Education Endowment Foundation

PROJECT TITLE	Supporting Teachers And childRen in Schools (STARS): Incredible Years® Teacher Classroom Management
DEVELOPER (INSTITUTION)	University of Exeter
EVALUATOR (INSTITUTION)	National Foundation for Educational Research
PRINCIPAL INVESTIGATOR(S)	Dr. Ben Styles
PROTOCOL AUTHOR(S)	Gemma Stone, Maria Galvis, Connie Rennie
TRIAL DESIGN	Two arm randomised controlled trial. Random allocation at school level
PUPIL AGE RANGE AND KEY STAGE	Ages 5-6, Key Stage 1
NUMBER OF SCHOOLS	140
NUMBER OF PUPILS	5,880 (based on a mean cohort size of 42 pupils per school)
PRIMARY OUTCOME	Maths attainment: Maths Key Stage 1 raw scores
SECONDARY OUTCOMES	- Pupil emotional and social well-being - Pupil concentration - Pupil prosocial behaviour - Pupil classroom behaviour - Teacher-pupil relationship

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.1	24 May 2019	- updated list of recruitment areas - updates to randomisation strategy - updates to CLASS observer training
1.0	09 November 2018	N/A

Table of Contents

Intervention

TIDieR items

Brief name

Supporting Teachers and Children in Schools (STARS): Incredible Years® Teacher Classroom Management (IY-TCM)

Why

Rational/theory/goal of the intervention

The Incredible Years programme is a series of parent, teacher and child training packages that focus on developing positive parent-teacher-child relationships as a way of building resilience in young people for their social and emotional wellbeing, potentially positively impacting on their future academic achievement. The STARS: Incredible Years Teacher Classroom Management (IY-TCM) programme is a training programme for classroom teachers on techniques and behaviours that support and encourage these positive relationships, with a specific focus on managing challenging behaviour.

Evidence suggests that between 10% and 20% of children and adolescents worldwide suffer with some form of psychological disorder which may result in challenging behaviour, such as attention-deficit hyperactivity disorder or autism spectrum disorder (Ford *et al*, 2018[1]). The goal of the IY-TCM intervention is to support teachers in managing challenging behaviour, giving them the tools to improve pupils' behaviour in their classroom by adopting specific behaviours and techniques.

A recent systematic review that evaluated eight interventions aiming to improve children's social and emotional wellbeing through training teachers found that IY-TCM is one of the school-based programmes supported by more substantial evidence (Whear *et al.*, 2013)[2]. Some studies have found early evidence of the impact of IY-TCM on pupil social-emotional development. Ford *et al.* (2018)[3] report a small short-term improvement to children's mental health, particularly for those who are already struggling. Nye (2017)[4] also shows an overall effect on pupil mental health, with a significant effect on children with behaviour problems.

This intervention has been trialled recently (Ford et al, 2018[5]), in order to assess, amongst other elements, impact on pupil mental health and wellbeing. This independent evaluation is

---

[1] Ford, T., Hayes, R., Byford, S., Edwards, V., Fletcher, M., Logan, S., & Ganguli, P. (2018). The effectiveness and cost-effectiveness of the Incredible Years® Teacher Classroom Management programme in primary school children: results of the STARS cluster randomised controlled trial. Psychological medicine, 1-15.

[2] Whear, R., Thompson-Coon, J., Boddy, K., Ford, T., Racey, D., & Stein, K. (2013). The effect of teacher-led interventions on social and emotional behaviour in primary school children: a systematic review. British Educational Research Journal, 39(2), 383-420.

[3] ibid.

[4] Nye, E. (2017). Classroom behaviour management to support children's social, emotional, and behavioural development (Doctoral dissertation)

[5] ibid.

necessary in order to both scale up the research, and to assess impact on pupils academic attainment.

Who

Recipients of the intervention

The training (see details below) will be delivered to all teachers of Year 1 and Year 2 pupils within the target regions in mainstream primary schools i.e. all schools other than independent/selective, special, alternative provision and schools in special measures.

The target regions are Reading, Cornwall, Bristol, Southampton, Dorset and Liverpool. The training will be provided to a total of 70 schools, expecting an average participation of three teachers per school (based on 1.5 average form entry). To be eligible to receive the training, teachers must have at least four days of teaching per week. In cases where teachers are on a job share, both teachers sharing the job will be trained and their combined teaching time must be at least four days per week.

In order for one year group to receive the intervention for two consecutive years, training will be delivered to both Year 1 and Year 2 teachers, so that when the Year 1 pupils move up to Year 2, they are taught by teachers who have been trained and have previously practised the intervention (on their class during the first year of the trial). As a by-product, pupils in Year 2 at the beginning of the trial will receive the intervention as their teachers will be receiving the training. Though these pupils are not part of the trial, both Year 1 and Year 2 teachers are part of the trial. Only pupils who start the intervention in Year 1 and continue on to Year 2 will have their data analysed for the impact evaluation. For the sake of clarity, in this document we will refer to the groups as Cohorts 1 and 2.

	2019 – 20	2020 – 21
Cohort 1 (study group)	Year 1 (teacher receiving training)	Year 2 (with trained teacher)
Cohort 2	Year 2 (teacher receiving training)	Year 3

What

Materials

Exeter will recruit 'group leaders' to deliver the training to teachers. These group leaders will attend a consecutive three-day training course, where they will be provided with a training manual that includes suggested scripts, videos, a day-by-day programme, posters, stickers and worksheets. These manuals will be used by group leaders to deliver the IY-TCM training to the study teachers.

Teachers will receive six one-day training sessions, delivered by the group leaders once a month for six months. Group sizes will be determined by the number of teachers being trained per school but are targeted at an ideal number of four schools per session i.e. twelve teachers. At commencement they will be provided with the intervention handbook, which contains the whole course. At each training session, each teacher will receive handouts for

the session and to take away, as well as 'buzz' documents (summaries of key learning points).

Procedures

IY-TCM's explicit goals are to enhance teacher classroom management skills and improve teacher-pupil relationships, assist teachers to develop effective proactive behaviour plans (plans for teachers to use to deal proactively with behaviour), encourage teachers to adopt and promote emotional regulation skills, and encourage teachers to strengthen positive teacher-parent relationships. The programme supports teachers in this through cognitive and emotional self-regulation training, and through goal-setting, reflective learning, video-modelling and role play, at the training sessions and in tasks to be undertaken between training sessions.

