Using fuzzy-set qualitative comparative analysis to explore causal pathways to reduced bullying in a whole-school intervention in a randomized controlled trial

Warren, Emily* (1); Melendez-Torres, G.J. (2) ; Bonell, Chris (1)
1-Public Health, Environments and Society, London School of Hygiene and Tropical Medicine, London, UK

2-College of Medicine and Health, University of Exeter, Exeter, UK

*15-17 Tavistock Place, London WC1H 9SH, UK; Emily.Warren@lshtm.ac.uk;
+44(0)20 7927 2408

Funding: This work was supported by the National Institute of Health Research (NIHR) in England under its Public Health Research Board (12/153/60) and the Educational Endowment Foundation.

Competing interests: The authors declare that they have no competing interests.

Ethics approval and consent: Ethics approval was obtained from the University College London ethics committee (ref5248/001). Written, informed consent was obtained from head teachers for random allocation. Intervention facilitators, staff, and students were provided with information sheets and consent forms, and were given the opportunity to ask questions and opt out.

Acknowledgments: We would like to acknowledge the participants for their contribution to this study. We would also like to acknowledge Adam Walden for his assistance designing the tables and figures.

Using fuzzy-set qualitative comparative analysis to explore causal pathways to reduced bullying in a whole-school intervention in a randomized controlled trial

Learning Together is a whole-school intervention, evaluated using a randomized controlled trial in southeast England, which reduced bullying and improved physical and mental health. This paper examines trial data using fuzzy-set qualitative comparative analysis to test hypotheses derived from embedded qualitative research about potential causal pathways. Analyses suggested that the intervention worked via three mechanisms: improving student commitment to school; improving student pro-social skills; and de-escalating conflict and bullying. Evidence also suggests that these mechanisms may have been activated via other resources in schools not receiving Learning Together resources. The analysis suggests which contextual features may be important for activating these mechanisms.

Keywords: Qualitative comparative analysis; school environment; bullying; whole-school interventions; realist evaluation; adolescent mental health; mechanisms

Introduction

Around one third of UK young people report bullying (Brooks et al. 2015) with impacts on physical and mental health decades later.(Copeland et al. 2013; Price et al. 2013) Bullying involves intentionally hurtful, repetitive, physically, verbally or socially aggressive behaviour targeting those with less power.(Cantone et al. 2015) Systematic reviews (Evans, Fraser, and Cotter 2014; Cantone et al. 2015; Gaffney, Ttofi, and Farrington 2021) have examined the impact of school-based anti-bullying interventions,

such as social and emotional learning (SEL) curricula, restorative practices (RP) and whole-school interventions. One meta-analysis found SEL results in significantly lower rates of bullying.(Durlak et al. 2011) Two previous randomized controlled trials (RCTs) of RP found no significant decrease in bullying(Acosta et al. 2019) and positive impacts which were not sustained.(Cross et al. 2018) Despite evaluating complex multi-component interventions, few studies have explored intervention mechanisms or how effects vary with context, which is key to assessing intervention transferability.(Hanckel et al. 2021)

The Learning Together intervention

Learning Together (LT) was a whole-school intervention that aimed to decrease bullying and improve physical and mental health among secondary-school students. LT was delivered over three years and involved both RP and SEL. Resources included an intervention manual; annual reports of student needs from an annual survey; a facilitator to guide delivery in the first two years; a SEL curriculum; and staff training in RP. These aimed to enable: regular staff/student AG meetings to review needs, review rules and policies and oversee delivery of SEL and RP. RP could be preventive or in response to bullying and conflict.(Bonell et al. 2018) Each school received the same intervention resources but schools could tailor implementation to local needs (Figure 1).

The theory of change was informed by the theory of human functioning and school organisation (Markham and Aveyard 2003), which proposes that schools aim to develop students' "practical reasoning" and "social affiliations" via schools' "instructional" (curriculum) and "regulatory" (social norms and behavioural expectations) orders. Schools can "reframe" school practices to re-centre on student needs and erode "boundaries" between students and between students' academic and

broader development. This is theorised to build student commitment to the instructional and regulatory orders and thereby build practical reasoning and affiliation, and ensure students make healthier choices.

[INSERT Figure 1]

An RCT of LT found it effective in reducing bullying victimization as well as substance use and contact with police, and improving health-related quality of life, mental wellbeing and psychological difficulties compared to the control group. The RCT involved qualitative research which was used to build theory about intervention mechanisms and how these might vary with context.(Warren et al. 2020) Firstly, in some schools, bullying may have decreased by building students' participation in and commitment to school via students and staff sharing experiences and developing empathy working on the AG. Secondly, bullying appears to have decreased via improving students' pro-social behaviors via RP and SEL. Thirdly, bullying may have been reduced via de-escalation among students involved in bullying incidents via RP, ensuring bullies understood the consequences of their actions.(Warren et al. 2020) However, the qualitative data were from three case-study schools and we were interested in whether these mechanisms might have operated in other schools within the trial.

