – children who are present during bullying but not actively involved – can contribute to the maintenance of bullying by assisting or reinforcing the perpetrator’s behaviour, giving the bully a position of power (Salmivalli, 1996). Defending the victim, on the contrary, can help to make bullying an unsuccessful strategy for attaining high social status. By influencing the behaviour and norms of all students, the social rewards gained by perpetrators are reduced and, consequently, reduces their motivation to bully. The KiVa programme has two distinct components; universal actions that target the whole school and indicated actions for tackling specific cases of bullying. Universal actions at the class and school level help students to recognise bullying, providing them with safe ways of responding to incidents, and teaching them to empathise with, and support, victims. This includes structured lessons, delivered by teachers. The manualised curriculum targets students in school Years 3 and 4 (aged 7-9) (Unit 1), and Years 5 and 6 (aged 9-11) (Unit 2). Although the lesson plans target KS2 classes, the programme is introduced to all staff, parents and students and is visible across the school. Each unit contains ten structured 90-minute lessons, typically delivered fortnightly as twenty 45-minute lessons throughout the school year by class teachers. The curriculum encourages student engagement via oral

presentations, role-play, videos, group work and whole class activities. Online games, that support lessons can be played at home or in school. Other universal actions include posters for school corridors, high-visibility vests for break-time supervisors to highlight the presence of supervision, and a parent's guide. A trained KiVa team (2 to 3 staff members) address confirmed bullying incidents with children involved using the structured and scripted indicated actions, with individual follow-up meetings 1-2 weeks later to ensure that the situation has improved. Two members of the teaching/management team from each intervention school attended a local two-day training course delivered in June 2021 by accredited KiVa trainers at each site and then led school-wide implementation. Due to ongoing disruption from the COVID-19 pandemic, some training was delivered in an online format.

The intervention delivery was over one academic year. However, the programme is designed to be embedded into ongoing school practice.

Comparison

The mandatory PSE/PSHE curricula in Wales and England aim to develop students' pro-social values and attitudes and empower participation in school and community life as responsible citizens. Comparison schools continued to use existing methods to cover this curriculum, and to address bullying. UP schools continued with their standard practice and were free to implement any other programmes or strategies they wished to in their school. At baseline, all schools reported which existing policies and practices they used to prevent and deal with bullying. At the follow-up point all schools were asked if there had been any changes to their policies or practices during the trial period.

Outcomes

Primary outcome

The primary study outcome was self-reported bullying victimisation, measured by student responses on the Olweus Bully/Victim Questionnaire (OBVQ) (Olweus, 1996). The OBVQ measures different forms of bullying, including verbal, physical, relational and cyber-bullying. The global item: "How often have you been bullied at school in the last couple of months?" was used to measure victimisation. Students respond on a five-point scale (0 = "not at all", 1 = "once or twice", 2 = "2 or 3 times a month", 3 = "about once a week", 4 = "several times a week"). The response was dichotomised with those scoring 2 to 4 classified as victimised and those scoring 0 to 1 as not victimised (20)

Secondary outcomes

i) Student-reported outcomes:

1. The OBVQ global item: “How often have you bullied others at school in the last few months?” measures bullying perpetration. Students' responses, using the same five-point scale reported above were analysed continuously and dichotomised, with scores of 2 to 4 classified as perpetration (Solberg, 2003).

2. Subjective student wellbeing in school was measured using the “How I feel about my school” (HIFAMS) (Allen, 2018). The survey is comprised of seven items which asks students how they feel about various aspects of school life, for example how they feel when completing their schoolwork or when they are in the playground.

3. Student empathy was measured using the Empathy Toward Victim Scale (Pöyhönen V., 2010). The seven-item measure asked respondents to rate their level of empathy towards victims on a Likert scale anchored from ‘never’ (0) to ‘always’ (3).

4. Roles in bullying situations was measured using the “Participant Role Questionnaire” (PRQ) (Salmivalli, 2004), designed to identify the different roles that peers play in bullying situations. The questionnaire identifies five different bullying roles; bully, assistant, reinforcer, defender and outsider. We were particularly interested in the role of defender, given that KiVa is intended to raise awareness among children about the importance of everyone taking a stand against bullying. The respondent rates how often they behave in the ways described for each role on a three-point scale (Never, Sometimes, Often).

ii) Teacher-reported student outcomes (measured through teacher survey reports for each student):

5. Teacher-reported Strengths and Difficulties Questionnaire (TSDQ) (R., 2001) Click here to enter text. was reported at baseline and follow-up. The TSDQ measure provides a Total Difficulties score from 20 behavioural and emotional problems items as well as a pro-social behaviour score from five items. Subscales include a hyperactivity score, emotional score, conduct problems score and peer problems score.

We intended to collect data from the Absence and Exclusion, and KS2 attainment datasets for English pupils from the National Student Database (NPD) and from the Welsh Government SAIL database for students in both trial arms. However, due to COVID 19 pandemic-related delays in obtaining data, these data are not currently available. For a full list of tertiary measures (not included in this paper), see study protocol (Clarkson et al., 2022).

iii) Staff-level outcomes.

6. Adapted version of the 5-item Challenging Behaviour Self-Efficacy Scale, designed as a measure of teacher self-efficacy related to challenging behaviours and adapted to specifically refer to bullying behaviours (Hastings & Brown, 2002).

7. Warwick-Edinburgh Mental Wellbeing Scale (Tennant et al., 2007). This 14-item positively worded scale measures adult mental wellbeing.

8. Maslach Burnout Inventory-Educator

Survey (MBI-ES), a 22-item psychological inventory measuring emotional exhaustion, depersonalisation, and personal accomplishment (Maslach C, 1996).

Process evaluation

The methods for the process evaluation are reported in our [protocol paper](#) (Clarkson et al., 2022) and the results will be in a separate publication.

Economic evaluation

In line with guidance for evaluating complex public health interventions (Skivington et al., 2021), a broad economic evaluation using cost-consequence analysis (CCA), cost-effectiveness analysis (CEA), and cost-utility analysis (CUA) frameworks explored the cost-effectiveness of the KiVa intervention compared to UP. An education-sector perspective (Edwards RT, 2019) was taken to evaluate incremental effects at 12 months after randomisation. These results are reported separately.

