



EVALUATION PROTOCOL

Grassroots pilot evaluation

IOE – UCL's Faculty of Education and Society

Principal investigator: Becky Taylor

IOE – UCL’s Faculty of Education and Society

Becky Taylor, Mark Hardman, Sal Riordan, Nicola Abbott, Keri Wong, Jeremy Hodgen & Victoria Showunmi

Project title	Grassroots pilot evaluation
Developer (Institution)	Behavioural Insights Team (BIT)
Evaluator (Institution)	IOE – UCL’s Faculty of Education and Society (IOE)
Principal investigator(s)	Becky Taylor
Evaluation plan author(s)	Becky Taylor, Mark Hardman
Evaluation setting	School-based
Target group	11-14 year olds in mainstream schools in England and Wales ¹
Number of participants	4 schools

¹ The pilot study will be conducted in England only.

Contents

Contents.....	3
Study plan version history.....	4
Evidence.....	5
Rationale.....	6
Research questions and/or objectives.....	8
Development phase.....	8
Pilot phase.....	9
Success criteria and/or targets.....	12
Methods.....	15
Methods and data collection.....	15
Methods overview.....	17
Data analysis.....	18
Outputs.....	18
Ethics and registration.....	19
Data protection.....	19
Personnel.....	20
Risks.....	20
Timeline.....	22
References.....	24
Appendix A: Theory of change.....	27

Study plan version history

Version	Date	Reason for revision
1.2 [latest]		
1.1	26/04/23	Revised following YEF review
1.0 [original]		<i>[leave blank for the original version]</i>

Any changes to the design need to be discussed with the YEF Evaluation Manager (EM) and the developer team prior to any change(s) being finalised. Describe in the table above any agreed changes made to the evaluation design, research questions and approach, and the rationale for these.

Intervention

Grassroots: empowering socially influential pupils to reduce bullying (Grassroots) is an anti-conflict programme that aims to reduce bullying and conflict in schools by empowering pupils to positively impact their fellow pupils' prosocial attitudes and behaviours. It has international evidence from a large-scale RCT in the USA but has not been delivered previously in England and Wales.

Research Assistants (RAs) complete three days of training delivered by the project team. The project team help administer a survey ('Grassroots survey') asking pupils to identify the other pupils with whom they have recently chosen to spend time. The project team conducts network analysis on the survey data to identify the best-connected pupils in each school (termed 'social referents'), who are prioritised for inclusion in the school's 'seed group' alongside less well-connected pupils chosen to be representative of the school demographic. RAs then convene each of these seed groups of approximately 30 pupils for 10 fortnightly sessions over the course of the rest of the school year.

In the sessions with 'seeds', RAs will help pupils to:

- a. Identify areas for improvement in student interactions
- b. Generate possible solutions (for example activities they may do, what they might encourage their friends to do)
- c. Provide opportunities for action (for example weekly or fortnightly challenges)
- d. Make initiatives visible to others (for example putting up posters and handing out wearable items, like wristbands, to peers for prosocial behaviour)
- e. Help pupils use online platforms to reach others (for example videos, social media content).

The intervention has not been delivered in England and Wales and will require adaptation to the England/Wales context. With this in mind the delivery partner (The Behavioural Insights Team (BIT)) proposes a 'developmental phase' in February-March 2023 to adapt the intervention, followed by a pilot in four schools in April-July 2023.

Evidence

In a large RCT in the USA (Paluck et al., 2016), Roots reduced disciplinary events related to conflict and bullying among adolescents by 25% over one year in treatment schools. This makes it one of the only evidenced anti-bullying interventions for adolescents, along with the more resource-intensive INCLUSIVE intervention (Bonell et al., 2015).

The US Roots trial also found that, on average, students in treatment schools reported higher levels of talking with friends about how to reduce conflict, and also wore anti-bullying wristbands more often than control schools.

Stronger effects on school-reported numbers of disciplinary incidents involving peer conflict were found in schools where seed groups had a higher proportion of 'social referents' – in the US study, the proportion of social referents in the seed group was varied randomly across schools (from 0 to 37%). The authors of the US study recommend that future interventions include as many social referents in their seed groups as possible. BIT intend to include as high a proportion of social referent pupils in each seed group as possible, while still ensuring that seed groups are representative of school populations by ethnicity. It is anticipated therefore that the seed groups will include approximately 75% social referents.

Two recent pilots of Roots in Indonesia (Bowes et al., 2019) found mixed results, but were deemed promising enough for the programme to be rolled out nationally. In South Sulawesi, mean bullying perpetration decreased by 29% and mean victimisation by 20%. However in Central Java, bullying increased slightly from the baseline due largely to situational factors (including increased reporting, among other factors). Both pilots yielded useful insights on effective implementation.

Rationale

A systematic review conducted by Kowalski and Limber in 2013 concluded that bullying is prospectively associated with school absenteeism, that bullying victimisation is prospectively associated with later educational achievement, and that bullying perpetration is prospectively associated with later aggression and/or violence. Adolescents are particularly influenced by peers. Indeed, the EEF's rapid evidence assessment, conducted in advance of this funding round, found that peer networks may be more powerful levers for this group than the risk of negative consequences from anti-bullying sanctions (Education Endowment Foundation, 2022).