Building on the theory of change and this qualitative analysis, we developed hypotheses about intervention mechanisms and how these interacted with context to generate outcome. These were that bullying could be decreased via: i) building student commitment to school (in schools with a participative ethos); ii) teaching pro-social skills (in schools where there was evidence for deficits in these and/or where students felt unsafe); and/or iii) de-escalating bullying among a core group of students (in

schools with a high baseline rate of victimisation).(Warren et al. 2020) These hypotheses were framed in terms of how mechanisms might generate outcomes contingent on context. This approach was informed by realist evaluation, which seeks to examine what works for whom under what conditions, and therefore inform assessments of transferability. Within this approach, interventions are understood to "work" via introducing resources into contexts which agents then use and which triggering cognitive or social process that generate outcomes in interaction with local context.(Pawson and Tilley 1997)

To assess whether the above hypotheses appear plausible in other trial schools, we used fuzzy set qualitative comparative analysis (fsQCA). We were interested in whether these mechanisms occurred in intervention schools via provision of intervention resources but also whether some control schools were activating the overall mechanism described in the above hypotheses via existing resources.

Qualitative comparative analysis

QCA was developed by Charles Ragin as a tool for understanding macrosociological change.(Ragin 2004, 2009; Ragin and Becker 1992) In contrast to regression-based analysis, which examines statistical associations between multiple variables, QCA employs Boolean algebra (combinations of conditions linked by AND, NOT and OR) to examine what 'pathways' (complex combinations of the presence or absence of factors) co-occur with certain outcome among a set of cases. Researchers using QCA assume a "configurational" view of causation whereby multiple conditions combine to generate change.(Melendez-Torres et al. 2019; Thomas, O'Mara-Eves, and Brunton 2014) Within QCA, conditions are understood to be either "*sufficient*" or "*necessary*" for an outcome to occur. A condition is *sufficient* if the outcome is always

present with the condition regardless of other factors. A condition is *necessary* if the outcome cannot occur when the condition is absent.(Rihoux and Ragin 2008)

The first step in QCA is constructing a *data table* with conditions and outcomes. In this analysis, each school formed a row with columns indicating conditions (contextual features, markers of hypothesized mechanisms and outcomes). In crisp-set (cs) QCA, conditions and outcomes are dichotomised as 1 or 0, respectively indicating they are fully cases or not cases. However, in fuzzy-set (fs) QCA, data are directly or indirectly *calibrated* so they have a value somewhere between 0 and 1. *Direct calibration* involves researchers examining the empirical distribution of a condition and selecting a threshold above which change in plainly evident, a threshold below which change is plainly not evident, and the cross-over point at which the magnitude of change might cross over from more evident than not to less evident. *Indirect calibration* involves researchers or experts assigning scores (also called truth values) based on their knowledge of a subject area: for example, 0 (fully not a case), 0.33 (more not a case), 0.67 (more a case) or 1 (fully a case). By transforming all scores into truth values which fit between 0 and 1, QCA generates comparable values of 'caseness.'

After completing the data table, analysis moves from individual cases to understanding the combinations of conditions associated with the outcome using a *truth table*.(Thomas, O'Mara-Eves, and Brunton 2014) A truth table presents each different configuration of conditions as a row and reports how many cases are within each set. In both fuzzy- and crisp-set QCA, data in truth tables are presented as binary, with truth values <0.5 being reduced to 0 and values >0.5 becoming 1. Truth tables show the combinations of conditions which do and do not co-occur with outcomes. Configurations can be positive (all cases within the set have the outcome), negative (all cases within the set do not have the outcome), contradictions (the same combination of

conditions produce different results) and remainders (possible configurations with no empirical manifestations to test them). Truth tables report *consistency* (i.e. the proportion of cases within each set that also have the same outcome) and *coverage* (i.e. how much of the outcome is explained by the model). If consistency is low, there is weak or contradictory evidence that a combination of conditions co-occurs with the outcome. If coverage is low, this suggests that the model is missing key explanatory conditions. QCA is iterative, with conditions added to models to improve consistency and coverage.(Thomas, O'Mara-Eves, and Brunton 2014) QCA models become difficult to interpret with too many conditions so additional conditions are chosen judiciously based on detailed knowledge of the topic.(Rihoux and Ragin 2008) The final step is *Boolean minimization* whereby a condition is removed from a combination if neither its presence nor absence affects the emergence of the outcome.

Aims

This paper seeks to elaborate and test whether the above hypotheses appear consonant with the pattern of contingencies found in fsQCA. We developed a series of context-mechanism-outcome configurations (CMOC) to elaborate the above hypotheses, which provided the basis for fsQCA. The first CMOC explored data from both arms of the trial to assess if the overarching mechanism might be reducing bullying across some control as well as intervention schools:

> Overarching mechanism: When students' commitment to school increased (M), students' pro-social skills were improved (M), and/or bullying was de-escalated among a core group of students (M), bullying would decrease (O) regardless of which arm schools were allocated to within the trial (C).

We then further explored each of these three sub-mechanisms as CMOCs in relation to LT resources and hypothesized that:

> Sub-mechanism 1) In schools with a pre-existing ethos of wanting to involve students in decision-making (C), improving relationships between students and staff on the AG (M), students feeling like they made a positive contribution to the school via implementation of AG activities (M), and/or feeling the AG connected them to other people in the school to make positive changes (M) would increase student participation in school decisions (O).
>
> Sub-mechanism 2) In schools where students lack strong pro-social skills or where the development of pro-social skills was a staff priority (C), and/or where students felt unsafe in school (C), delivering a social and emotional skills-based curriculum (M) and/or implementing preventative RPs (M) would improve pro-social skills (O).
>
> Sub-mechanism 3) In schools with high bullying victimization at baseline (C), sufficient staff trained in responsive RP (M), high incidence of the use of responsive RP (M), perpetrators feeling empathy (M), and/or accepting responsibility and punishment for their actions (M) would decrease bullying (O).