Cost analyses from an education-sector perspective focused on two out of three components of the KiVa programme; i) class-level delivery of KiVa lessons and, ii) whole-school

activities to promote KiVa to the school population. KiVa cost components were categorised into recurrent and non-recurrent costs and unit costs were derived using appropriate UK estimates for the trial year (2021/2022) (Appendix X, Table 1). Resource use and cost data for KiVa were collated from the Children's Early Intervention Trust (CEIT) charity that holds the license for KiVa dissemination in the UK and process evaluation school checklists. Resource use data for UP were collected via PSE/PSHE school checklists. Cost calculations for usual practice considered the school-reported overlap between the KiVa and PSE/PSHE curriculums. Where there were missing data at the school-level mean imputation was used unless there was evidence that these costs had not been incurred. Discounting was not applied.

Safety and adverse events

There were no anticipated risks to participants or schools.

Statistical analyses

Our sample calculation was based on previous research (reduction from 18% to 14% in rates of victimisation) (Skivington et al., 2021), and a similar baseline victimisation rate of 18% from a small UK based pre-post study (Edwards RT, 2019). Assuming 111 students in Years 3 to 5, an intracluster correlation coefficient (ICC) of 0.02, and allowing for one school dropout per arm, 10% student dropout due to either opt-out or loss to follow-up, an 18% rate of victimisation, and a relative reduction of 22%, a trial involving 116 schools (58 per arm) would provide 90% power at a 5% significance level (a total of 12,828 students).

All analyses were intention to treat without imputation (a complete case analysis restricting to pupils with responses at both baseline and follow-up), with outcomes compared between KiVa and UP groups using three-level regression models (allowing for clustering between students within school, and between schools within sites). Analyses controlled for school level stratification variables (school size, proportion of students eligible for FSM), key student characteristics (age, sex) as well as baseline outcome measures (where collected).

For binary outcomes a logistic model was used, and the result presented as adjusted odds ratios (ORs) comparing the odds of an event in KiVa schools compared with UP schools. For continuous outcomes, we fitted a linear-regression model and presented results as difference

in adjusted means (KiVa minus UP). Multilevel ordinal logistic regression model was used to compare TSDQ scales. Due to skewness in the TSDQ scales, data were categorised according to the clinical cut-offs (normal, borderline, abnormal). Box-cox transformations were applied to skewed data when necessary and Glass's delta standardised effect size calculated as the difference in means (KiVa - UP) divided by the standard deviation of the UP group. Between-group comparisons are presented with two-sided 95% confidence intervals (CI). Pre-specified exploratory sub-group analyses investigated the effect of the intervention by student gender and school-level percentage of students eligible for FSM status. Three additional post-hoc subgroups were performed: Pupils self-reported age (6-7, 8, 9, 10/11 years of age), whether the school previously recruited before COVID-19 pandemic for the Stand Together trial, and the timing of baseline data collection. For all subgroup analyses, an interaction term was added to the main comparison model between the study arm and each variable and results presented using 95% CIs.

The impact of missing outcome data on the trial conclusions, missing mechanisms were explored, and appropriate imputation methods applied via sensitivity analyses.

The effect of KiVa lesson dosage on the primary outcome was explored in a sensitivity analysis and estimated in a way that preserves randomisation using complier-average causal effect estimates, using a two-stage least squares instrumental variable regression model (Dunn G, 2005).

The secondary outcome of the OBVQ bullying victimisation and perpetrator score was originally omitted from the SAP but is included as a post-hoc analysis (ordinal logistic regression).

In exploratory analyses, mediation analyses were used to determine whether the effectiveness of the intervention on bullying victimisation occurred through change in affective empathy and secondly with change self-efficacy in defending. The mediation analysis was parameterised as a 2-1-1 model where the independent variable (intervention) is at the school level, while the mediator (affective empathy or self-efficacy) and the outcome (bullying victimization) are at the pupil level. Both models adjusted for pupil-level characteristics (age, gender), school level characteristics (FSM, KS2) along with corresponding baseline measures.

Analyses were done by two statisticians (MB and RCJ) using IBM SPSS Statistics version 20.0 (IBM Corporation, Armonk, NY, USA), and Stata® version 17 (StataCorp LP, College Station, TX, USA).

An independent steering committee oversaw the study. The trial is registered at the ISRCTN 12300853.

Results

See Figure 1 for the CONSORT Flow Diagram. Between 8th November 2019 and 12th February 2021, 118 schools were recruited from 396 schools approached across four trial sites. From a potential of 11,922 students, 11,111 (93.2%) students completed a baseline questionnaire. Twelve month-follow up data was collected from 9,981 (90.1%) students, 5,321 (90.0%) students in the KiVa schools compared to 4,660 (90.2%) in UP schools. In addition, a total of 263 (KiVa) and 202 (UP) questionnaires were completed by students who had not completed a baseline questionnaire. Baseline student and teacher characteristics were well balanced across trial arms (Table 1). Forty-nine percent of students were female and on average aged 8.7 years.

The total number of students available for the complete case analysis (baseline and 12-month follow-up) was 9,981 (Table 2). At baseline, 20.3% of students in KiVa schools and 21.6% of students in UP schools experienced bullying victimisation. After 12-month follow-up, the bullying victimisation rate was reduced to 17.7% in the KiVa arm compared to 20.7% in the UP arm. The adjusted odds ratio indicated that the students from schools randomised to receive KiVa had a 13% lower odds of reporting bullying victimisation compared to students in schools randomised to continue with UP (Table 2) (odds ratio (OR): 0.87, 95% CI: 0.78, 0.97, p-value=0.009). Pre-specified sensitivity analyses were performed to investigate the effect of dosage (at least seven out of the 10 annual lessons delivered over a full school year to each class) on intervention effect, and imputation for missing data. The intervention effect was consistent across all sensitivity analyses (appendix, Table S1.1). No differential effects were found for any primary outcome in subgroup analyses (appendix, Table S2.1). The additional post-hoc analysis on the maximum OBVQ bullying victimisation and perpetrator score, supported the primary analysis (Table 3).