Teachers are encouraged to practise strategies between sessions and discuss their experiences with each other at training, aiming to build teacher-pupil relationships through social and emotional coaching, praise and incentives. The programme is a manualised series of tools and strategies, allowing for flexible, contextually sensitive implementation.

Manualised curriculum components:

- Building positive relationships with students and being a proactive teacher
- Teacher attention, coaching, encouragement and praise
- Motivating students through incentives
- Decreasing inappropriate behaviour (ignoring and redirecting, follow through with consequences)
- Emotional regulation, social skills and problem solving

Intervention delivery techniques:

- Experimental learning (observe, discuss, rehearse, reflect)
- Timetabled sessions to ensure adequate time to develop and embed new practices
- Peer support
- Collaborative learning
- Expert facilitation and support

Who

Providers/Implementers

Group leaders (trainers) will deliver the intervention, as described above. An experienced group leader will run half of the IY-TCM groups alone. The other half will be run both by an experienced and a less experienced group leader. The IPE will monitor comparability through observations (see the IPE section for more detail).

Teaching Assistants are not directly trained as part of the intervention, but teachers are encouraged to share information and training documents with them.

How

Modes of delivery

The University of Exeter will recruit schools to the trial in the target areas. The group leader will train groups of minimum eight and maximum 15 teachers; the minimum is set to retain a group dynamic. Each group should include a mix of teachers from at least four different schools from the geographical hub.

Group leaders and their group will stay together throughout the training period (other than emergencies). It is important that teachers do not miss sessions or send a replacement. Nevertheless, according to a previous IY-TCM evaluation (Ford *et al.*, 2018[6]), attending at least four out of the six sessions is likely to be sufficient.

The intervention itself is delivered to pupils through and during normal teaching; there is no change to curriculum or content of teaching. The techniques and behaviours are not dependent on subject and will be encouraged/supported throughout the full school day.

Where

Location and setting

The intervention will take place in Reading, Cornwall, Bristol, Southampton, Dorset and Liverpool. Training will be delivered at a regional location, away from the participating schools. The school settings are described above ('Who; receipients of the intervention').

When and how much

Duration and dosage of the intervention

The training consists of six one-day sessions spread across October 2019 – March 2020. The intervention itself should be integrated into the usual teaching practice, meaning that the dosage to pupils is continuous (daily). At the monthly training sessions, teachers will indicate to what extent and how frequently they use the IY-TCM strategies using a template designed by Exeter and NFER collaboratively. The expectation is that teachers use the strategies and behaviours every day, but this is dependent on the level of challenging behaviour in the class. For classes with little challenging behaviour, strategies may be different and used less frequently to those classes with more challenging behaviour.

Tailoring

Adaptation of the intervention

The IY-TCM programme is highly manualised with clear criteria for training, supervision and fidelity. Nevertheless, it allows for "adaptation with fidelity", meaning that group leaders can select from a range of techniques to deliver the prescribed curriculum according to what is most acceptable to their context. For example, a teacher may find that setting up incentives for individuals such as stickers or prizes is not appropriate for their school or class context, and choose instead to incentivise with special privileges such as extra computer time or being the teacher's 'special helper'.

---

[6] ibid.

How well

Strategies to maximise adherence and fidelity

The developer team has previously delivered the intervention successfully in their own trial of the intervention (Ford et al, 2018[7]). This trial was administered and school engagement was high, as such the developers have a good track record in engaging schools and delivering the intervention to a high standard. There are no changes to the intervention in this trial, only to the outcomes and the size of the sample. Training sessions, materials and manuals provided as part of the intervention are very structured and the intervention is highly manualised, which supports delivery with fidelity at all levels.

Some techniques to ensure success in the implementation include selecting highly experienced group leaders. Some of these will be individuals who have previously delivered the IY-TCM training, and some will be new to the project. Those new to the project will be carefully selected based on training and classroom experience; trained to the programme and standardised in order that they are able to deliver to the same standard as the existing trainers.

The relationship with schools is a top priority for the developer to ensure successful delivery. This will take the format of established regular contact as soon as a school is allocated to the intervention group and throughout the first year of the programme, with named and easily contactable individuals to respond to queries and comments from participants. Schools will be allocated into groups or 'hubs' and have an allocated trainer; this establishes a small network of support between schools and the trainer which supports positive relationships.The developer will ensure that schools feel fully confident in delivering the intervention and provide any additional support and advice throughout the training and thereafter.

---

[7] ibid.

Theory of Change

Assumptions

Evidence suggests that some types of negative or coercive interaction between children and adults reinforces disruptive behaviour (Patterson, 1982[i]). In previous trials, the STARS intervention has successfully impacted upon social and emotional development (Ford et al, 2018) in children.

STARS supports teachers in learning and applying behavioural management techniques and building strong pupil teacher relationships. It draws on the importance of modelling and self-efficacy (Bandura, 1977[ii]); and developmental interactive learning methods (Piaget and Inhelder, 1962[iii]), and incorporates cognitive behavioural approaches and Bowlby's attachment theory on the importance of positive relationships (Bowlby, 1951[iv]). The hypothesis is that this will lead to more productive time in class and a more trusting learning environment; and thus to higher attainment.

For the intervention to achieve success, teachers must attend a minimum of two thirds of the training (as evidenced in Ford et al, 2018)

Strategies and activities

What is the approach?

This is a highly manualised programme. Teachers in the intervention group receive six day-long sessions of CPD, delivered by group leaders, once per month for six months.

Teachers then use the techniques learnt in training during usual teaching in the classroom. The CPD trains teachers in specific techniques to build teacher-pupil relationships, and effectively manage and decrease inappropriate behaviour. These techniques are practised in the classroom and reflected upon at training events.