Methods

Trial Methods

Details about intervention and trial methods are published elsewhere.(Bonell et al. 2014; Bonell et al. 2018) Briefly, LT was evaluated using a cluster RCT in 40

mainstream state secondary schools in south-east England with a government inspection rating higher than "inadequate". Baseline surveys involved paper-based questionnaires completed by students in year 7 (age 11-12). Similar surveys were conducted at 24- and 36-months post-baseline when students were in years 9 and 10 to assess short/medium term impacts of institutional change. Staff were also surveyed at these time-points.

Data sources

Data used in the overarching model were drawn from baseline and endline student surveys, and included various measures for each mechanism and bullying victimization as the primary outcome (Table 1). The Beyond Blue School Climate Questionnaire(Sawyer et al. 2010) was used to assess the mechanism involving improving student commitment to school. The Strengths and Difficulties Questionnaire(Goodman 1997) was used to assess the mechanism involving improved pro-social skills. The Edinburgh Study of Youth Transitions and Crime (ESYTC) scale(Smith 2006) was used to assess the mechanism involving de-escalating bullying. The Gatehouse Bullying Scale (GBS) (Bond et al. 2007) was used to assess the outcome of reduced bullying.

Data sources used in the sub-mechanisms were drawn from the above sources plus the 24-month student survey, a survey of AG members, staff surveys, interviews with students involved in restorative conferences, process evaluation records on intervention fidelity, and staff surveys reports on use of RP (Table 1). Measures of context were collected in the first year of the trial. Outcomes used either endline scores or changes between baseline and endline. Mechanisms were represented by change over time or data collected at 24-months post-baseline (Tables 2-4).

QCA

To assess CMOC 1, an overarching model assessed whether the three hypothesized sub-mechanisms, individually or in combination, led to decreased bullying across all schools. As explained above this encompassed both arms as we were interested in whether schools could trigger such mechanisms drawing on intervention and/or existing resources.

For the models examining what sub-mechanisms were important in different contexts, we examined whether mechanism markers co-occurred with markers of outcomes relating to increasing participation in school decision-making, improving pro-social skills, and decreasing bullying, respectively. The analysis of these sub-mechanisms focused only on intervention schools because we aimed to understand whether mechanisms arising directly from the use of LT resources were accompanied by changes in the hypothesised proximal outcomes.

Data tables

It is important that those establishing anchor points and interpreting the data have 'thick' knowledge of the cases.(Hanckel et al. 2021) The lead author of this paper was responsible for collecting much of the qualitative data, had worked directly with all the schools in the trial, and had led the earlier qualitative analyses informing our hypotheses. Two study authors (*initials to be added later*) examined data to decide anchor points based on two key criteria. Cut-off points needed to represent a change which would be of public health significance and to provide us with a reasonable degree of distribution as to whether schools met or did not exhibit the conditions of interest. For example, we started by examining, as our overall outcome, the school average change in bullying victimization (using the GBS), which varied from -62 to +7%. After examining any natural gaps in the data which may indicate qualitatively different levels of casesness and the conditions' distribution, we discussed what level of bullying

reduction might be of public health significance. We decided that schools with greater than 50% reductions were fully cases, schools with less than 15% decrease in bullying were not fully cases and schools with 30% reductions were most ambiguous. It is important to note that even though a 15% reduction in bullying is notable, schools achieving less than this were the least successful in our sample, and cut-offs must be established based on the included data so that analysis can continue. The schools' GBS scores were then directly calibrated in STATA, giving every school a truth value between 0 and 1. This process of examining scores, establishing cut-offs, and calibrating data was repeated for all other conditions in all our models.

Truth tables

Generated using the Tosmana (Cronqvist 2011) software, we assessed each model's consistency and coverage in their respective truth tables. Rihoux and Ragin recommend that consistency scores should be >0.75 and coverage scores should be >0.85.(Rihoux and Ragin 2008) We valued higher consistency over coverage because consistency is a better tool for showing whether the data supported our hypotheses. While low coverage may be a problem, it also indicates that other explanations outside the model may contribute to the outcome. This was expected because schools take diverse action to reduce bullying not restricted to those enabled by LT resources. When consistency or coverage were too low, new concepts suggested by our intervention theory of change and qualitative research, were added. To avoid data-dredging, we stopped adding conditions when there were no further measures that aligned with the hypotheses emerged directly from the qualitative findings. For example, in the first iteration of the overarching mechanism, consistency was high at 90% and coverage was moderate at 55%. Therefore, we added indicators for the learning of conflict resolution skills and the decreasing of conduct problems, which are both important for improving pro-social skills.(Goodman 1997)

Our first iteration of sub-mechanism 3, lacked sufficient explanatory power. Therefore, we added measures of perpetrators coming to feel empathy and accept responsibility through RP. This meant we reduced the model to include only the 14 schools for which interview data provided markers of these. In these schools, interviews were conducted with students who had been involved in a restorative conference either as a bully or victim. To quantify interview data, we created spreadsheets identifying which school the data came from, key quotes explaining the situation, and any data that expressed feeling empathy (or not) and accepting responsibility and, when applicable, accepting responsibility (or not). These quotes were then given a score of 0 if they did not express any change in attitude, .33 if they expressed very limited change in attitude, .67 if they recognized a change in their attitudes but it was not complete, and 1 if they described the intervention as having a meaningful change in their views and actions.(Pratchett et al. 2009) When multiple accounts were taken from one school, the scores were averaged and then directly calibrated in STATA.