Some secondary outcomes suggested positive impacts of the KiVa intervention (table 2). Students in schools randomised to KiVa had increased empathy towards victims compared to students in the UP arm (adjusted mean difference: 0.45, 95% CI: 0.20, 0.69), however they reported significantly less self-efficacy in defending other children from being bullied compared to students in the UP arm (OR: 0.76, 95% CI: 0.59, 0.97). We found no evidence of a difference in school-related happiness (HIFAMS) or self-reported bullying perpetration. Students in the KiVa arm had 15% lower odds of having teacher-reported TSDQ peer problems as compared to students in the UP arm (OR: 0.85, 95% CI: 0.75, 0.96). Follow-up exploratory analysis suggested that this effect was driven by a reduction in teacher-reported peer victimization (see Table S3.2). We found no evidence of trial arm differences in the other TSDQ subdomains (including emotional problems, conduct problems, hyperactivity and prosocial behaviour). We found no evidence of intervention effects on teacher self-efficacy in dealing with bullying, wellbeing, or burnout (Table S3.1).

Table S4.1 presents the direct, indirect and total effect of intervention and corresponding mediator on bullying victimisation. These models indicated that as compared to usual practice, 22% effect of KiVa is mediated through change in affective empathy whereas 12.5% of this effect mediated through change in self-efficacy in defending.

Economic evaluation

Full economic evaluation results will be reported in a separate paper. The main cost components of the KiVa intervention were training costs and staff-time spent delivering the KiVa curriculum (Appendix 5, Table S5.1). Mean total education sector-related costs of the KiVa intervention were £38.18 (SD £11.36) per pupil in the first year of KiVa set-up and implementation (i.e. non-recurrent and recurrent costs). Means costs were £19.05 (SD £3.94) per pupil for recurrent costs only. An average 40% overlap between the KiVa and PSE/PSHE curriculums was reported by intervention schools. When considering the mean cost of delivering 40% of the PSE/PSHE curriculum (£17.40 (SD £6.74) per pupil), the cost of KiVa compared to usual practice is estimated to be £20.78 more per pupil in the first year of KiVa set-up and implementation and £1.65 more per pupil when only recurrent costs are considered. Further details of resource use and costs are reported in the appendix

Discussion

The results of the study showed that the KiVa intervention reduced bullying victimisation among students from KiVa schools reporting a 13% reduction in the odds of victimisation compared with students from schools continuing with usual practice. The reduction in odds of bullying victimisation observed in our trial (13%) was lower than reported in other studies (Finland – 30% (Karna, Voeten, Little, Poskiparta, Kaljonen, et al., 2011); Netherlands 43.4% (Huitsing et al., 2020) , Italy 50.9% (Nocentini & Menesini, 2016). This may reflect the significant challenges in implementing a whole-school intervention in the context of the COVID-19 global pandemic, positive effects of UP, or both. It is unclear whether a 13% reduction in bullying could be considered meaningful at the school level. The concept of identifying ‘clinically meaningful’ reductions in symptoms is common in clinical psychology but is yet to be addressed in anti-bullying interventions. (Carey et al., 2023).

In addition to reducing bullying victimisation, students in the intervention group demonstrated increased empathy. Indeed, 22% effect of the KiVa was found to be mediated through change in affective empathy. These positive findings further support the efficacy of the KiVa programme in improving the overall social dynamics within schools. Students in the intervention group had fewer reported peer relationship problems, however this effect was no longer significant when adjusting for multiple testing. We also observed some small, unexpected effects on student reported self-efficacy in defending victims of bullying, which decreased in the intervention group. These findings should be interpreted with caution and require replication in independent samples. It is also important to note that no significant differences were observed in teacher-reported emotional or conduct problems, or child-reported well-being. Whilst the original evaluation of KiVa in Finland observed a significant reduction in symptoms of anxiety (but not depression), both the Dutch and Italian KiVa evaluations found small, largely non-significant effects on internalizing symptoms (Carey et al., 2023; Huitsing et al., 2020; Nocentini & Menesini, 2016). These results suggest that while the KiVa programme may reduce bullying, effects may not extend to mental health outcomes.

Our exploratory findings also indicate that the effectiveness of the KiVa programme may be consistent across different demographic groups, including those from more socioeconomically deprived settings and by gender. This suggests that the intervention may have the potential to benefit students from diverse backgrounds and is not limited to specific

subgroups – an important finding when attempting to ensure equitable access to effective anti-bullying interventions. Furthermore, the economic evaluation suggests that the KiVa programme is a low-cost intervention compared to many other UK school-based interventions. This is a significant finding and suggest that implementing whole-school approaches like KiVa is a feasible and efficient strategy for reducing bullying in schools.

Our study had several strengths. This is the largest randomised controlled trial in the world to date of the KiVa anti-bullying programme and included a broad and diverse range of schools and students. We included an embedded economic evaluation that collected data on the wider costs of bullying to schools and families. School and student participation were high. All analyses were pre-registered and were performed blind by statisticians at the Centre for Trials Research, a UKCRC Registered Trials Unit.