The Roots intervention (Paluck et al., 2016) was implemented with typical adolescent school conflict in mind, including verbal and physical aggression, spreading rumours about peers, and social exclusion. Paluck et al. used a working definition of conflict as including 'harassment or antagonism from a high-power or high status person aimed at a person with lower power or status (i.e., bullying), but also conflict between or among people with relatively balanced levels of social power and status' (page 567). By encouraging highly-connected pupils to take a public stance against typical forms of conflict at school, behaviour change can be maximised across the pupil population. This approach has been successful in other domains, for example Campbell et al. (2008) recruited socially influential adolescents

as anti-smoking campaigners and found a reduction of 22% in regular smoking among 12–13-year-olds in England and Wales.

Research questions and/or objectives

Development phase

The developers are adapting a successful intervention from the US for use in UK schools. The purpose of the development phase evaluation is to document the process of adaptation and to update the theory of change logic model in preparation for the pilot evaluation.

The developers intend to review and revise the US Roots curriculum materials for use in English and Welsh schools. Alongside their own internal review meetings, the developers intend to work with two small, diverse groups of young people (young people advisory groups, YPAGs) during Spring 2023. The developers intend to recruit a panel of eight 11-14 year olds to participate in two 1.5 hour face-to-face intervention development workshops and an additional eight 16-19 year olds to participate in three 1.5 hour online workshops.

Members of the evaluation team will participate in the development process as ‘critical friends’ to the development team, acting as participant observers in development meetings and observing YPAG meetings. This will enable the developers and evaluators to work together with the YPAG in a coproductive relationship to support adaptation of the intervention.

Our research question for the development phase is summarised in Table 1.

Table 1. Research questions and data collection methods for the development phase.		
Focus	Research questions	Data collection methods
D1. Adaptation of the intervention to UK context.	What adaptations are needed to the Roots intervention to make it suitable for implementation in UK schools?	Participation in adaptation activities/workshops

During this phase we will also:

- Update the theory of change and logic model with the developers;

- Identify which measures might be needed in the evaluation and whether any adaptation or additional validation is required;
- Refine research questions for the pilot phase, especially in relation to any issues raised by the YPAG.

To achieve these goals Wong and Abbott will attend a sample of adaptation meetings between the developers and the YPAG. We will also hold a workshop with the developer team to achieve the other goals.

Pilot phase

The pilot phase is intended to assess the suitability of the intervention components, and the adaptations made through the development phase, for use in England and Wales. It is also a mean by which the indicators and processes of evaluation may be refined. This phase is therefore a pre-cursor to the efficacy trial rather than a full pilot evaluation. During the pilot phase we will focus on:

Intervention feasibility. We will consider what are the facilitators and barriers to the intervention that may affect implementation in schools. To support this we will examine school level factors, developer factors including the use of research assistants, the characteristics of the intervention and support for the intervention in schools. This will provide initial feedback on elements of the intervention and support definition of evaluation dimensions (such as fidelity and dosage) in readiness for the later efficacy trial.

Evaluation feasibility: data management and quality. We will pilot the measures we plan to use in the impact evaluation and IPE to check reliability, validity and practicality, including looking for ceiling/floor effects and to check the burden on participants and likelihood of missing data. Measures to be piloted would include: measures of bullying victimisation and perpetration (including the Olweus Bully/Victim Questionnaire – Revised), measures of disciplinary incidents in school; school-reported measures of pupil absence (to establish whether we can distinguish between COVID-related and other absence, and the feasibility of using school data). We will pilot survey and interview instruments with young people to ensure they understand the wording of questions and to see whether they would foresee any confusion or issues with the measures being used. We will also discuss with the developers to ensure that the elements of the evaluation do not interfere with the intervention.

Evidence of promise. We believe it will be essential to collect initial evidence around aspects of the logic model to assess promise, particularly as a number of aspects of the model have been highlighted amber/red risk by the developers (see Appendix A), indicating high impact on outcomes if there is no causal influence of the intervention. We will also check for unintended or negative effects of the intervention, particularly with attention to attitudes

towards pupils identified as social referent seeds and pupils from minority ethnic groups. We will ask questions via Teacher Tapp’s panel survey to check for evidence as to whether aspects of the intervention might already be commonly delivered in schools, so as to establish differentiation from usual practice and approaches to evaluating this further.

Readiness for trial. We would work with the developers to refine the theory of change and logic model and ensure that the intervention is well-defined in preparation for the efficacy trial.

The impact evaluation design has already been decided, with randomisation unlikely to be problematic, and as recruitment will be ongoing during the development and pilot phases, it is not necessary to pilot the impact evaluation design or procedures. The short duration of the pilot also makes piloting retention strategies and assessing potential attrition rates challenging. We will however discuss with pilot schools the potential issues around retention that they foresee and work with them to suggest strategies to keep a high level of retention.

This evaluation presents challenges in working simultaneously with English and Welsh pupil data which we will continue to investigate during the pilot phase. For example, there is no data sharing agreement between the Department for Education (England) and Welsh Government, meaning that Welsh and English data may need to be analysed separately and subsequently combined. Welsh data may not be available within the same timeframe as English data. Additionally, we will need to verify that variables in the English National Pupil Database (NPD) and Welsh SAIL database are directly comparable. We will investigate during the pilot phase whether it may be more feasible to collect attendance data from schools.

We will work with the young person advisory group (YPAG, as described in the development phase, above) and our race equity advisor to ensure that the evaluation is racially and culturally sensitive.