When new conditions were added to the data tables, they were subjected to the aforementioned calibration process, and truth tables were re-run to assess the impact of their inclusion on coverage. Tables 1-4 show which variables were included in the original models, and which were added later.

[INSERT TABLES 1-4]

Boolean minimization

We identified configurations in which the same outcome appeared and involved the same pathway except for the presence of one condition, indicating its lack of causal impact. We then reported the simplified solution. When reading QCA solutions, the presence of conditions are written in capital letters, the absence of conditions are written in lower-case letters, and * is read as “and.”

Results

Overarching mechanism

Our first model explored our hypothesis that schools decreased bullying by improving students' commitment to school, improving pro-social skills, and/or de-escalating bullying, regardless of trial arm. We identified 13 pathways (data from 21 schools) that did not decrease bullying and 15 pathways (data from 19 schools) that did (Online Appendix Table 1.) These 15 effective causal pathways were minimized to nine solutions as an automatic output of the Tosmana software. Consistency across solutions was very high (97.43%) meaning that all of the schools following one of these pathways reduced bullying. Coverage was moderate at 62% meaning that 62% of the decrease in bullying could be explained by these combinations of conditions (Table 5). While 13 schools in both trial arms decreased bullying, schools in the intervention arm had higher truth values (0.67085332 compared to 0.5589047), indicating that they experienced greater decreases in bullying victimization. Moreover, across the intervention schools, there was evidence that mechanisms activated 63 times compared to 58 times in control schools.

[INSERT TABLE 5]

School 31 aligned most closely to our hypothesis, having all conditions except the learning of conflict resolution, and experienced one of the greatest decreases in bullying (truth value= 0.9766525). The pathway with the greatest explanatory power (role*rpsolving*CP*aggress) suggested that in schools that did not improve student commitment, implement learning of conflict resolution or de-escalate bullying, but that

did decrease conduct problems, bullying was meaningfully lessened. This pathway explained 20% of the bullying decrease in the trial. A similar configuration was also effective (belong*role*prosocial*rpsolving*CP; coverage = 0.14792643). Decreasing conduct problems appeared to be the most important mechanism for decreasing bullying, activated in 12 of 15 effective solutions. Other configurations required the activation of multiple mechanisms. For example, in schools that did not improve student commitment and did not de-escalate bullying but did improve pro-social skills and taught students to resolve conflict (belong*role*PROSOCIAL*RPSOLVING*aggression), this explained 14% of the model's effectiveness.

Sub-mechanisms

We then examined each sub-mechanism, including contextual features.

Sub-mechanism 1: Improved commitment

Our analysis found evidence for two pathways that improved participation. However, the contextual feature (baseline ethos of wishing to involve students in decision-making) was not necessary in either effective pathway. Excluding this contextual feature, schools 22 and 27 met the conditions for our hypothesized mechanisms and both increased students' participation in decisions. After Boolean minimization, the reduced causal pathway could be expressed as ACTIONS*ATTITUDE CHANGE→PARTICIPATION meaning that students feeling they made a positive contribution via implementation of AG activities and the AG initiating a change in student attitudes to school led to increased student participation in decision-making. This effect was felt in four schools. Consistency was good at 84.10% but coverage was low at 32.40%. This means that the majority of the ways through

which increased student participation in decisions was achieved outside of these conditions. (Table 5; Online Appendix Table 2)

Sub-mechanism 2: Improved pro-social skills

No school had all of our hypotheses' conditions (students lack pro-social skills, students feel unsafe at school, the school delivering the SEL curriculum with fidelity and using preventative RPs). Two sets of configurations were similar to our model, each containing one of the two contextual features and both mechanisms activated, but only one configuration (FEEL UNSAFE*CURRICULUM*PREVENTATIVE RP) led to improvements in pro-social behaviour (Online Appendix Table 3). Seven schools improved pro-social skills via five other pathways, while six pathways (with data from 13 schools) did not lead to improvements in pro-social skills.

The data indicated that students feeling unsafe at school was a more important contextual feature than students lacking pro-social skills or having staff value their development. Feeling unsafe was an important condition in four of five configurations while lacking pro-social skills was active in only two. There was one configuration in which none of our hypothesized mechanisms activated but pro-social skills were still improved, indicating that other mechanisms, unconnected with the trial, were simultaneously occurring.