Limitations of our trial include our reliance on self-reported bullying victimisation and perpetration, which may inflate response bias. Whilst peer-nomination is often considered the gold-standard method of measuring bullying, this is not acceptable to many schools in the UK, nor does it typically allow for measurement of bullying outside of the classroom context. It also does not capture student perceived threat. Given that teachers also reported fewer peer relationship difficulties in the intervention arm, we can be confident that our findings support our student-reported outcomes. Secondly, some schools were informed of their assignment to the intervention before student reported baseline data were collected due to the impact of COVID-19 pandemic related school closures on our data collection timetable and the need for schools to train staff and plan for the next academic year. We mitigated the impact of this by asking teachers not to inform students or research staff of their assignment, and by performing sensitivity analyses to check that this did not impact on our findings (see Supplementary Table S2.1). Additionally, some schools were recruited to our trial in 2019, prior to the COVID-19 pandemic. For these schools, baseline data collection needed to be repeated. We mitigated any impact of this repeated data collection by performing additional sensitivity analyses to ensure that intervention effects were similar for these schools (see Supplementary Table S2.1). Finally, we collected follow-up data before the end of the school year whilst pupils remained in the same classes when schools had delivered the KiVa intervention for less than one academic year. Complex, whole-school interventions typically take time to bed in, which may have contributed to the lower than predicted intervention effect sizes observed. In support of this, we found that intervention schools that had delivered more of the KiVa lessons showed a greater reduction in bullying. It is plausible that the

effects of the intervention may increase as the programme gets embedded within the broader school system – as found in an RCT in the Netherlands which found greater effect sizes after two years of programme implementation (Huitsing et al., 2020), as well as the broader rollout of the KiVa programme in Finland (Karna, Voeten, Little, Poskiparta, Alanen, et al., 2011).

Our economic evaluation compared costs of KiVa training and delivery including whole-school activities against the mean cost of delivering 40% of the PSE/PSHE curriculum in UP schools. However, it is a mandatory requirement for all schools in the UK to have anti-bullying policy and practices in place, and we were unable to cost any additional school time focused specifically on bullying in UP schools, so these costs are likely to be an underestimate of the true costs of tackling bullying in UK schools. Thus, it is probable that the relative additional costs of delivering KiVa are even smaller than reported here.

Our study was conducted during the COVID-19 pandemic, when schools were experiencing increased pressure, closures, staff and student absences and other COVID-19 restrictions. This impacted all elements of implementation, from teacher training through to lesson delivery and data collection, primarily through staff and student absences. Given this unusual context, it is notable that only one school dropped out of the intervention (but remained in the data collection), and that our data return rates were high.

Overall, the results support the effectiveness and low cost of the KiVa anti-bullying programme as a public health intervention. The findings add to evidence for the importance of bystander effects and underscore the importance of implementing evidence-based whole-school approaches to address bullying. The results provide a valuable foundation for further research on the longer-term integration and embedding of KiVa into school practices.

References

- Allen, K., Marlow, R., Edwards, V., Parker, C., Rodgers, L., Ukoumunne, O. C., ... & Ford, T. (2018). 'How I feel about my school': the construction and validation of a measure of wellbeing at school for primary school children. *Clin Child Psychol & Psych*, 23(1), 25-41.
- Axford, N., Bjornstad, G., Clarkson, S., Ukoumunne, O. C., Wrigley, Z., Matthews, J., Berry, V., & Hutchings, J. (2020, Jul). The Effectiveness of the KiVa Bullying Prevention Program in Wales, UK: Results from a Pragmatic Cluster Randomized Controlled Trial. *Prev Sci*, 21(5), 615-626. <https://doi.org/10.1007/s11121-020-01103-9>

- Bowes, L., Joinson, C., Wolke, D., & Lewis, G. (2015, Jun 2). Peer victimisation during adolescence and its impact on depression in early adulthood: prospective cohort study in the United Kingdom. *BMJ*, *350*, h2469. <https://doi.org/10.1136/bmj.h2469>
- Brimblecombe, N., Evans-Lacko, S., Knapp, M., King, D., Takizawa, R., Maughan, B., & Arseneault, L. (2018, Jul). Long term economic impact associated with childhood bullying victimisation. *Soc Sci Med*, *208*, 134-141. <https://doi.org/10.1016/j.socscimed.2018.05.014>
- Brown, V. C., E & Ferguson, C. . (2011). Estimating the prevalence of young people absent from school due to bullying. . *National Centre for Social Research.*, *1*, 1-61.
- Carey, E. G., Ridler, I., Ford, T. J., & Stringaris, A. (2023, Nov). Editorial Perspective: When is a 'small effect' actually large and impactful? *J Child Psychol Psychiatry*, *64*(11), 1643-1647. <https://doi.org/10.1111/jcpp.13817>
- Clarkson, S., Bowes, L., Coulman, E., Broome, M. R., Cannings-John, R., Charles, J. M., Edwards, R. T., Ford, T., Hastings, R. P., Hayes, R., Patterson, P., Segrott, J., Townson, J., Watkins, R., Badger, J., Hutchings, J., & Stand Together, T. (2022, Mar 29). The UK stand together trial: protocol for a multicentre cluster randomised controlled trial to evaluate the effectiveness and cost-effectiveness of KiVa to reduce bullying in primary schools. *BMC Public Health*, *22*(1), 608. <https://doi.org/10.1186/s12889-022-12642-x>
- Dunn G, M. M., Tomenson B. . (2005). Estimating treatment effects from randomized clinical trials with non-compliance and loss to follow-up: the role of instrumental variable methods. *Statistical Methods Medical Research.*, *14*, 369–396. .
- Edwards RT, M. E. (2019). *Applied health economics for public health practice and research*. Oxford University Press.
- Fisher, H. L., Moffitt, T. E., Houts, R. M., Belsky, D. W., Arseneault, L., & Caspi, A. (2012, Apr 26). Bullying victimisation and risk of self harm in early adolescence: longitudinal cohort study. *BMJ*, *344*, e2683. <https://doi.org/10.1136/bmj.e2683>
- Ford, T. M., O; Wolpert, M. . (2014). *Children and Young People's mental health. The Annual Report of the Chief Medical Officer 2013*. (Public Mental Health Priorities; investing in the evidence., Issue.
- Gaffney, H. F., D.P. & Ttofi, M.M. . (2019). Examining the Effectiveness of School-Bullying Intervention Programs Globally: a Meta-analysis. *International Journal of Bullying Prevention*, *1*, 14–31.
- Hadjar, A., & Uusitalo, E. . (2016). Education systems and the dynamics of educational inequalities in low educational attainment: A closer look at England (UK), Finland, Luxembourg, and German-speaking Switzerland. *European Societies*, *18*(3), 264-287.
- Hastings, R. P., & Brown, T. (2002, Feb). Behavioural knowledge, causal beliefs and self-efficacy as predictors of special educators' emotional reactions to challenging behaviours. *J Intellect Disabil Res*, *46*(Pt 2), 144-150. <https://doi.org/10.1046/j.1365-2788.2002.00378.x>