Resources include:
- Intervention handbook
- Handouts
- 'Buzz' sheets

Short-term outcomes (1-2 years)

Pupil impacts:

Primary outcome
Maths attainment

Secondary outcomes
Pupil emotional and social wellbeing
Pupil concentration
Pupil prosocial behaviour
Pupil classroom behaviour
Teacher-pupil relationship

School and teacher level impacts:

Reduced Stress

Improved school behaviour policies

Longer-term outcomes (2-5 years) (not evaluated in this trial)

Improved long term attainment

Improved health and self-esteem

Improved staff retention

Target groups

Schools: Mainstream schools, not in special measures.
Regions: Reading, Cornwall, Bristol, Southampton, Liverpool, Dorset
Pupils: Year 1 (Cohort 1)

Study rationale and background

IY-TCM's explicit goals are to: enhance teacher classroom management skills and improve teacher–student relationships; assist teachers to develop effective proactive behaviour plans; encourage teachers to adopt and promote emotional regulation skills; and encourage teachers to strengthen positive teacher–parent relationships. The theory is that pupils in classes taught by IY-TCM-trained teachers will have the capacity for higher attainment as their engagement with the teaching content is higher, less time is spent on behaviour management, and they feel more positive and confident about actively participating in learning. A recent high-quality meta-analysis of teacher classroom management approaches (Korpershoek, 2016)[8] found that classroom management interventions have a significant effect on various student outcome measures, including attainment. Programmes such as IY-TCM, which incorporate social and emotional learning (SEL), were found to have the biggest effects, particularly on SEL outcomes but also on attainment. This programme is perhaps the best established, having been trialled in several countries, as well as being previously trialled in England recently by the developers for this study.

IY-TCM is a popular programme internationally that has been implemented in England, Wales, Scotland, Ireland, Jamaica, New Zealand, Norway and Portugal as well as its country of origin, the USA. It has a 3+ rating from the Early Intervention Foundation, indicating strong evidence of short-term effects on behavioural outcomes. There are three RCTs that feed into this high rating, from Wales, Ireland and the USA. The Welsh study (Hutchings *et al.*, 2013[9]) involved 107 children aged 3-7 years old from 12 classes and found significant reductions in off-task behaviour and 'child negatives' (negative behaviours) to teacher and vice versa, and an increase in child compliance (in terms of classroom behaviour), all measured through independent coded observation one year after baseline. Hickey *et al.* (2017)[10] recruited 445 children from 11 schools in Ireland and found significantly positive outcomes in children's emotional self-regulation as measured through the teacher-reported Strengths and Difficulties Questionnaire (SDQ). Reinke *et al.* (2016)[11] randomised 1,817 students aged 5-8 in the USA to intervention or waitlist conditions, and found significant improvements in prosocial behaviour, emotional regulation, and social competence.

The delivery team have also recently conducted a RCT with 2,075 children in 80 schools randomised into two arms (control and intervention) in the South West of England. They trained one teacher per school in 3 cohorts between 2012 and 2015, and found significant

[8] Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. Review of Educational Research, 86(3), 643-680.

[9] Hutchings J, Martin-Forbes P, Daley D, Williams ME. A randomized controlled trial of the impact of a teacher classroom management program on the classroom behavior of children with and without behavior problems. J School Psychol. 2013;51(5):571–85

[10] Hickey, G, McGilloway, S, Hyland, L, Leckey, Y., Kelly,P., Bywater, T., Comiskey, C., Lodge, A., Donnelly, M., and O'Neill, D.. "Exploring the effects of a universal classroom management training programme on teacher and child behaviour: A group randomised controlled trial and cost analysis." Journal of Early Childhood Research 15, no. 2 (2017): 174-194.

[11] Reinke, W. M., Herman, K. C., & Dong, N. (2016). The Incredible Years Teacher Classroom Management Program: Outcomes from a group randomized trial. Prevention Science, 1-12.

effects on mental health and behaviour outcomes measured by the SDQ after 9 months, though these did not persist at 18 and 30 months (Ford *et al*, 2018[12]).

The focus of this evaluation is the potential impact on mathematics attainment. There is convincing evidence that behavioural regulation has a robust relationship with numeracy. For example, Ponitz *et al.*, (2009[13]) found that behavioural regulation predicted gains in mathematics, but not language and literacy, for six-year-old children. Numeracy skills have previously been shown to be increased after successful self-regulation interventions in 6th graders (Perels *et al.,* 2009[14]). In addition, assessing gains in maths skills makes the research more inclusive of SEN/EAL children as language and literacy ability will not interfere with pupil ablity to demonstrate knowledge and understanding against the construct. Maths skills and knowledge in early school years (end of reception year) potentially predict maths performance at end of Key Stage 1 as well as end of Key Stage 2, so there is a significant long-term reason for improving numeracy skills early on (Aubrey *et al.*, 2006[15]) This study also sets out to demonstrate impact on social and emotional wellbeing, pupil concentration, prosocial behaviour, classroom behaviour and teacher-pupil relationships as secondary outcomes, and as such will represent the largest-scale study of IY-TCM against these dimensions.

Impact Evaluation

Research questions

Primary question

RQ1: What is the impact of the IY-TCM intervention on pupils KS1 Maths attainment (collected directly from schools), compared to 'business as usual'?

Secondary questions

RQ2a: What is the impact of the IY-TCM intervention on pupils' emotional and social wellbeing as measured by the Total Difficulties Score of the SDQ, compared to 'business as usual'?

RQ2b: What is the impact of the IY-TCM intervention on pupils' concentration, as measured by the hyperactivity/innatention sub-scale of the SDQ, compared to 'business as usual'?

RQ2c: What is the impact of the IY-TCM intervention on pupils' prosocial behaviour, as measured by the prosocial sub-scale of the the SQD, compared to 'business as usual'?

---

[12] ibid.

[13] Ponitz, C. C., McClelland, M. M., Matthews, J. S. and Morrison, F. J. (2009). 'A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes', Developmental Psychology, 45, 3, 605-619.

[14] Perels, F., Dignath, C.and Schmitz, B. (2009). 'Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes', European Journal of Psychology of Education, 24, 1, 17

[15] Aubrey, C., Godfrey, R. and Dahl, S. (2006). 'Early mathematics development and later achievement: Further evidence', Mathematics Education Research Journal, 18, 1, 27-46.