The delivery of the SEL curricula with fidelity was only present in one effective pathway but was present in half of the ineffective pathways, indicating that this delivery of curriculum had a negligible impact on improving pro-social skills. Consistency was acceptable at 76.34% and coverage was moderate at 54.27%. The pathway with the greatest explanatory power (38%) was students lacking pro-social skills or their development was seen as a priority by staff, not delivering the curriculum with fidelity, and using preventative RP (Table 5). This aligned closely with our qualitative research

suggesting that teachers often did not like the curriculum but felt RP was useful in improving student behaviour.(Warren et al. 2019; Warren et al. 2020)

Sub-mechanisms 3: De-escalate conflict amongst a core group of students

Of the 14 schools with pertinent data, we identified five pathways (with data from six schools) that decreased bullying (Online Appendix Table 4). School 3 contained all our hypothesized conditions and bullying in that school decreased. High baseline bullying victimization was present in five of the ineffective combinations, indicating that it may not be important for the activation of the investigated mechanisms. Consistency was high at 90.24% and coverage was moderate at 59.7%. The pathway that explained the greatest decrease in bullying (33% coverage) was not having high bullying victimisation at baseline, having sufficient staff trained in RP, but not needing perpetrators of bullying to accept responsibility. Another effective configuration was not having high bullying victimization at baseline, sufficient staff training in RP, and perpetrators not feeling empathy or accepting responsibility (23% coverage) (Table 5 and Online Appendices Table 4). In both of the above configurations, having sufficient staff train in RP may have caused a decline in bullying by modifying and correcting many small instances of poor behaviour which can create a climate where bullying is tolerated.

Discussion

Summary of key findings

This analysis suggested that bullying can be reduced via improving commitment to school, improving pro-social skills and de-escalating bullying among a core group of offending students. It also suggested that the provision and use of LT resources enabled the activation of mechanisms which contributed to reducing bullying. Methodologically, this analysis suggests that QCA is a useful approach for determining what combinations of conditions co-occur with an impact on outcomes. Our analyses suggested that control schools were able to activate the same mechanisms using other resources (albeit to a lesser extent), indicating the possible validity of the mid-range theory which informed this intervention's theory of change. The model also indicates that of the three investigated sub-mechanisms, the most consistently effective appeared to be improving students' pro-social skills, since at least one indicator of improved pro-social skills was indicated in all of the effective solutions, the most common being decreased conduct problems. Indicators of improving commitment were present in six pathways and de-escalated bullying was found in four. While no school met all of the conditions of our overarching hypothesis, the school that most nearly did so experienced one of the largest decreases in bullying victimization, suggesting that our overarching hypothesis is a plausible pathway through which bullying can be decreased. Coverage was lower, advancing our belief that schools were likely undertaking other activities not related to our intervention which contributed to decreased bullying.

When exploring the sub-mechanisms, de-escalating bullying had the highest coverage in the combined solution, indicating that having sufficient staff trained in RP and ensuring perpetrators feel empathy and accept responsibility may be effective strategies to decrease bullying. The presence of hypothesized contextual features was

less important than the activation of hypothesized mechanisms in the generation of improved outcomes, suggesting that these strategies may be useful across school contexts.

To increase commitment, the two most important mechanisms appeared to be creating new roles for students and the AG changing student attitudes towards school. Evidence also indicates that preventative RPs help develop student pro-social skills. Finally, the evidence fully supported our hypothesis about de-escalating bullying.

QCA enabled us to look beyond single-causal explanations of causality to focus on generative explanations in terms of combinations of contextual features and mechanisms, suggesting there were multiple pathways to the same outcome.(Sager and Andereggen 2012) Our research also suggested that, while one part of a sub-mechanism may be sufficient to decrease bullying in some schools, multiple mechanisms have to be activated together to disrupt the mechanisms which generate bullying in other schools.(Bonell et al. 2020) Within our three sub-mechanisms, we found evidence of impact for 15 of 20 intervention schools, with only two schools activating more than one sub-mechanism. In practice this may indicate that schools can select which activities to focus on depending on their needs and abilities. In some schools this may mean focusing on students who are regularly aggressive while in others it might mean creating opportunities for students and staff to build bonds outside of hierarchical classroom settings.

Weaknesses and strengths of this study

Our study has important limitations. Firstly, our over-arching model did not account for the contextual features earlier qualitative analysis indicated would be important for the activation of mechanisms. Within an already-large model, additional variables would

have made interpretation impossible. Secondly, some of the mechanisms we identified as important were not predicted before the trial began so we lacked measures and therefore used imperfect proxy indicators in these analyses. In two instances, we had to quantify interview data to develop markers of the mechanisms that the qualitative data indicated were crucial for change. This may mean that the measures are of limited validity as the interview guide did not include specific questions about feeling empathy or accepting responsibility. Thirdly, our truth tables show that we identified contradictory configurations whereby the same combination of conditions led to different outcomes in different schools. As we were unable to expand the models further without data dredging, they remain unresolved. Fourthly, our process evaluation asked school staff about whether addressing bullying was a priority but we lacked sufficient detail to report on what other activities may have been ongoing to reduce bullying at the same time as the intervention. Finally, it is possible that some pathways that were identified as effective were coincidental rather than causal.

Where possible, we sought to bolster our research against QCA's well-known shortcomings. For example, accounting for the passage of time is difficult in QCA where data are generally cross-sectional. To maximize the strengths of the longitudinal data, all of our contextual features in our CMOs were taken from data in the first year of the study, and our outcomes related to either the percent change over three years or prevalence at endline.