- Huitsing, G., Lodder, G. M. A., Browne, W. J., Oldenburg, B., Van der Ploeg, R., & Veenstra, R. (2020, Jul). A Large-Scale Replication of the Effectiveness of the KiVa Antibullying Program: a Randomized Controlled Trial in the Netherlands. *Prev Sci*, 21(5), 627-638. <https://doi.org/10.1007/s11121-020-01116-4>
- Karna, A., Voeten, M., Little, T. D., Poskiparta, E., Alanen, E., & Salmivalli, C. (2011, Dec). Going to scale: a nonrandomized nationwide trial of the KiVa antibullying program for grades 1-9. *J Consult Clin Psychol*, 79(6), 796-805. <https://doi.org/10.1037/a0025740>
- Karna, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011, Jan-Feb). A large-scale evaluation of the KiVa antibullying program: grades 4-6. *Child Dev*, 82(1), 311-330. <https://doi.org/10.1111/j.1467-8624.2010.01557.x>
- Maslach C, J. S., Schwab RL. . (1996). Maslach Burnout Inventory-Educators Survey (MBI-ES). In J. S. Maslach C, Leiter MP, (Ed.), *MBI Manual*. (3rd ed.). Consulting Psychologists Press.
- Nocentini, A., & Menesini, E. (2016, Nov). KiVa Anti-Bullying Program in Italy: Evidence of Effectiveness in a Randomized Control Trial. *Prev Sci*, 17(8), 1012-1023. <https://doi.org/10.1007/s11121-016-0690-z>
- Olweus, D. (1994, Oct). Bullying at school: basic facts and effects of a school based intervention program. *J Child Psychol Psychiatry*, 35(7), 1171-1190. <https://doi.org/10.1111/j.1469-7610.1994.tb01229.x>
- Olweus, D. (1996). *Revised Olweus Bully/Victim Questionnaire (OBVQ)*. (2006 ed.). APA PsycTests.
- Pöyhönen V., J. J., Salmivalli C. . (2010). What does it take to stand up for the victim of bullying? The interplay between personal and social factors. . *Merrill-Palmer Quarterly*, 56(2), 143–163.
- R., G. (2001). Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry*, 40(11), 1337–1345.
- Salmivalli, C. L., K; Bjorkqvist, K; Osterman, K. . (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, 22, 1-15.
- Salmivalli, C. V., M. . (2004). Connections between attitudes, group norms, and behaviors associated with bullying in schools. . *Int J Behav Dev*, 28, 246–258.
- Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021, Sep 30). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*, 374, n2061. <https://doi.org/10.1136/bmj.n2061>

- Solberg, M. E. O. D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior*, 29(3), 239–268.
- Takizawa, R., Maughan, B., & Arseneault, L. (2014, Jul). Adult health outcomes of childhood bullying victimization: evidence from a five-decade longitudinal British birth cohort. *Am J Psychiatry*, 171(7), 777-784.
<https://doi.org/10.1176/appi.ajp.2014.13101401>
- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007, Nov 27). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes*, 5, 63. <https://doi.org/10.1186/1477-7525-5-63>
- Ttofi, M. M., Farrington, D. P., Losel, F., Crago, R. V., & Theodorakis, N. (2016, Mar). School bullying and drug use later in life: A meta-analytic investigation. *Sch Psychol Q*, 31(1), 8-27. <https://doi.org/10.1037/spq0000120>
- Williford, A., Boulton, A., Noland, B., Little, T. D., Karna, A., & Salmivalli, C. (2012, Feb). Effects of the KiVa anti-bullying program on adolescents' depression, anxiety, and perception of peers. *J Abnorm Child Psychol*, 40(2), 289-300.
<https://doi.org/10.1007/s10802-011-9551-1>

Contributor's statement

JH, RPH, LB, and SC conceived the study and led the funding applications. JH, LB, RPH, TF, RH, and SC participated in the design of the study. LB led the drafting of the manuscript. JS designed the process evaluation. RCJ contributed to the design of the study and RCJ and MB wrote the analysis plan. RTE and JC contributed to the design of the study and design of the cost-effectiveness evaluation. RTE took over the economic evaluation and RTE and HW wrote the health economics protocol and conducted the HE analyses, EC and EOJ led the trial management. MRB led the Birmingham site, RH the South West England site, LB the South East England site and JH the North Wales site. JB led data collection at two sites and contributed to study material design. FLW led on data linkage, and MS on data cleaning. RCW led the PPI work. MB, JB, MRB, RCJ, SC, EC, RTE, TF, RPH, RH, EOJ, PP, JS, JT, RCW, HW, MEW and JH inputted on all drafts of the paper, and our discussion of findings. All authors had full access to all the data in the study and LB had final responsibility for the decision to submit for publication.

Declaration of interests

JH manages the employees in the CEIT registered charity. SC is employed as the UK KiVa Coordinator. RW receives a salary from GwE. The funding bodies were not involved in the study design, in collection, analysis and interpretation of data or in the writing of the manuscript. All other authors declare no competing interests.

Data sharing

De-identified patient data from this study along with a data dictionary can be made available after a signed data access agreement upon reasonable request to the principal investigators of the trial (j.hutchings@bangor.ac.uk and lucy.bowes@psy.ox.ac.uk).

Acknowledgments

The research study has been funded by the UK National Institute for Health and Care Research (NIHR) Public Health Research (17-92-11). We obtained separate funding for the intervention costs from the Rayne Foundation (for participating schools in Southeast England, from GwE – School Effectiveness and Improvement Service for north Wales (for participating schools in Wales), from Children’s Services, Devon County Council (for participating schools in Exeter) and from HSPC Global Services (UK) Ltd (for participating schools in Birmingham and Warwick). We would like to thank all schools that participated in the Stand Together Trial.