RQ3: What is the impact of the IY-TCM intervention on pupils' classroom behaviour, as measured by the Pupil Behaviour Questionnaire, compared to 'business as usual'?

RQ4: What is the impact of the IY-TCM intervention on the teacher-pupil relationships, as measured by a revised version of the Student Teacher Relationship Scale (STRS) compared to 'business as usual'?

RQ5: Are effects on KS1 Maths attainment (as per RQ1) different for pupils eligible for FSM?

RQ6: Is the potential impact on KS1 Maths attainment (as per RQ1) different for 'struggling' pupils whose Total Difficulties score of the SDQ is grater than or equal to 12 (a score which represents above the 80th percentile for the British school-age population), compared to 'non-struggling' pupils?

RQ7: Is the potential impact on KS1 Maths attainment (as per RQ1) different for pupils with different Total Difficulties score of the SDQ? This research question will be assessed using a similar model to that used for RQ5, however the SDQ responses will be included as a continuous variable, not a categorical variable.

Design

Table 2. Design of the intervention

Trial type and number of arms		Two arm cluster randomised controlled trial
Unit of randomisation		School
Stratification variables (if applicable)		Geographical area
Primary outcome	variable	Maths attainment
	measure (instrument, scale)	Baseline: Early Years Foundation Stage Profile (EYFSP) End-point: KS1 Mathematics raw scores
Secondary outcome(s)	variable(s)	Pupil emotional and social wellbeing Pupil concentration Pupil prosocial behaviour Pupil classroom behaviour Teacher-pupil relationship
	measure(s) (instrument, scale)	Strengths and Difficulties Questionnaire (SDQ) Pupil Behaviour Questionnaire (PBQ) Revised version of the Student Teacher Relationship Scale (STRS)

Randomisation

Randomisation will be carried out by a statistician at NFER using a full SPSS syntax audit trail. Randomisation will be conducted at school level, with region and school size used as stratifiers. Six target regions will be used in order to facilitate training administration: Reading, Cornwall, Bristol, Southampton, Dorset and Liverpool. Furthermore, a proxy of

school size will be used: the number of reception teachers per school. The schools will be randomised 1:1 to intervention/business as usual. The analysts will not be blind to the randomisation result, however a full paper trail of all pre-specified analysis will be archived after the trial as per EEF guidance. Baseline measurements will be recorded prior to the school being informed of their group allocation, and follow up measures will be recorded by blind test administrators.

Participants

- Eligible schools are all mainstream primary schools within the target regions: Reading, Cornwall, Bristol, Southampton, Dorset and Liverpool. This is all schools other than independent/selective, special, alternative provision and schools in special measures.
- Eligible teachers are all Year 1 and Year 2 teachers who have at least four days of classroom responsibility per week, or both partners of a job share.
- The intervention is delivered to whole classes of Year 1 and Year 2 in selected schools, hence no pupil selection is necessary.

Exeter will recruit schools to the trial. Eligible schools are mainstream primary schools i.e. all schools other than independent/selective, special, alternative provision and schools in special measures.

Recruitment documents will consist of the following:

- Memorandum of Understanding (prepared jointly)
- Privacy Notice for parents (prepared by NFER)
- Privacy Notice for schools/teachers (prepared by NFER)
- School information letter (prepared jointly)
- Parent information leaflet (prepared by Exeter)
- Parental letter to opt out of data sharing (prepared jointly)

Exeter will begin contacting schools in January 2019. When a school indicates their interest to Exeter, they will send them the School information letter, the Memorandum of Understanding and the School privacy notice. When the school sends a signed Memorandum of Understanding back to Exeter, Exeter will record them as recruited to the trial. At three time points during recruitment, Exeter will notify NFER of recruited schools and pass MOUs to NFER via a secure portal. The batches will be sent in early March, mid April and late May.

On receipt of a batch of recruited schools, NFER will contact the school and send them, via a secure school portal, the Parent information leaflet, Parent withdrawal from data processing letter, and the Parent privacy notice. The school will be advised to print them out for all pupils who will participate in the trial and ensure they are taken home. If a parent withdraws their child from data processing at any time, the school will be required to inform NFER, who will then delete the pupil's data. NFER will also send the data template, requesting that the school return all pupil data within two weeks. Once data are received, NFER will use it to prepopulate baseline (SDQ/PBQ) test papers, which are then sent to schools via the secure online portal. (For further details about these tests, see the outcome measures section below). Schools are given a three week window to complete them before

sending them back, via the secure online portal. Once the tests are received from the school, schools are put forward to randomisation. Randomisation will take place in early July, in order to notify schools of group allocation in mid-July, before schools break up for the summer.