One key strength of this study is the use of data from both trial arms. Other QCA studies have been nested within larger RCTs but have only focused on pathways to change within the intervention arm.(Short, Eadie, and Kemp 2020) By exploring the overarching mechanism also drawing on data from control schools, we are able to examine the social processes through which improvements might be made without

focusing solely on intervention resources, enabling us to assess whether or not our hypotheses were generalizable to schools in the control arm. Another strength of this analysis is that plausible pathways outside of our hypotheses and ineffective pathways were also highlighted. Finally, this study is part of a theoretically informed evaluation and builds on an earlier qualitative study which drew on 66 interviews in order to inform the CMOs tested here.

Conclusion

While our sample was not always large enough to have empirical manifestations of our exact hypotheses, our CMO configurations were generally supported. Evidence from this study suggests that student participation in decision-making may be an avenue through which bullying can be decreased. Preventative RPs may be sufficient to improve pro-social skills among students, and training staff in responsive RP appears to de-escalate conflict amongst a core group of aggressive students. The SEL curricula appears to be the least effective resource provided through the LT intervention. Even when the hypothesized contextual features were not present, mechanisms were often still able to activate, indicating that a wide range of schools could benefit from the implementation of either preventative and/or responsive RP depending on the local manifestations of bullying. Our analysis, informed by realist evaluation, suggests that when given resources, agents will deploy them in locally relevant ways. While some schools needed to activate numerous mechanisms to improve outcomes, in others, changes could be achieved more easily. Our analysis also showed that using different resources, control schools also achieved reductions in victimization- although slighter- via the same mechanisms, indicating that our hypotheses are plausible and potentially generalizable. QCA is a useful approach within trials and it can be used to explore phenomena in both trial arms.

References

Acosta, Joie, Matthew Chinman, Patricia Ebener, Patrick S Malone, Andrea Phillips, and Asa Wilks. 2019. "Evaluation of a Whole-School Change Intervention: Findings from a Two-Year Cluster-Randomized Trial of the Restorative Practices Intervention." *Journal of youth and adolescence* 48 (5):876-890.

Bond, Lyndal, Sarah Wolfe, Michelle Tollit, Helen Butler, and George Patton. 2007. "A comparison of the Gatehouse Bullying Scale and the Peer Relations Questionnaire for students in secondary school." *Journal of School Health* 77 (2):75-79.

Bonell, Chris, Elizabeth Allen, Deborah Christie, Diana Elbourne, Adam Fletcher, Richard Grieve, Rosa LeGood, Anne Mathiot, Stephen Scott, and Meg Wiggins. 2014. "Initiating change locally in bullying and aggression through the school environment (INCLUSIVE): study protocol for a cluster randomised controlled trial." *Trials* 15 (1):1.

Bonell, Chris, Elizabeth Allen, Emily Warren, Jennifer McGowan, Leonardo Bevilacqua, Farah Jamal, Rosa Legood, Meg Wiggins, Charles Opondo, and Anne Mathiot. 2018. "Effects of the Learning Together intervention on bullying and aggression in English secondary schools (INCLUSIVE): a cluster randomised controlled trial." *The Lancet* 392 (10163):2452-2464.

Bonell, Chris, Audrey Prost, GJ Melendez-Torres, Calum Davey, and James R Hargreaves. 2020. "Will it work here? A realist approach to local decisions about implementing interventions evaluated as effective elsewhere." *J Epidemiol Community Health*.

Brooks, F, J Magnusson, E Klemera, K Chester, N Spencer, and N Smeeton. 2015. "HBSC England national report 2014." *Hatfield, UK: BMJ Publishing Group*.

Cantone, Elisa, Anna P Piras, Marcello Vellante, Antonello Preti, Sigrun Daníelsdóttir, Ernesto D'Aloja, Sigita Lesinskiene, Mathhias C Angermeyer, Mauro G Carta, and Dinesh Bhugra. 2015. "Interventions on bullying and cyberbullying in schools: A systematic review." *Clinical practice and epidemiology in mental health: CP & EMH* 11 (Suppl 1 M4):58.

Copeland, William E, Dieter Wolke, Adrian Angold, and E Jane Costello. 2013. "Adult psychiatric outcomes of bullying and being bullied by peers in childhood and adolescence." *JAMA psychiatry* 70 (4):419-426.

Cronqvist, Lasse. 2011. "Tosmana." *Tool for Small-N Analysis [Computer Programme], Version* 1 (2.0).

Cross, Donna, Therese Shaw, Melanie Epstein, Natasha Pearce, Amy Barnes, Sharyn Burns, Stacey Waters, Leanne Lester, and Kevin Runions. 2018. "Impact of the Friendly Schools whole-school intervention on transition to secondary school and adolescent bullying behaviour." *European Journal of Education* 53 (4):495-513.

Durlak, Joseph A, Roger P Weissberg, Allison B Dymnicki, Rebecca D Taylor, and Kriston B Schellinger. 2011. "The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions." *Child development* 82 (1):405-432.

Evans, Caroline BR, Mark W Fraser, and Katie L Cotter. 2014. "The effectiveness of school-based bullying prevention programs: A systematic review." *Aggression and Violent Behavior* 19 (5):532-544.