Collaborating author names

The Stand Together Group consists of:

Sadja Butt⁵

Chris Chiswell⁵

Mackenzie Fong⁴

Hayley Gains⁹

Helin Gosalia⁴

Anwen Jones³

Bryony Longdon⁹

Siobhan B. Mitchell⁹

Caitlin Murray⁹

Naomi Rose¹

Melanie Varley⁴

*Corresponding author: lucy.bowes@psy.ox.ac.uk

1. Department of Experimental Psychology, Oxford University, OX2 6GG
2. NIHR Oxford Health Biomedical Research Centre, Oxford University, OX3 7JX
3. Bangor University, Bangor, Gwynedd LL57 2DG
4. Centre for Trials Research, Cardiff University, Neuadd Meirionnydd, Heath Park, Cardiff, CF14 4YS, UK
5. Institute for Mental Health, University of Birmingham, Edgbaston, Birmingham B15 2TT and Birmingham Women's and Children's NHS Foundation Trust, Steelhouse Lane, Birmingham, B4 6NH
6. Department of Psychiatry, University of Cambridge, Hershel Smith Building, Cambridge Biomedical Campus, Robinson Way Cambridge CB2 0SZ
7. School of Educational Learning and Communication Sciences, University of Warwick, CV4 8UW
8. Faculty of Health and Life Sciences, Department for Public Health and Sport Sciences, South Cloisters, St Luke's Campus, University of Exeter, Exeter EX1 2LU
9. Birmingham Women's and Children's NHS Foundation Trust
10. GwE, North Wales, Regional School Improvement Service

Research in context

Evidence before this study

Bullied children access more school health, primary care and specialist child mental health services than their counterparts (5) and experience poor mental health into adulthood (3). Systematic reviews indicate that whole-school, multi-component interventions can be effective at reducing bullying. The KiVa intervention was developed and taken to scale in Finland and requires rigorous evaluation in different contexts.

Added value of this study

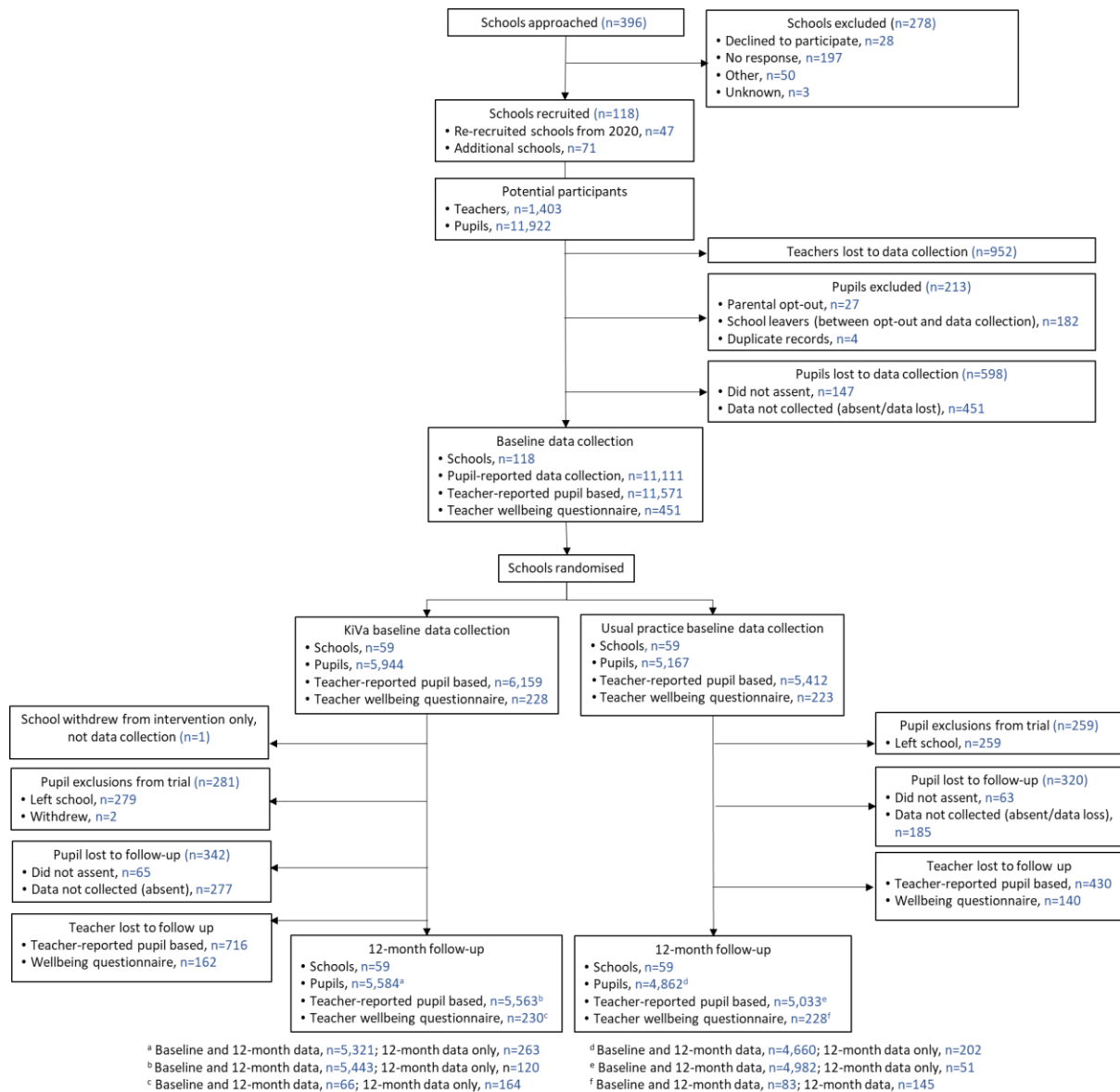
We present evidence from the largest randomised controlled trial to date of the KiVa anti-bullying intervention across 118 primary schools in England and Wales. We found that KiVa

reduced child-reported bullying victimisation by 13%, reduced peer problems and promoted increased empathy among children.

Implications of all the available evidence

Results support the effectiveness and low cost of the KiVa anti-bullying programme as a public health intervention.