Sample size calculations

Sample size calculations were run in a bespoke NFER-designed Excel spreadsheet. A sample size of 140 schools with 42 pupils each is sufficient to detect an effect size of 0.17. This minimum detectable effect size (MDES) is achieved at 0.8 power by using the following assumptions: intra-cluster correlation of 0.15; correlation between Key Stage 1 and EYFSP of 0.51, and average cohort size of 42 pupils per school. The assumptions were drawn from the values of a previous study using KS1 arithmetic scores. The result of the power analysis is illustrated in Figure 1.
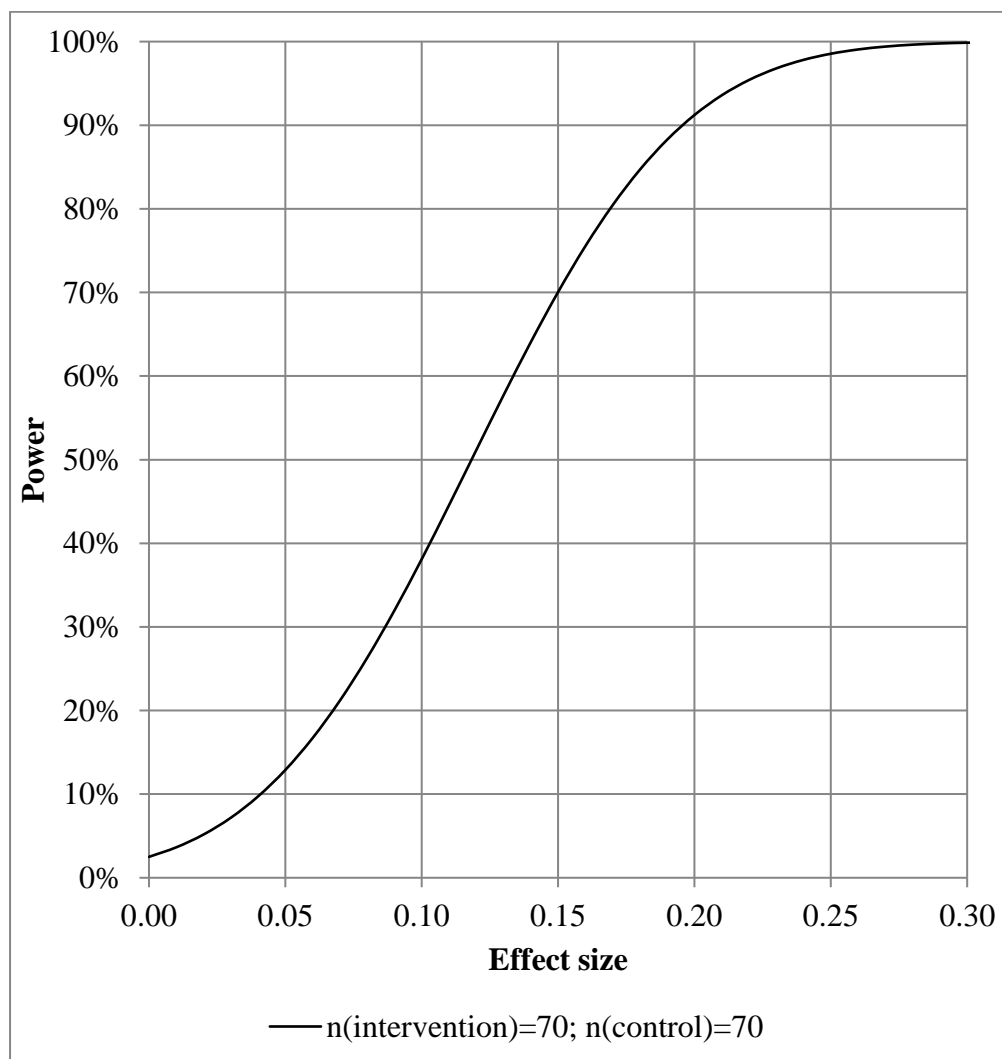
Figure 1. Sample power analysis

Table 3. Sample size calculations

		OVERALL	FSM*
MDES		0.17	0.20
Pre-test/ post-test correlations	level 1 (pupil)	0.51	0.51
	level 2 (class)	-	-
	level 3 (school)	-	-
Intracluster correlations (ICCs)	level 2 (class)	-	-
	level 2 (school)	0.15	0.15
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two	Two
Average cluster size		42	10.5
Number of schools	Intervention	70	70
	Control	70	70
	Total	140	140
Number of pupils	Intervention	2,940	735
	Control	2,940	735
	Total	5,880	1470

*assuming 25% FSM.

Outcome measures

Primary outcome

The primary outcome measure will be pupils' mathematics ability as measured by the *mathematics KS1 raw scores* -composite of both maths papers- collected directly from schools. Collecting raw data directly from the schools will provide continuous variables for the analysis, while eliminating the potential bias involved in recoding the results into categories by the teachers.

Secondary outcomes

Secondary outcome 1 will be *pupil emotional and social wellbeing* as measured by the Total Difficulties Score (which is the sum of the emotional, behaviour, peer relationship problems and attention/concentration subscales) of the Strengths and Difficulties Questionnaire (SDQ). This will be completed by the teacher about each individual pupil. The SDQ is a mental health instrument developed as a clinical measure, not focused on low-level disruption, but used widely to identify children for targeted interventions and differential impact in whole-class interventions. The SDQ will be administered online at baseline, midpoint and end-point for Cohort 1.

Secondary outcome 2 will be *pupil concentration* as measured by the hyperactivity/inattention subscale of the SDQ.

Secondary outcome 3 will be *pupil prosocial behaviour* as measured by the prosocial behaviour subscale of the SDQ.

Secondary outcome 4 will be *pupil classroom behaviour* as measured by the Pupil Behaviour Questionnaire (PBQ), which will be completed by the teacher about each individual pupil. The PBQ is better suited for identifying low-level disruption, but cannot be used to identify subgroups for analysis, which the SDQ can. The PBQ will be administered online at baseline, midpoint and end-point for Cohort 1.

Secondary outcome 5 will be *teacher-pupil relationship* as measured by a revised version of the Student Teacher Relationship Scale (STRS). In December 2018 NFER's Code of Practice group advised that NFER would not be prepared to use the STRS in its current form due to language it felt was inappropriate for a UK classroom context. As such, EEF, Exeter and NFER agreed to draft a revised version which could be tested by Exeter during the first year of the trial (2019). The testing would not be adequate to fully validate the test, aiming for approximately 200 teachers to feed back on the test, but would assess compatibility and user acceptance and allow confidence (or lack of) in the new scale to be identified. If the revised version of the scale shows good user acceptance, the protocol will be updated to include the scale in an Appendix with links where appropriate to a description of the testing process.

Analysis

Analysis will be pre-specified in a Statistical Analysis Plan published within three months of randomisation.

Analysis of the primary outcome will be intention-to-treat analysis of KS1 maths raw scores (a weighted average of the two papers) regressed on intervention group. We will use multi-level models containing two levels (school and pupil) to account for the cluster randomisation. The model will include EYFSP as a covariate. Regions will also be included as covariates (dummy variables) as these were strata in the randomisation. The analysis will be aimed at testing for an overall IY-TCM effect over the Business as Usual (BaU) control group.

Analysis of the secondary outcomes 1 through 4 will be intention-to-treat analysis of the respective outcome, regressed on the intervention group. We will use multi-level models containing three levels (school, pupil and timepoint) to account for cluster randomisation and measure the differential effect of time. Regions will be included as covariates (dummy variables). These models will be run to test for an overall IY-TCM effect over the BaU control group.