Gaffney, Hannah, Maria M Ttofi, and David P Farrington. 2021. "Effectiveness of school-based programs to reduce bullying perpetration and victimization: An

updated systematic review and meta-analysis." *Campbell Systematic Reviews* 17 (2):e1143.
Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: a research note." *Journal of child psychology and psychiatry* 38 (5):581-586.
Hanckel, Benjamin, Mark Petticrew, James Thomas, and Judith Green. 2021. "The use of Qualitative Comparative Analysis (QCA) to address causality in complex systems: a systematic review of research on public health interventions." *BMC public health* 21 (1):1-22.
Markham, Wolfgang A, and Paul Aveyard. 2003. "A new theory of health promoting schools based on human functioning, school organisation and pedagogic practice." *Social science & medicine* 56 (6):1209-1220.
Melendez-Torres, GJ, Katy Sutcliffe, Helen ED Burchett, Rebecca Rees, and James Thomas. 2019. "Developing and testing intervention theory by incorporating a views synthesis into a qualitative comparative analysis of intervention effectiveness." *Research synthesis methods* 10 (3):389-397.
Pawson, Ray, and Nick Tilley. 1997. *Realistic evaluation*: Sage.
Pratchett, Lawrence, Leila Thorp, Melvin Wingfield, Vivien Lowndes, and Ruby Jabbar. 2009. "Preventing Support for Violent Extremism through Community Interventions: A Review of the Evidence-Rapid Evidence Assessment Full Final Report."
Price, Lawrence H, Hung-Teh Kao, Darcy E Burgers, Linda L Carpenter, and Audrey R Tyrka. 2013. "Telomeres and early-life stress: an overview." *Biological psychiatry* 73 (1):15-23.
Ragin, Charles C. 2004. "Turning the tables: How case-oriented research challenges." *Rethinking social inquiry: Diverse tools, shared standards* 123.
Ragin, Charles C. 2009. *Redesigning social inquiry: Fuzzy sets and beyond*: University of Chicago Press.
Ragin, Charles C, and Howard Saul Becker. 1992. *What is a case?: exploring the foundations of social inquiry*: Cambridge university press.
Rihoux, Benoît, and Charles C Ragin. 2008. *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*: Sage Publications.
Sager, Fritz, and Céline Andereggen. 2012. "Dealing with complex causality in realist synthesis: The promise of qualitative comparative analysis." *American journal of evaluation* 33 (1):60-78.
Sawyer, Michael G, Sara Pfeiffer, Susan H Spence, Lyndal Bond, Brian Graetz, Debra Kay, George Patton, and Jeanie Sheffield. 2010. "School-based prevention of depression: a randomised controlled study of the beyondblue schools research initiative." *Journal of Child Psychology and Psychiatry* 51 (2):199-209.
Short, Kate, Patricia Eadie, and Lynn Kemp. 2020. "Influential factor combinations leading to language outcomes following a home visiting intervention: A qualitative comparative analysis (QCA)." *International Journal of Language & Communication Disorders* 55 (6):936-954.
Smith, David John. 2006. *School experience and delinquency at ages 13 to 16*: University of Edinburgh, Centre for Law and Society Edinburgh.
Thomas, James, Alison O'Mara-Eves, and Ginny Brunton. 2014. "Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example." *Systematic reviews* 3 (1):67.
Warren, Emily, Leonardo Bevilacqua, Charles Opondo, Elizabeth Allen, Anne Mathiot, Grace West, Farah Jamal, Russell Viner, and Chris Bonell. 2019. "Action groups

as a participative strategy for leading whole-school health promotion: Results on implementation from the INCLUSIVE trial in English secondary schools." *British Educational Research Journal*.
Warren, Emily, GJ Meledez-Torres, Russell Viner, and Chris Bonell. 2020. "Using qualitative research to explore intervention mechanisms: findings from the trial of the Learning Together whole-school health intervention." *Trials* 21 (1):1-14.

Table 1: Overall mechanism, indicators, variables, and whether they were included in the original model

	Sub-mechanism 1		**Sub-mechanism 2**			**Sub-mechanism 3**	**Outcome**
	Improving commitment		**Improving social skills**			**De-escalating bullying**	**Decreased bullying victimization**
Indicator	Increasing overall student belonging	Creating a role for to participate in school	Improving pro-social skills	Learning conflict resolution	Decreasing conduct problem	Reducing perpetration	Decreased victimization
Abbreviation	belong	role	prosocial	rp solving	cp	aggression	decreased bullying
Data source (difference between 1-36 months)	BBSCQ belonging subscale	BBSQC student active participation at school subscale	Selected SDQ prosocial items (1,4,9,17,20)	Students who report teachers help resolve conflict	SDQ conduct problems subscale	ESYTC measure of bullying perpetration	GBS (School % difference 0-36 months)
Included in original model or added to improve coverage	Original	Original	Original	Added to improve coverage	Added to improve coverage	Original	Original

Table 2 for improved commitment: Indicators, data sources, and whether they were included in the original model