We will produce a slide deck near the end of the pilot phase to assist YEF with the decision about whether the intervention is ready for efficacy trial, and a short report on the development and pilot phases in late summer 2023. The research questions for the pilot evaluation are summarised in table 2.

Table 2. Research questions and data collection methods for the pilot evaluation.
--

Focus of pilot evaluation	Research questions	Data collection methods
P1. Intervention feasibility	<p>How manageable is each aspect of the Grassroots intervention for schools?</p> <p>What are the most important facilitators and barriers to successful implementation?</p> <p>What are the likely moderators and mediators of impact?</p> <p>Are there any additional facilitators or barriers to implementation in Welsh schools?</p>	<p>Observations of training sessions and Grassroots survey processes</p> <p>Interviews with teachers</p> <p>Focus groups with seed pupils</p> <p>Survey of teachers</p> <p>Telephone interviews with school leaders in Welsh schools.</p>
P2. Evaluation feasibility: data management and quality	<p>What are the most reliable, valid and practical measures for bullying victimisation/perpetration, disciplinary incidents in schools and for pupil absence that can be used in the impact evaluation phase?</p> <p>What are barriers and facilitators are there to data collection in schools?</p> <p>Are counts of behavioural incidents and school attendance suitable measures to conduct further evaluations of the intervention?</p>	<p>Surveys of pupils</p> <p>Interviews with teachers</p>
P3. Evidence of promise	<p>Are there any early indicators of promise?</p> <p>Which, if any, aspects of the intervention are being delivered in control schools as part of their business-as-usual practices?</p>	<p>Survey of pupils</p> <p>Interviews with teachers</p> <p>Teacher Tapp survey</p>

P4. Readiness for trial	<p>Is the Grassroots intervention ready for trial in English and Welsh schools?</p> <p>Does the logic model accurately capture the processes and outcomes of the intervention?</p> <p>How can the intervention be specified ready for efficacy evaluation?</p>	<p>Survey of pupils</p> <p>Workshop with developers</p>
P5. Race equity	<p>What racial equity or diversity challenges might be faced in the delivery of the project and the evaluation? How can these be addressed?</p> <p>Are activities and materials or surveys accessible, inclusive and culturally and racially sensitive?</p> <p>How does the intervention address racist or racialised bullying?</p> <p>To what extent is race-cognizance evident in the intervention?</p>	<p>Discussion with developers, YPAG and race equity advisor</p>

Success criteria and/or targets

Success criteria for the development phase:

1. The Grassroots curriculum is ready for the pilot phase. Materials and processes are appropriately adapted for English schools.
2. The theory of change logic model for the intervention is updated for pilot phase.

Success criteria for the pilot phase:

Dimension	Detail	RAG stop/go criteria
P1. Intervention feasibility	1. The Grassroots intervention (network survey, network	RED: The network survey and network analysis process are not ready.

	analysis and Grassroots curriculum) is adapted ready for trial in English and Welsh schools, including translation into Welsh for Welsh-medium schools.	AMBER: The network survey and network analysis process are ready. <50% of the Grassroots curriculum is ready in English and Welsh. GREEN: All elements of the Grassroots intervention are ready.
	2. School leaders in English and Welsh schools report that the intervention can be feasibly implemented in schools like theirs.	RED: One or more schools report critical concerns that cannot be mitigated. AMBER: One or more schools report critical concerns with mitigations in place but not tested. One or more schools reports non-critical concerns. GREEN: No schools report critical or non-critical concerns, or mitigations are in place and have been successfully tested for any critical concerns reported.
P2. Evaluation feasibility: data management and quality.	1. A suitable measure of behavioural incidents (from analysis of disciplinary report data), which is practical to create is identified.	RED: N/A
		AMBER: A measure of behavioural incidents cannot be feasibly produced. Disciplinary report data would be omitted from the efficacy trial, but the trial could proceed without this outcome measure.
		GREEN: A suitable measure is developed. Any limitations relating to the measure's reliability and validity are documented.
	2. A suitable measure of attendance is identified for use as the primary outcome measure at the short-term and long-term follow-up.	RED: A suitable measure cannot be identified.
		AMBER: A measure is identified. Any limitations relating to the measure are identified and documented, including relating to equivalence between England and Wales. GREEN: A suitable measure is identified and known to be equivalent for England and Wales.
	3. Data collection is deemed feasible in schools.	RED: Barriers to the collection of primary outcome data are identified that cannot be mitigated.
AMBER: Barriers to the collection of secondary outcome data are identified that cannot be mitigated.		

		GREEN: Any barriers to collection of primary or secondary outcome data are identified, documented and mitigated.
P3. Evidence of promise.	1. Business-as-usual (BAU) practices likely to be observed in control schools are identified.	RED: >75% of schools report BAU practices that are very similar to the intervention.
		AMBER: >50% of schools report BAU practices that are very similar to the intervention.
		GREEN: >25% of schools report BAU practices that are very similar to the intervention.
P4. Readiness for trial.	1. The Grassroots intervention is ready for trial in English and Welsh schools.	Please see P1.1, P1.2.
	2. The primary and secondary outcome measures (attendance and disciplinary events) are operationalised and confirmed as appropriate for the impact evaluation.	Please see P2.1, P2.2, P2.3.
	3. The intervention is fully specified, including criteria for compliance. Evaluation dimensions (such as fidelity and dosage) are specified using evidence from the pilot.	RED: Criteria for compliance, fidelity and dosage have not been identified.
		AMBER: Criteria for compliance, fidelity and dosage have been identified.
		GREEN: Robust criteria for compliance, fidelity and dosage have been identified.
P5. Race equity.	1. Racial equity and diversity challenges for the intervention delivery and evaluation are identified.	RED: The intervention is deemed not to meet racial equity standards by the YPAG or race equity advisor. Mitigations or adaptations are not possible.
		AMBER: Concerns are raised by the YPAG or race equity advisor about race equity. Mitigations are put in place but adaptations are not possible.
		GREEN: No concerns about race equity are raised by the YPAG or race equity advisor in relation to the intervention, or suitable adaptations are put in place to address concerns.