Analysis of secondary outcome 5 will be intention-to-treat analysis of teacher-pupil relationships regressed on the intervention group. A multi-level model containing two levels (school and pupil) will be used to account for cluster randomisation. The model will include the baseline measurement of the STRS scale as a covariate as well as the dummy region variables. This model will test for an overall IY-TCM effect over the BaU control group.

Complier Average Causal Effect analysis will be carried out on the primary outcome only. The fidelity measure for the CACE model will be a binary measure of teacher attendance at training; both teachers (Year 1 and Year 2) will need to each attend 4 out of the 6 sessions.

As FSM-eligible pupils represent a particularly important subgroup, a separate analysis of FSM-eligible pupils will also be carried out as per standard EEF practice if at least 1470 FSM eligible pupils are recruited, ensuring enough power to detect an MDES of 0.2 (see Table 3). This model will mimic the model used to asses the primary outcome, however with FSM pupils only.

Lastly, as per RQ 5, 6 and 7, three models with interaction terms will be run. These models will mimic the model used to assess the primary outcome, with the addition of the respective interaction terms. 1) FSM eligibility will be interacted with the intervention. 2) 'Struggler/non-struggler' (categorical), as described in the research question section) will be interacted with the intervention. 3) TDS (continuous) will be interacted with the intervention. These models will be used to asses any potential differential effect of the IY-TCM over BaU between pupils with different levels of disadvantage, and different levels of social and emotional well-being.

Implementation and process evaluation

The implementation and process evaluation (IPE) for this trial will cover the relevant dimensions set out in the EEF IPE introductory handbook: fidelity/adherence, dosage, quality, reach, responsiveness, programme differentiation, monitoring of comparison groups, and adaptation (Humprey *et al.*, 2016)[16]. The following research questions capture these dimensions along with particular areas of interest:

RQ1: To what extent is fidelity to the intervention maintained?

RQ2: How much of the intended intervention has been delivered? *Teacher attendance to the training sessions, time period that the pupil was taught by a trained teacher*

RQ3: To what extent did teachers and pupils engage with the intervention? *Classroom observations, training observations*

RQ4: What does 'business as usual' consist of for the comparison group?

RQ5: What level and type of support does the developer team provide to intervention schools?

RQ6: Does the intervention have a perceived impact on teacher self-efficacy?

RQ7: Does the intervention have a perceived impact on teacher stress?

IDEA workshop

The IDEA workshop was held on 9th October 2018, and provided an opportunity for the University of Exeter development team and the NFER evaluation team to explore the intervention in depth, and to develop an effective implementation and process evaluation (IPE) plan. Attendees first developed the Template for Intervention Description and Replication (TIDieR) framework, discussing key features of the intervention; then explored the Theory of Change, reflecting on the developer team's original ToC and amending to ensure a mutual understanding of the research process and its aims. Attendees then discussed fidelity/compliance criteria, including identifying the critical components of implementation and how to identify 'success' within each. This included monitoring requirements for both Exeter and NFER.

Tool 1: Practitioner Surveys – RQ3, RQ4, RQ6, RQ7

Baseline Survey

All Year 1 and Year 2 teachers in both control and intervention schools will complete a baseline survey in September 2019. This survey will be delivered online to teachers, via an email link. It will cover the following areas:

---

[16] Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). Implementation and process evaluation (IPE) for interventions in educational settings: An introductory handbook. *Education Endowment Foundation*, *1*.

- The teacher's confidence in behaviour management, and any previous CPD they have had on the topic
- The school's existing policy and practice for behaviour management, any strategies in place for challenging behaviour or individuals
- Teacher self-efficacy. This will consist of some items (not the full scale) from the Teacher Self Efficacy Scale (Bandura, 1997[17])
- Work-related stress. This will consist of some items (not the full scale) from the Teacher Concerns Inventory (Kyriacou, 2001[18])

Endpoint Survey

The same sample of teachers as captured in the baseline survey will complete an online endpoint survey in June 2021. This survey will cover the following areas:

- Teacher confidence in behaviour management
- Any changes to the school's approach/policies for behaviour management
- Any additional CPD that the school/teacher has received during the two year trial
- Teacher self-efficacy
- Work-related stress

In addition, intervention group teachers will answer questions relating to the following:

- Engagement and responsiveness to the intervention; particularly take up of the techniques during and after the training
- Costs of the intervention

To account for the intensive data collection burden for teachers, the University of Exeter and EEF will offer incentives to both intervention and control schools. Each school will receive £400 for each study teacher who is taking part.

Tool 2: Training Observation and Monitoring – RQ1, RQ2, RQ3, RQ5

A sample of training days will be observed in order to gather information about the trainer's fidelity to programme, the quality of the training provision and the intervention characteristic. The sample will ensure that each trainer is observed at least once, to ensure we can observe any variance in quality of delivery, content and engagement of participants across the trainers.

In addition, Exeter will provide us with the attendance registers of teachers at each training session, in order to measure the dosage of training delivery to teachers. It is pre-specified that compliance to the intervention will be indicated by attendance at four out of the six sessions.

In each of the six training sessions, teachers will also be asked to indicate, using a proforma created jointly by Exeter and NFER, the frequency of which they have used each of the

---

[17] Bandura, A. (1977a). Self efficacy: Toward a Unifying Theory of Behavioral Change. Psychological Review, 84, 2, pp 191–215.
[18] Kyriacou, C (2001) Teacher stress: Directions for future research, Educational Review, 53,1, pp 27-35.

intervention's strategies in the classroom since the last training session and how helpful they found it. This will provide a reported measure of both dosage and responsiveness during delivery of the intervention, which will be added to data from the survey on these dimensions.

Tool 3: Case studies – RQ1, RQ2, RQ3, RQ5, RQ6, RQ7

In addition to collecting information through the surveys and training observation, we will conduct a series of qualitative case studies and telephone interviews in order to explore the following aspects of the implementation factors set out in the IPE guidance in greater depth, particularly:

- Preplanning and foundations: what is the level of need, readiness and capacity for integrating the intervention's approaches to behaviour management in trial schools?
- Implementation support system: what training and support is available from the University of Exeter delivery team, and how is it perceived?
- Implementation environment: what is the context in the schools, e.g. how does it fit/differ from the school's usual behaviour strategy/practice including any senior leader support; any barriers to delivery?
- Implementer factors: who delivers the intervention – is there any deviance from what is stated in the Memorandum of Understanding? What is teachers' level of qualification and years of experience?
- Intervention characteristics that can affect implementation – the key criteria agreed for fidelity regarding dosage, weekly pattern, permissible tailoring, who implements, and any aspects of the programme or materials recently revised.