Context (C), Mechanism (M), or Outcome (O)	C	M	M	M	O
Indicator	Pre-existing ethos of wishing to involve students in decision-making	Good relationships between staff and students on AGM	Students feeling they made a positive contribution to the school via implementation of AG activities	AG participants initiating change in students' attitudes to school	AGM increased participation of students in school decisions
Abbreviation	decision-making	relationships	actions	attitude change	participation
Relevant intervention resource		Facilitator, annual student needs survey, preventative RP training	Facilitator, NAR	Facilitator, NAR	
Data sources	BBSCQ participation subscale at baseline	AGM survey (end of year 1) Score with a point for agreeing with the any of the following: “I got positive responses when I expressed my own attitudes and ideas on the Action Group”; “I found the Action Group to be exciting and energizing”; “This Action Group taught me how to work well together with others”; and “This Action Group helped me connect with other people in my school to help others”.	AGM surveys (end of year 1) “Do you think the AG made sure that these actions were implemented?" Percentage of students who answered “Yes” vs “No” and “Not sure”	AGM survey (year 1) "This Action Group helped me connect with other people in my school to help others" Percentage of students who answered “Yes” vs “No” and “Not sure”	BBSCQ (school % difference 0-36 months)

Included in original model or added to improve coverage	Original	Original	Original	Added to improve coverage	Original

Table 3 for improved social skills: Indicators, data sources, and whether they were included in the original model

Context (C), Mechanism (M), or Outcome (O)	C	C	M	M	O
Indicator	Students lack strong pro-social skills, or their development is seen as a priority by staff	Students feel unsafe in school	Delivering SEL skills curriculum with fidelity	Preventative RPs being used	Improved pro-social skills
Abbreviation	weak pro-social	feel unsafe	curriculum	preventative rp	improve prosocial
Relevant intervention resource			Curriculum	Preventative and responsive RP training	
Data source	School average baseline SDQ	Student survey question 55 at baseline: "Do you feel safe in school?" Percentage of students who responded "Never" vs "some of the time", "most of the time", and "all of the time"	Delivered 5+ hours or units of SEL curriculum in Y1 and Y2	Staff survey question Q32 at endline: "Teachers and students at this school get together to build better relationships" and question 33 "Teachers and students at this school get together to discuss their views and feelings" (Answers: "Often" vs "Sometimes" or "Never")	SDQ-pro-social subscale (school % difference 0-36 months)

Included in original model or added to improve coverage	Original	Added to improve coverage	Original	Original	Original

Table 4 for De-escalated bullying: Indicators, data sources, and whether they were included in the original model

Context (C), Mechanism (M), or Outcome (O)	C	M	M	M	M	O
Indicator	High baseline bullying victimization	Sufficient staff trained in RP	Implementing responsive RP	Perpetrators feeling empathy	Perpetrators accepting responsibility and accepting punishment	Decreased bullying
Abbreviation	bullying	rp training	responsive rp	empathy	contrition	decreased bullying
Relevant intervention resource		Responsive RP training	Responsive RP training	Responsive RP training	Responsive RP training	
Variable	GBS at baseline. Threshold: School score at baseline > median across all schools	At least 6 members of staff participated in 3-day training	From the endline staff survey: "If there is trouble at this school, staff respond by": and anyone who answers "Talking to those involved to help them get on better"	Student interviews	Student interviews	GBS (School % difference 0-36 months)

Included in original model or added to improve coverage	Original	Original	Original	Added to improve coverage	Added to improve coverage	Original

Table 5: Consistency and coverage scores for effective solutions

Mechanism	Consistency	Coverage
Overall mechanism: Effective solutions to reduce bullying victimization		
Combined solutions	***0.97437423***	***0.61665392***
belong*ROLE*prosocial*RPSOLVING*cp*AGGRESS	0.99540198	0.02784096
belong*ROLE*PROSOCIAL*RPSOLVING*CP*AGGRESS	1	0.05866053
BELONG*ROLE*PROSOCIAL*rpsolving*cp*aggress	0.99647295	0.0465431
belong*role*prosocial*rpsolving*CP	0.95749176	0.14792643
belong*role*PROSOCIAL*RPSOLVING*aggress	0.98276007	0.14313276
belong*ROLE*prosocial*CP*aggress	0.97100782	0.12016959
BELONG*PROSOCIAL*rpsolving*CP*AGGRESS	1	0.11216037
BELONG*ROLE*rpsolving*CP*AGGRESS	1	0.1324797
role*rpsolving*CP*aggress	0.95651352	0.20270701
Sub-mechanism 1: Effective solutions for improving commitment		
decision-making*ACTIONS*ATTITUDE CHANGE	0.84102321	0.32393599
Sub-mechanism 2: Effective solutions for improving pro-social skills		
Combined solutions	***0.76434785***	***0.54527509***
weak pro-social*curriculum*PREVENTATIVE RP	0.83939826	0.3778989
weak pro-social*FEEL UNSAFE*PREVENTATIVE RP	0.75935143	0.30734947

WEAK PRO-SOCIAL*FEEL UNSAFE*curriculum	0.86973155	0.25098297
Sub-mechanism 3: Effective solutions for de-escalating conflict		
Combined solutions	***0.90241832***	***0.59704226***
bullying*rp training*responsive rp*EMPATHY*CONTRITION	0.88330477	0.10710326
BULLYING*RP TRAINING*IMPLEMENT RESPONSIVE RP*EMPATHY*CONTRITION	0.85339141	0.10188263
bullying*RP TRAINING*empathy*contrition	0.90876937	0.3278009
bullying*RP TRAINING*responsive rp* empathy	0.93179297	0.22316746

Capital letters indicate the presence of a condition; lowercase letters indicate the absence of a condition. * = and.