	2. Evaluation activities, materials and data collection instruments are judged by YPAG and by our race equity advisor to be accessible, inclusive and culturally and racially sensitive.	RED: The evaluation is deemed not to meet racial equity standards by the YPAG or race equity advisor. Mitigations or adaptations are not possible.
		AMBER: Concerns are raised by the YPAG or race equity advisor about race equity. Mitigations are put in place but adaptations are not possible.
		GREEN: No concerns about race equity are raised by the YPAG or race equity advisor in relation to the evaluation, or suitable adaptations are put in place to address concerns.

a. Methods

Methods and data collection

The developers have committed to recruiting four pilot schools to participate in the Grassroots pilot. These schools will be selected through the developer’s networks, for ethnic diversity and for proximity to the developer’s London base. The sample size was determined based on cost-effectiveness given that the main focus of the pilot is on feasibility and qualitative rather than quantitative evidence of promise.

All pupils in years 7-9 in the pilot schools will be invited to participate in the pilot, as this is the target age group for the intervention. BIT intend to pilot the involvement of a small number of year 10 pupils as special advisors in one or two schools, but year 10s will not complete the Grassroots survey and year 10 pupil data will not be collected in the initial data upload. We will aim to include year 10 special advisors in the pilot evaluation where possible.

We will conduct a mixed-methods pilot evaluation with a survey of pupils and light touch case studies in four English schools experiencing elements of the intervention. The developers are planning to pilot the network analysis in two schools and aspects of the programme delivery in two further schools. They have indicated that the materials will not have been translated into Welsh in time for the pilot phase and they will not by that point have recruited a Welsh speaking research assistant. During the pilot phase we will conduct telephone interviews with school leaders in two Welsh secondary schools to establish whether there are any particular issues for feasibility.

The developers have already produced a theory of change and logic model during the proposal phase (see Appendix A). We will revisit this at the end of both the development and pilot phases to finalise the model for the next stage. This will take place in meetings involving

the developers and evaluation teams. Additionally, the theory of change and logic model will be presented to the YPAG and their feedback sought. This will include attention to race equity questions.

Data collection instruments will include interview and focus group schedules, observation pro formas and surveys.

Interview and focus group schedules will be developed in anticipation of the evaluation dimensions for the following efficacy trial and tested for clarity and focus during the pilot. We will also include a short section of further items to explore the research questions of the pilot, whilst ensuring the interviews and focus groups are not too arduous for participants in the pilot. Observation protocols will developed inductively through each observation, with an initial focus on dimensions of evaluation as explored within interviews.

We will pilot surveys with young people, including to select a suitable measure of bullying perpetration and victimisation. We will administer the Olweus Bully/Victim Questionnaire (Revised Version) in all four pilot schools (Kyriakides, Kaloyirou and Lindsay, 2006). This will establish the feasibility of conducting this survey and how it can be administered to reduce the burden on schools.

During the pilot phase, BIT intend to work with four schools. BIT intend to pilot all components of the intervention (i.e. Grassroots survey, network analysis and seed group allocation, and seed group meetings following the adapted Grassroots curriculum), but due to the limited time available, not all schools will receive all elements of the intervention.

The Grassroots survey and network analysis will be piloted in two schools, followed by seed group identification and piloting of some seed group activities over a shorter period of time. In the other two schools, seed groups will be nominated by school staff and BIT will aim to pilot more seed group activities over a longer period of time. Due to constraints of time BIT will only be able to pilot a sample of the Grassroots curriculum.

Methods used during the pilot evaluation will include observations, interviews with teachers and pupils, surveys of teachers and pupils, scrutiny of the Grassroots survey processes and network analysis, YPAG discussions and workshops with the developer team.

Methods overview

Research methods	Data collection methods	Participants/ data sources (type, number)	Data analysis methods	Research questions addressed
	Observations of Grassroots activities	Research assistants and seed group pupils in four pilot schools	Inductive analysis initially focused on the anticipated dimensions of the efficacy trial.	P1
	Administrative data for Grassroots activities	BIT	Quantitative analysis	P1, P4
	Teacher interviews	One teacher in each pilot school	Inductive analysis. Dual focus on pilot RQs and efficacy trial dimensions.	P1, P2, P3
	Pupil focus groups	One focus group of 4-6 seed group pupils in each pilot school At least one focus group of year 10 special advisors, where appointed.	Inductive analysis. Dual focus on pilot RQs and efficacy trial dimensions.	P1
	Teacher surveys	Teachers and leaders in each pilot school	Descriptive statistics and thematic analysis.	P1
	Teacher Tapp survey	Teacher Tapp panel	Descriptive statistics.	P3
	Teacher interviews	School leaders in two Welsh schools	Inductive analysis.	P1