Figure 1 The CONSORT Flow Diagram



* Of the 118 schools recruited to the study, 47 had previously completed baseline data for the Stand Together Trial prior to the COVID 19 pandemic and were re-recruited when the study restarted.

Table 1: Baseline characteristics for students and teachers (Data sources: Pupil, TSD and teacher questionnaire)

	KiVa	Usual practice
Number of schools	59	59
Cluster size (Students per school) Median (IQR)	81 (45, 114)	82 (40, 152)
Number of students	5,944	5,167
Trial site		
North Wales	1,556 (26.1%)	1,148 (22.2%)
Southwest England	1,225 (20.6%)	1,106 (21.4%)
Southeast England	1,168 (19.7%)	1,161 (22.5%)
Birmingham	1,995 (33.6%)	1,752 (33.9%)
Student reported characteristics		
Females	2,905/5,933 (49.0%)	2,560/5,166 (49.6%)
Age (in years)*	N=5,937	N=5,165
Mean (SD)	8.7 (0.97)	8.7 (0.94)
OBVQ - Self-reported bullying victimisation[†]		
Yes	1,028/5,074 (20.3%)	955/4,418 (21.6%)
No	4,046/5,074 (79.7%)	3,463/4,418 (78.4%)
OBVQ - Self-reported bullying perpetration[‡]		
Yes	209 (4.0%)	168 (3.7%)
No	5,012 (96.0%)	4,395 (96.3%)
HIFAMS[§] Mean (SD)	11.04 (2.3)	11.01 (2.4)
ETV Mean (SD)	11.68 (5.0)	11.74 (5.0)
PRQ		
Bully Scale		
Never (score=0)	4,541 (86.1%)	4,017 (87.3%)
Sometimes/Always (score>0)	735 (13.9%)	586 (12.7%)
Assistance Scale		
Never (score=0)	4,340 (81.4%)	3,729 (81.5%)
Sometimes/Always (score>0)	995 (18.7%)	846 (18.5%)
Reinforcer Scale		
Never (score=0)	1,026 (19.1%)	947 (20.3%)
Sometimes/Always (score>0)	4,342 (80.9%)	3,722 (79.7%)
Defender Scale		
Never (score=0)	108 (2.1)	116 (2.5)
Sometimes/Always (score>0)	5,173 (97.9)	4,515 (97.5)
Outsider Scale		
Never (score=0)	274 (5.8%)	262 (6.4%)
Sometimes/Always (score>0)	4,432 (94.2%)	3,864 (93.7%)
Teacher reported student based		

Number of students	6,159 ^{††}	5,412 ^{††}
TSDQ** Median (IQR)		
Total difficulties score ^{‡‡} (range 0-40)	4 (1, 9)	4 (1, 9)
Prosocial behaviour score (range 0-10)	9 (6, 10)	9 (6, 10)
Hyperactivity score (range 0-10)	2 (0, 5)	2 (0, 5)
Emotional problems score (range 0-10)	1 (0, 2.5)	1 (0, 2)
Conduct problems score (range 0-10)	0 (0, 1)	0 (0, 1)
Peer problems score (range 0-10)	0 (0, 2)	0 (0, 2)
Teacher reported characteristics		
Number of teachers	228	223
Job role n (%)		
Headteacher/Executive headteacher	3 (1.3%)	0 (0.0%)
Deputy or assistant Headteacher/Senior Leadership teacher/Key stage and subject specific leads	31 (13.7%)	30 (13.6%)
Teacher or associate teacher	116 (51.3%)	116 (52.7%)
Higher Level Teaching Assistant or trainee teacher	4 (1.8%)	12 (5.5%)
Teaching, Classroom or Learning support assistants	55 (24.3%)	53 (24.1%)
Special Educational Needs coordinator / Pastoral Lead / Behaviour Mentor / designated safeguarding leader / Wellbeing	11 (4.9%)	6 (2.7%)
1:1 Supports / Learning Mentors / Additional learning needs	6 (2.7%)	3 (1.4%)
Duration in role n (%)		
Up to 10 years	168 (74.3%)	178 (80.9%)
11 – 20 years	45 (19.9%)	35 (15.9%)
Greater than 20 years	13 (5.8%)	7 (3.2%)
MBI-GS Median (IQR)		
Emotional exhaustion ^{§§}	15 (9, 22)	16 (10, 22)
Cynicism ^{§§}	8 (2, 14)	7 (3, 13)
Professional efficacy	28 (24, 31)	28 (24.5, 31)
WEMWBS** Mean (SD)	49.0 (8.2)	49.2 (8.0)
CBSES^{†††} Median (IQR)	27 (24, 30)	27 (23, 29)

Data are n (%), mean (SD) or median (IQR). UP=Usual practice. OBVQ= Olweus Bully/Victim Questionnaire. HIFAMS= How I Feel About My School. ETV= Empathy Toward Victim Scale. PRQ= Participant Role Questionnaire. TSDQ=Teacher Strength and Difficulties Questionnaire. MBI-GS= Maslach Burn-Out Inventory General Survey. WEMWBS=Warwick-Edinburgh Mental Well-being Scale. CBSES=Challenging Behaviour Self-Efficacy Scale. * Five children self-reported being six years of age despite being in school years 3-5. † Data were missing for 870 (KiVa) and 749 (UP) students. Bullying victimisation defined as “How often have you been bullied at school in the past couple of months?” = 2 or 3 times a month/About once a week/Several times a week. ‡ Data were missing for 723 (KiVa) and 604 (UP) students. Bullying perpetration defined as “How often have you taken part in bullying another pupil(s) at school in the past couple of months?” = 2 or 3 times a month/About once a week/Several times a week. § Data were missing for 269 (KiVa) and 211 (UP) students. HIFAMS scores range from 0 to 14 - higher scores reflect greater happiness. || Data were missing for 338 (KiVa) and 310 (UP) students. ETV scores range from 0 to 21 - higher scores indicate greater empathy for the victim. ** Data were missing for between 2 and 5 students. For all subscales, higher scores indicate more strength or difficulty. †† Teachers reported TSDQs for a higher

number of students compared to the pupils self-reported questionnaires. ‡‡ Total of all subscales excluding Prosocial behaviour subscale. §§ Data were missing for between 10 and 15 teachers. Scores range from 0 to 30. ||| Data were missing for 14 (KiVa) and 11 (UP) teachers. Scores range from 0 to 36. *** Data were missing for 7 (KiVa) and 6 (UP) teachers. WEMWBS scores range from 14 to 70. ††† Data were missing for 8 (KiVa) and 9 (UP) teachers. CBSES scores range from 5 to 35.