We will undertake qualitative work in a sample of 10 intervention schools, which will be selected randomly based on achieving a broad geographical spread. Four of these schools will be case-study schools, and six will be invited to take part in lighter-touch telephone interviews.

We will conduct case studies on a sample of four schools. These case studies aim to collect rich, descriptive data on classroom climate, teacher efficacy and relationships between teachers and pupils. In order to facilitate this and also bringing the study in line with previous evaluations of Incredible Years – Teacher Classroom Management (for example Murray *et al*, 2017[19]), we will be conducting a series of classroom observations at case study schools, and using CLASS – the Classroom Assessment Scoring System (Pianta *et al*, 2008[20]) to code the observations in order to gain a measure of any change over the two years of the evaluation. The CLASS tool produces a set of scores against 10 dimensions reflective of the class climate, teacher efficacy and teacher-pupil relations. Each dimension is scored out of 7 marks. These are totalled to give an overall score. Observations will be conducted at baseline, midpoint and endpoint, on different combinations of teacher-pupil, as described below. This will enable us to see any change over the time points, as measured by the CLASS tool.

---

[19] Murray, D, Rabiner, D, Kuhn, L, Pan, Y, Sabet, R (2018) Investigating teacher and student effects of the Incredible Years Classroom management program in early elementary school. *Journal of School Psychology*, 67, pp 119-133

[20] Pianta, R. C., La Paro, K. M. & Hamre, B. K. (2008). The Classroom Assessment Scoring System, Baltimore: Brookes. Manual, pre-K

The observer will take part in two days of CLASS K-3 training in order to develop an understanding of the tool, followed by 10 hours of extra reading and practice (as specified by the trainers). However, they will not be required to be certified or pass the base or annual associated reliability tests, and as such the outcomes of the observations must be treated as unstandardized; outcomes will, though, reflect on change during the intervention.

Developers suggest (Pianta, Hamre and Mintz, 2012[21]) that coding is most reliable when the classroom is observed four times within the same day, each time for approximately 20 minutes, which is immediately followed by ten minutes of coding, and this will be adhered to in the design. The frequency and timing of the data collection has been designed in order to track change between two subgroups (Table 4):

- *Class*: the study group and their Year 1 teacher (observed at two timepoints). This will enable us to track change in the relationship between the trained teacher and one group of children, observing difference in CLASS scores at the beginning and end of the active intervention.
- *Teacher*: the Year 2 teacher, first with their group at the beginning of year 1 of the evaluation, and second with the study group at the end of the evaluation. This will enable us to track change in CLASS scores whilst working with two different groups of children, and to follow the target pupils into their second year with a different, but still trained, teacher. It will also enable us to observe what impact/change in CLASS scores remains nearly a full year after the training ends.

Table 4. Timing of CLASS observations

Timepoint		Observations where the class stay the same	Observations where the teacher stays the same
Baseline	Sep 2019	Year 1 teacher and the study group	Year 2 teacher and their class at the time
Midpoint	June 2020	Year 1 teacher and the study group	
End point	February 2021		Year 2 teacher and the study group

Interviews

All teachers who have been observed as part of the case studies in the four selected schools will be interviewed to explore implementation factors in greater depth. Interviews with teachers will address similar elements/dimensions to the practitioner survey,. We will also explore their views of the quality of training received, the usefulness and success of the intervention techniques and behaviours adopted as part of the intervention, ways in which they are using the techniques, any changes they have made from the recommended approach and why, their views of the impact of the intervention on pupils, and cost.

---

[21] Pianta, R.C., Hamre, B.K. & Mintz, S.L. (2012). The CLASS-secondary manual. Charlottesville, VA: University of Virginia

The same interviews will be conducted via telephone with Year 1 and Year 2 teachers in a further six randomly selected schools.

In June 2021, we will conduct up to five telephone interviews with members of the development team and group leaders to explore: perceptions of the effectiveness of the training, intervention and follow-up support; perceptions of school or regional differences; challenges encountered, and how these were overcome.

Cost evaluation

The cost of programme delivery will be explored from the school's and developer's perspectives. Information will be collected about the cost of the intervention as it was delivered in the evaluation, and about what it would cost a school to self-fund the entire costs of delivering the intervention. As the programme is funded for intervention schools by the EEF and the University of Exeter, further cost information will be sought from them if needed. Costs will then be calculated as a cost per pupil from the school's perspective, as if schools were paying for the intervention, based on marginal financial costs. We do not propose to collect Business as Usual (BaU) cost data from control schools, as this may be unnecessarily costly for an efficacy trial.

Questions will be administered in the endpoint survey for schools, during the telephone interviews with intervention teachers, and during the telephone interviews with the development team. We will explore direct, marginal costs including: training costs, staff salary costs if over and above the hours of current staff; purchasing costs for resources, meals, subsistence, travel and any out of hours room hire. We will also report 'time' in terms of the amount of hours spent by staff and any other volunteers; and any re-allocation of existing resources (e.g. allocation of a named contact for the programme). We will report pre-requisite costs, if any. The intervention will be considered within the wider context of the costs of other behaviour programmes; taking into account existing costing methods and published costs. Costs per pupil will be estimated in terms of the overarching experiment i.e. what is the cost per randomised pupil regardless of their having received the intervention or not. Costs per pupil will also be estimated per school year, and then over multiple years (up to three years.)

Ethics and registration

The trial will be designed, conducted and reported to CONSORT standards (http://www.consort-statement.org/consort.statement/) and registered on http://www.controlled-trials.com/. The evaluation will be conducted in accordance with NFER's Code of Practice, available at: http://www.nfer.ac.uk/nfer/about-nfer/code-ofpractice/nfercop.pdf. The Code of Practice group will review the research at proposal and reporting stage, reviewing any instruments not previously used by the organisation and signing off on research methodology and outcomes. NFER, the University of Exeter, and EEF will work together to ensure each organisations' policies can be applied in practice.