	Pupil surveys	All KS3 pupils in four pilot schools (approximately n=800)	Descriptive statistics and evaluation of standard instrument	P2, P3, P4
	Workshop	Workshop with developer and evaluator teams		P4, P5
	Focus groups with YPAG	All members of the YPAG		P5
	School data collection	Attendance and disciplinary data for KS3 pupils in four pilot schools	Pilot statistical analysis to identify challenges for impact evaluation.	P2, P4

Data analysis

We will work with all four pilot schools to investigate the availability and format of disciplinary and attendance data that schools are willing to share. We will use this data to trial and further develop the statistical models for the impact evaluation. This process will identify issues regarding non-comparability of data, missing data, and linking pupil data and pupil characteristics. We will also explore the feasibility of NESTA using machine learning to code the disciplinary report data.

Qualitative data will be analysed using thematic analysis (Braun & Clarke, 2006), using themes derived from the research questions.

Outputs

There will be four outputs from the pilot evaluation:

1. Presentation to YEF on findings from the development phase.
2. Written report to YEF on findings from the development and pilot phases.
3. Updated logic model.
4. Efficacy trial study plan.

Ethics and registration

The evaluation will be conducted in line with the BERA (2018) Guidelines on Research Ethics and will be approved by the IOE Research Ethics Committee. A number of important ethical issues are likely to be encountered during this research and we will plan for mitigation as follows:

Bullying is a sensitive issue and therefore young people must be assured of confidentiality, anonymity and protection from harm. We will ensure that individuals and schools cannot be identified in any reporting and ensure that data collection methods are appropriate and proportionate and unlikely to cause harm or distress. We will take advice from the YPAG to support this. Case studies will include focus groups, surveys and observations with young people and this raises the risk of disclosure of harm. We will inform young people that in case of making a disclosure we cannot keep confidentiality and will follow school safeguarding procedures.

Schools are busy and time-poor places and the research should not pose an excessive burden on teachers, administrators or young people. We will use the pilot phase of the study to establish whether we can sample young people within a school rather than surveying the whole cohort.

All participants should consent to taking part in the study. Headteachers will be invited to sign their school up to the study through returning a completed Memorandum of Understanding and Data Sharing Agreement. In addition, we will write to all target young people and their parents/carers to inform them about the study and offer them the opportunity to opt out of the research. We will not include data from any young person for whom consent has been withdrawn. We will ensure that young people and teachers in case study schools provide active opt-in consent before collecting any data.

Data protection

Data will be processed in line with data protection legislation (including GDPR) and in the interests of the participants. The project will be registered with the UCL Data Protection Officer. Personal data will be lawfully processed using GDPR Article 6(1)(e) Public Interest. Special category personal data (ethnicity) will be lawfully processed using Article 9(2)(j) under condition 4 (Research). We will provide an opportunity for parents/carers to discuss the research with their child and to withdraw their data from the research and any data processing. All results will be anonymised so that no schools or individual pupils or teachers will be identified in any report arising from the research. Further information about how UCL uses participant information can be found here: <https://www.ucl.ac.uk/legal-services/privacy/ucl-general-research-participant-privacy-notice>

We will establish data sharing agreements between schools and UCL (legal body for IOE) and between UCL and BIT. Each will act as data controllers for their own separate purposes and schools will sign the DSAs at the stage of joining the evaluation through submitting the MOU.

Personnel

Dr Becky Taylor is a Principal Research Fellow at UCL and would lead the evaluation, manage the evaluation team and lead on communications with YEF and the developer team. She would lead on writing of the study plan and the interim (pilot) and final reports.

Dr Mark Hardman is an Associate Professor at UCL and would be a co-investigator for the evaluation, playing a leading role in the qualitative aspects of the IPE and contributing to all aspects of the evaluation including reporting.

Dr Keri Wong is an Associate Professor at UCL and would be a co-investigator for the evaluation; she would play a leading role in working with the developers on the theory of change and logic model, developing outcome measures and contribute to all aspects of the evaluation including reporting.

Dr Nicola Abbott is a Lecturer in Psychology at UCL and would be a co-investigator for the evaluation; she would play a leading role in working with the developers on the theory of change and logic model, working with the YPAG, and contribute to all aspects of the evaluation including reporting.

Dr Sally Riordan is a Senior Research Fellow at UCL and would be a co-investigator for the evaluation; she would lead on writing the statistical analysis plan and conducting the impact analysis and would contribute to reporting.

Professor Jeremy Hodgen is Professor of Mathematics Education at UCL and would provide expert statistical support for designing the impact evaluation, including modelling and outcome variables.

Dr Victoria Showunmi is an Associate Professor at UCL, specialising in gender and race in educational contexts. She will act as race equity advisor to the evaluation team and provide critical feedback on all materials and methods.

Claire Pillinger is a Research Assistant at UCL and would support all areas of data collection and management for the evaluation and conduct initial cleaning and coding of qualitative data for the IPE.

An **administrator** would support all areas of the project, including liaising with schools to secure a high level of retention, oversee the transfer of data from schools to UCL, arrange fieldwork visits and survey administration plus any other necessary tasks to support the project.