Table 2: Primary and secondary outcomes for students (Data sources: Pupil and TSD questionnaire)

	12-month follow-up		Adjusted* intervention effect (95% CI)	p value
	KiVa	Usual practice		
Total students	N=5,584	N=4,862		
Primary outcome				
OBVQ - Bullying victimisation				
Self-reporting bullying victimisation	885 (17.7%)	873 (20.7%)	0.87 [†] (0.78, 0.97)	0.009
Not self-reporting bullying victimisation	4,109 (82.3%)	3,347 (79.3%)
Secondary outcomes				
OBVQ - Bullying perpetration				
Self-reporting bullying perpetration	141 (2.8%)	149 (3.4%)	0.90 [†] (0.80, 1.01)	0.08
Not self-reporting bullying perpetration	4954 (97.2%)	4238 (96.6%)
HIFAMS Mean (SD)	10.3 (2.6)	10.3 (2.6)	0.02 [§] (-0.08, 0.11)	0.68
ETVS Mean (SD)	11.0 (4.9)	10.8 (4.9)	0.45 [§] (0.20, 0.69)	<0.001
PRQ - Defender Scale				
Sometimes/Always (score>0)	4,784 (97.0)	4,190 (97.1)	0.76 [†] (0.59, 0.97)	0.030
Never (score=0)	148 (3.0)	127 (2.9)
Teacher reported student based				
Total students	N=5,563	N=5,033		
TSDQ				
Total score				
Normal	4,918 (88.4%)	4,426 (88.0%)		
Borderline	279 (5.0%)	273 (5.4%)	reference	
Abnormal	364 (6.6%)	332 (6.6%)	1.01 [†] (0.78, 1.30)	0.96
Prosocial score				
Normal	4,358 (78.4%)	3,835 (76.2%)		
Borderline	386 (6.9%)	385 (7.7%)	reference	
Abnormal	817 (14.7%)	811 (16.1%)	0.93 [†] (0.77, 1.11)	0.40
Hyperactivity score				
Normal	3,893 (70.0%)	3,502 (69.6%)		
Borderline	384 (6.9%)	341 (6.8%)	reference	
Abnormal	1,284 (23.1%)	1,187 (23.6%)	1.02 [†] (0.91, 1.14)	0.73
Emotional score				
Normal	4,147 (74.6%)	3,680 (73.2%)		
Borderline	470 (8.4%)	445 (8.8%)	reference	
Abnormal	944 (17.0%)	906 (18.0%)	1.01 [†] (0.83, 1.23)	0.93
Conduct problems score				
Normal	5,219 (93.9%)	4,696 (93.3%)		

Borderline	141 (2.5%)	149 (3.0%)	reference	
Abnormal	201 (3.6%)	186 (3.7%)	0.87 [†] (0.66, 1.16)	0.34
Peer problem score				
Normal	4,775 (85.9%)	4,242 (84.3%)		
Borderline	336 (6.0%)	357 (7.1%)	reference	
Abnormal	450 (8.1%)	432 (8.6%)	0.85 [†] (0.75, 0.96)	0.011

Data are n (%). Mean (SD) or median (IQR). OBVQ= Olweus Bully/Victim Questionnaire. HIFAMS= How I Feel About My School. ETV= Empathy Toward Victim Scale. PRQ=Participant Role Questionnaire. TSDQ=Teacher Strength and Difficulties Questionnaire. *Adjusted for free school meals, KS2 school size, corresponding baseline measure, age and gender of pupil, and clustered within 118 schools within 4 recruitment sites; [†]Adjusted odds ratio for KiVa versus usual practice; §Adjusted difference in means for KiVa minus usual practice.

Table 3: OBVQ maximum self-reported victimisation and perpetration scores – post-hoc analysis

	Baseline		12-month follow-up		Adjusted* intervention effect (95% CI)	p value
	KiVa	Usual Practice	KiVa	Usual practice		
Total students	5,944	5,167	N=5,584	N=4,862		
Bullying victimisation					0.80 (0.71, 0.90)	<0.001
1 – Not at all	2,818 (47.4)	2,382 (46.1)	1,314 (23.8)	946 (19.7)		
2 – Once or twice	1,228 (20.7)	1,081 (20.9)	2,113 (38.2)	1,810 (37.7)		
3 – 2 or 3 times a month	389 (6.5)	376 (7.3)	727 (13.1)	707 (14.7)		
4 – About once a week	259 (4.6)	264 (5.1)	576 (10.4)	543 (11.3)		
5 – Several times a week	380 (6.4)	315 (6.1)	803 (14.5)	795 (16.6)		
Bullying perpetration					0.76 (0.69, 0.84)	<0.001
1 – Not at all	4,387 (73.8)	3,830 (74.1)	3,470 (63.3)	2,756 (57.8)		
2 – Once or twice	625 (10.5)	565 (10.9)	1,546 (28.2)	1,511 (31.7)		
3 – 2 or 3 times a month	93 (1.6)	81 (1.6)	218 (4.0)	230 (4.8)		
4 – About once a week	58 (0.98)	47 (0.91)	104 (1.9)	130 (2.7)		
5 – Several times a week	58 (0.98)	40 (0.8)	146 (2.7)	144 (3.0)		

Data are n (%).OBVQ= Olweus Bully/Victim Questionnaire.

*Adjusted for free school meals, KS2 school size, corresponding baseline measure, age and gender of pupil, and clustered within 118 schools within 4 recruitment sites