Risks

Risk	Likelihood	Impact	Action
Failure to recruit	Low	High	<ul style="list-style-type: none">Establish timeline for recruitment involving a variety of methodsRegular developer and evaluator team contact

Failure to gain data from schools	Low	High	<ul style="list-style-type: none"> • Clear and simple data collection procedures (including explanations of GDPR conditions) • Include data sharing agreement in MOU • Ensure correct data (including UPNs and matching variables) as a pre-condition for randomisation • Allocate staff time to school liaison at key data collection points
Attrition of schools	Low	Moderate / High	<ul style="list-style-type: none"> • Regular contact with participating schools • Regular developer and evaluator team contact • Pilot schools will receive a financial incentive
Loss of staff	Low / Moderate	Low	<ul style="list-style-type: none"> • IOE has a large staff team and would reallocate staff
Poor response rate to surveys	Low	Moderate	<ul style="list-style-type: none"> • Monitor through regular contact with schools. • Regular developer and evaluator team contact • Clear and simple collection methods

Timeline

Dates	Activity	Staff responsible/ leading
27/3/23	First 11-14 yr-old YPAG session	BIT
29/3/23	Second 11-14 yr-old YPAG session	BIT
30/3/23	First 16-19 yr-old (online) YPAG session	BIT
27-30/3/23	Observation of YPAG session	IOE
17-21/4/23	Information sheets and privacy notices distributed to parents of all KS3 pupils by pilot schools (and potentially year 10 too, at least in the case of the two schools that are not completing the Grassroots survey)	IOE
1-5/5/23	<p>Pilot schools share data with IOE.</p> <p>The two schools that will complete the Grassroots survey must share at least pupil first name, last name, year group, tutor group, sex and ethnicity.</p> <p>Ideally all four schools would also share a historical sample of the year 7-9s disciplinary report data at this point in the pilot.</p>	IOE
8-12/5/23	<p>IOE share first name, last name, year group, tutor group, sex and ethnicity with BIT (and chase the two schools, if necessary, to ensure this is possible within this window)</p> <p>Ideally the disciplinary data would be shared with BIT at this point too, to give time for Nesta to explore the data, but a little more time may be required for anonymisation (date TBC)</p>	IOE

15-19/5/23	Project team prepares name lists for the network surveys	BIT
22-26/5/23	Grassroots surveys administered in two of the pilot schools (no time for slippage due to May half term)	BIT
22/5/23- 2/6/23	Project team cleans data, completes social network analysis and selects seed group pupils in these two schools	BIT
	Review social network analysis	IOE
5-9/6/23	The two pilot schools that completed the Grassroots survey share seed group pupil invitations and parental consent forms	BIT
12-16/6/23	These schools chase responses (it may be necessary to give one more week)	BIT/schools
19/6- 25/7/23	The two pilot schools that completed the Grassroots survey host 1-2 seed group sessions	BIT
1-5/5/23	The two pilot schools that are not completing the network survey randomly select seed pupils (and, hopefully, a small number of Year 10 special advisors)	BIT/schools
1/5-25/7/23	Piloting of survey instruments Focus groups and interviews	IOE
7-21/5/23	Teacher Tapp survey to check business-as-usual practices	IOE
15-26/5/23	The two pilot schools that are not completing the network survey share seed group pupil invitations and parental consent forms, and then chase responses	BIT

5-9/6/23	Training of facilitator for pilot	BIT
12/6-25/7/23	The two pilot schools that are not completing the network survey host seed group sessions	BIT
7/7/23	YPAG review pilot findings	IOE
	Presentation of pilot evaluation findings to YEF	IOE
	Submission of draft pilot evaluation report to YEF	IOE
	Submission of final pilot evaluation report to YEF	IOE

References

- Bonell, C., Fletcher, A., Fitzgerald-Yau, N., Hale, D., Allen, E., Elbourne, D., Jones, R., Bond, L., Wiggins, M., Miners, A., Legood, R., Scott, S., Christie, D., & Viner, R. (2015). Initiating change locally in bullying and aggression through the school environment (INCLUSIVE): Pilot randomised controlled trial. *Health Technology Assessment, 19*(53). <https://doi.org/10.3310/hta19530>
- Bowes, L., Aryani, F., Ohan, F., Haryanti, R. H., Winarna, S., Arsianto, Y., Budiawati, H., Widowati, E., Saraswati, R., Kristianto, Y., Suryani, Y. E., Ulum, D. F., & Minnick, E. (2019). The development and pilot testing of an adolescent bullying intervention in Indonesia – the ROOTS Indonesia program. *Global Health Action, 12*(1), 1656905. <https://doi.org/10.1080/16549716.2019.1656905>

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

Education Endowment Foundation. (2022). *Attendance interventions: Rapid evidence assessment*. EEF.

<https://d2tic4wvo1iusb.cloudfront.net/documents/pages/Attendance-REA-report.pdf?v=1679344158>

Paluck, E. L., Shepherd, H., & Aronow, P. M. (2016). Changing climates of conflict: A social network experiment in 56 schools. *Proceedings of the National Academy of Sciences*, 113(3), 566–571. <https://doi.org/10.1073/pnas.1514483113>

Appendix A: Theory of change

THEORY OF CHANGE

The Theory of Change (developed by the delivery team with facilitation from EEF/YEF ToC workshops) can be found here: [Roots anti-bullying intervention, Online Whiteboard for Visual Collaboration \(miro.com\)](#)

Assumed causal mechanisms (developed by the delivery team with facilitation from EEF/YEF ToC workshops)

#	Assumption/causal mechanism	Where in the ToC does the assumption apply	Evidence	Assumption risk	Evidence strength
---	-----------------------------	--	----------	-----------------	-------------------

1	<p>Pupils attend to the behaviour of certain other pupils in the school to understand what is socially normative and adjust their behaviour accordingly (and these 'social referent' pupils can be identified through network analyses based on surveys of who pupils have chosen to spend time with).</p>	Output	Short-term outcome	<p>The US RCT (Paluck et al., 2016) found that the seed group-led anti-conflict interventions produced a large (25%) reduction in peer conflict (measured using school disciplinary reports) and that the effect size varied with the proportion of 'social referents' in the seed group: the reduction in peer conflict was closer to 60% for schools with the highest proportion of social referents in their seed groups. It should be noted that the specifics of the mechanism by which the actions of the seed group lead to a reduction in peer conflict is not fully understood. What <i>is</i> known is that (a) the intervention did not (in the US RCT) make pupils (on average) perceive the student body in general as more anti-conflict or as engaging less in conflict behaviour and (b) while being directly connected to a social referent seed did seem to make a pupil more likely to perceive the student body in general as more anti-conflict, this did not look to be related to the reduction in peer conflict behaviour. Our current hypothesis is that pupils perceive the seed group (rather than all pupils) to be anti-conflict and that if the pupils care about this group's opinion (which is more likely to be the case if it includes a relatively large proportion of highly-connected pupils) then they will change their behaviour accordingly</p>	<ol style="list-style-type: none"> 1. There is a risk that the seed group pupils don't buy in 2. Relatedly, there is a risk that the anti-conflict messaging from the seed group is not convincing to other pupils. 3. There is a risk that the actions of the seed group are not sufficiently visible 4. There is a risk that the change in conflict behaviour is mediated via something other than perceptions of the seed group's attitudes to conflict (but that the reduction in conflict behaviour is still achieved via another mechanism). One possibility is that the social referent pupils are generally <i>perpetrators</i> of conflict and that the intervention changes their own conflict related behaviour. We feel this is unlikely. And it could be investigated using social network analysis data and school disciplinary records in intervention and control schools) 	Amber/ Green
2	<p>The reduction in peer conflict leads to reduced exclusions because perpetration of student conflict reduces, resulting in fewer disciplinary sanctions (and ultimately exclusion).</p>	Short term outcome	Long-term outcome	<p>As above, the intervention has been found to reduce school disciplinary records of peer conflict in the US. It is not yet known whether the intervention reduces exclusions, as this was not specifically measured in the RCT, but this seems plausible, as 40% of exclusions in England are due to some form of conflict.</p>	<ol style="list-style-type: none"> 1. There could be changes to school leadership that result in changes in disciplinary approach. This ought to be equivalent across arms but if not then we might observe changes in exclusions that aren't the result of the intervention). 2. As (baseline) school exclusions are 	Amber/Red

					relatively rare, it is possible that the evaluator will not be able to detect the impact of the intervention on exclusions (we	
--	--	--	--	--	--	--

					could potentially aim to recruit schools with relatively high exclusion rates).	
3	<p>Peer conflict is detrimental to pupils' feelings of safety and belonging in school, and to their mental health, particularly for pupils on the receiving end of the conflict, and this worsens their school attendance and engagement. Therefore, we can expect a reduction in peer conflict to lead to improvements in attendance and engagement (via improvements* in pupils' mental wellbeing and feelings about school) *We do not anticipate this</p>	Short term outcome	Long-term outcome	<p>This has not been tested directly, to our knowledge, but there is suggestive evidence. Being bullied is associated with increased absenteeism (Kowalski & Limber, 2013). Victims of bullying often experience decreased interest in academics and may skip school to avoid being bullied (Slee, 1994). Bullying can also cause health problems that may lead students to miss school (Ramya and Kulkarni, 2011). Pupils who are repeatedly bullied display elevated symptoms of depression and anxiety (Fonagy et al., 2005), increasing the risk of school dropout (Esch et al., 2014).</p>	<p>1. As (baseline) attendance in British schools is relatively high, it is possible that the evaluator will not be able to detect an effect of the intervention on attendance (we could potentially aim to recruit schools with relatively low attendance).</p> <p>2. If British school children's attendance was not being adversely affected by peer conflict at baseline, then reducing peer conflict will not improve attendance.</p>	Amber/Red

	being measured in this project					
--	--------------------------------	--	--	--	--	--

Notes: Evidence Strength: Green - the evidence base is very strong, there are peer reviewed academic studies, meta-analyses or independent experimental evaluations directly linked to the assumption. **Green/Amber** - the evidence base is strong, there are academic studies or independent evaluations linked to the assumption. **Red/Amber** - the evidence base is developing, there are academic studies, internal evaluations or recorded observational evidence that are adjacent to the assumption. **Red** - the evidence base is limited.

Figure 2. Contextual assumptions (developed by the delivery team with facilitation from EEF ToC workshops)

	Contextual Assumption	Assumption Strength	Assumption Risk
1	Student conflict is present in KS3 in British schools (we are relatively confident about this, as almost half of young people report being bullied at age 14 (DfE, 2010) and 1 in 4 adolescents report bullying occurring 2-3 times a month (Przybylski & Bowes, 2017). Furthermore, we could look to preferentially recruit schools that are likely to have relatively high conflict rates.)	Green	Red

2	We are able to recruit sufficient schools (we will use our networks and our protocols developed over many trials to try to ensure this is not the case. However, we may still struggle to recruit the kinds of schools we would like to (e.g. those with a high exclusion rate) - we could try to enlist the support of influential figures e.g. Tom Bennett)	Amber/Green	Red
3	The schools accept the terms of participation and do not adapt the intervention in ways that might undermine its effectiveness (we can minimise this risk by giving clear, simple instructions to schools and facilitators)	Green	Amber/Green
4	A focus group of young people at British schools, from diverse backgrounds, are able to support us to adapt the Roots curriculum to a UK context during the development / recruitment phase (we will use our existing contacts with schools to maximise the likelihood of this)	Green	Amber/Green
5	Pupils complete the survey that allows us to do network analyses and identify pupils for the seed groups (we will draw on the expertise and experience of our academic collaborator, who has run the intervention before, and we will look into acceptability and feasibility in the pilot stage)	Green	Red
6	Schools share the survey data (we will ensure that instructions given to schools are as clear, simple and actionable as possible and that information sheets and withdrawal forms for parents are clear and accessible)	Green	Red
7	We are able to recruit appropriate facilitators (we will draw on the expertise and experience of our academic collaborator, who has run this intervention before, and the experience that we develop during the EEF Stop and Think project (which also involves recruiting temporary members of staff to visit schools to give training to the people who will administer the intervention).	Green	Amber/Red
8	The facilitators are able to attend the training (this will be part of our facilitator recruitment criteria)	Green	Amber/Red

9	The facilitators are able to reach the recruited schools to deliver all seed group sessions (geographical spread will be a factor considered during recruitment, as will conscientiousness)	Green	Red
10	The majority of invited seed group pupils accept the invitation to participate (and their parents provide consent for them to do so). In the US RCT in which the intervention was found to have a high impact (in an intention-to-treat evaluation), 24% of seed group pupils did not accept the invitation. The risk of low rates of invitation acceptance/parental consent can be mitigated through clear information for seed group pupils and their parents. We will ensure that British school pupils and parents feed into the development of this text during the development phase	Green	Amber/Red
11	The majority of seed group pupils exhibit pro-social behaviour (1. we consider this unlikely given the outcomes of the US RCT, 2. good facilitation can reduce the risk that clustering together pupils with anti-social behaviour increases anti-social behaviour, 3. If we take action that increases this risk (e.g. if we decide that a specific proportion of the seed groups will be from the funding round's target group of interest), then we can consider asking schools to vet proposed seed group lists)	Green	Red
12	The Roots intervention sessions can be arranged, and are honoured by schools (advanced preparation should help ensure this is the case)	Green	Red
13	There is sufficient attendance (by seed group pupils) at the intervention development sessions (1. in the US RCT in which the intervention was found to have a high impact (in an intention-to-treat evaluation), attendance was 55%; 2. we will use the development and recruitment phase to work out times of day that optimise attendance, 3. we will seek to make the sessions appealing to attend e.g. through making the sessions fun and the provision of snacks/a free lunch)	Amber/Green	Amber/Red
14	The seed group sessions are facilitated effectively, and seed group pupils are willing to actively participate and work together to design their interventions (1. we will draw on the expertise and experience of our academic collaborator, who has previously implemented the	Amber/Green	Amber/Red

	intervention, 2. We will select for facilitation skills during facilitator recruitment)		
15	Seed group pupils are willing and able to deliver their anti-conflict interventions, with the endorsement of school staff (1. good facilitation will help ensure that the seed group pupils feel entirely happy with what they are doing and speaking in their own voice, 2. we will make clear the expectations of school staff at the point of recruitment and foster good relationships with school staff)	Green	Red
16	The seed group pupils' anti-conflict interventions are not met with conflict by the student body (1. this did not appear to be an issue in the US RCT, 2. the involvement of the social referent pupils likely helps; 3. good facilitation, ensuring seed group pupils feel entirely comfortable with what they are doing, and able to speak in their own voice, increases the likelihood of this - see previous point)	Green	Amber/Red
17	Covid doesn't causes schools to close again - there is currently no indication that this will happen and the vaccination programmes help mitigate the risk	Amber/Green	Red
18	The British school disciplinary report data and attendance data is fit for the purpose of this trial (we will speak with British schools during the development phase to ensure this is the case, in particular that school disciplinary reports are collected, and can be shared, at sufficient granularity for the evaluation of this intervention)	Amber/Green	Amber/Red

Notes: Assumption Strength: Green – This assumption will hold in the vast majority of circumstances where the programme is delivered. **Green/Amber** – This assumption will hold in most of the circumstances where the programme is delivered. **Red/Amber** – This assumption will often not hold in the circumstances where the programme is delivered. **Red** – There is a good chance of this assumption not holding / do not know whether this assumption will hold or not. **Assumption Risk: Green** – The programme could continue to be delivered with very minor impact. **Green/Amber** – The programme could continue to be delivered, but the impact would be substantial. **Red/Amber** – The programme could continue to be delivered, but without fidelity to original design. **Red** – The programme could not be delivered



youthendowmentfund.org.uk



hello@youthendowmentfund.org.uk



[@YouthEndowFund](https://twitter.com/YouthEndowFund)

The Youth Endowment Fund Charitable Trust

Registered Charity Number: 1185413