Ethical Agreement

Ethical agreement for a school's participation within the trial will be provided by the headteacher of the school. Parents will be provided with full details about the intervention,

and will be given the opportunity to withdraw their child from data processing if they have objections to this. Participant opt-in consent will be sought for participants (teachers) in the interviews/case studies that form the IPE. All data gathered during the trial will be held in accordance with the General Data Protection Regulation (2016, applicable in the UK from May 2018), and will be treated in the strictest confidence by the NFER, EEF and the University of Exeter. Pupil data collected from schools by NFER will not be shared with any other parties. Our legal basis for gathering and using personal data is legitimate interest, through our work as a research organisation. Our legal basis for gathering special data is covered by GDPR Article 9 (2) (j) (see below). For qualitative data collected as part of the IPE, opt-in consent will be sought from participants.

Data protection

For the purpose of the research, NFER will collect and process both personal and special data.

- **Personal data:**

  The legal basis for processing personal data is covered by GDPR Article 6 (1) (f):

  Legitimate interests: the processing is necessary for your (or a third party's) legitimate interests unless there is a good reason to protect the individual's personal data which overrides those legitimate interests.

  We have carried out a legitimate interest assessment which demonstrates that the evaluation fulfils one of NFER's core business purposes (undertaking research, evaluation and information activities) and is therefore in our legitimate interest, that processing personal information is necessary for the administration of the randomised controlled trial. We have considered and balanced any potential impact on the data subjects' rights and find that our activities will not do the data subject any unwarranted harm.

- **Special data:**

  The legal basis for processing special data is covered by GDPR Article 9 (2) (j):

  *Processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subjec*t.

  We have concluded that the special data can be lawfully processed under GDPR Article 9 (2) (j), given that the two conditions specified by Sections 19.2 and 19.3 of the Data Protection Act 2018 are met: 1) the research is not likely to cause substantial damage or distress to the data subjects and 2) the purpose of the research is not to make decisions about particular data subjects. We considered conducting a Data Protection Impact Assessment for this matter, but concluded that it did not meet the criteria specified. As such, we carried out a Data Protection Advice

Record in order to explore issues. The DPAR can be accessed by request to incredibleyears@nfer.ac.uk

In setting out the roles and responsibilities for this trial, the three parties (NFER, the University of Exeter and EEF) will sign a Data Sharing Agreement (DSA). This includes a description of the data being collected and how it will be shared and stored by each party. In addition, the University of Exeter, supported by NFER, will provide a Memorandum of Understanding to schools, explaining the nature of the data being requested, how it will be collected, and how it will be passed to and shared with the University of Exeter (at recruitment stage) and NFER (during the trial).

For the purpose of the research, NFER will link data for all pupils in the trial with background and assessment information from the National Pupil Database, held by the Department for Education (DfE).

Personnel

Name	Institute	Roles and responsibilities
Dr Ben Styles (BS)	NFER	Trial Director, responsible for leading the NFER team and project delivery.
Gemma Stone (GS)	NFER	Trial manager, responsible for overseeing the day to day running of the trial and process evaluation
Suzanne Straw (SS)	NFER	Process Director, responsible for overseeing the process evaluation
Kathryn Hurd (KH)	NFER	Test and Schools administration lead, responsible for overseeing school contact and testing
Lisa Kuhn (LK)	NFER	Assessment advisor, responsible for guiding the team on selection and marking of appropriate assessments
Connie Rennie (CR)	NFER	Statistician, responsible for statistical analysis
Tamsin Ford (TF)	University of Exeter	Lead developer, responsible for delivery of the intervention
Rachel Hayes (RH)	University of Exeter	Lead developer, responsible for delivery of the intervention

Risks

Risk	Likelihood/ Impact	Mitigation
Insufficient schools recruited to the study	Likelihood: moderate Impact: high	NFER could help with recruitment if this becomes problematic.

School, teacher or pupil attrition	Likelihood: moderate Impact: moderate	Clear information provided to schools explaining the principles of the trial and expectations. Schools sign MOU with clear identification of requirements. Use of KS1 outcomes will result in minimal pupil attrition.
Intervention is not implemented well	Likelihood: low Impact: moderate	Clear information provided to schools explaining the principles of the trial and expectations. Both 'intention to treat' and 'on-treatment' analysis will be used. Good communication with delivery team to aid strong implementation. Process evaluation will monitor implementation.
Control group adopts similar treatments (contamination issues)	Likelihood: low (school randomisation) Impact: moderate	Process evaluation/email survey will monitor extent and nature of contamination in control group
Researchers lost to project due to sickness, absence or staff turnover	Likelihood: moderate, especially over 3 years Impact: moderate	NFER has a large research department with numerous researchers experienced in evaluation who could be redeployed.

Timeline

Dates	Activity	Staff responsible/ leading
Jun-Oct 2018	Set up meetings, IDEA Workshop, develop TIDieR and Theory of Change	GS
Oct-Dec 2018	Develop trial protocol, Data Sharing Agreement, Data Protection Impact Assessment/ Data Protection Advice Record, MoU, school information sheet, parental withdrawal form	GS, RH/TF
Jan-May 2019	School recruitment Develop IPE surveys	RH/TF GS, SS
Mar-Jun 2019	Collect pupil data, teachers complete online SDQ and PBQ for all pupils. Collect parental withdrawal forms	GS, KH
Jul 2019	Randomisation Schools informed of allocation	GS, KH
Sep 2019	All teachers complete baseline Practitioner survey	GS, KH
Oct 2019 - Mar 2020	Teacher training workshops (six in total, one per month) IPE case studies and telephone interviews	RH/TF GS
June 2020	Teachers complete midpoint online SDQ and PBQ for all pupils.	GS, KH
Sep 2020	Year two of intervention commences (Cohort 1 pupils begin Year 2)	GS

	Catch-up email to all schools to remind of responsibilities IPE case studies and telephone interviews	
May 2021	Pupils sit KS1 tests	(schools)
Jun-Jul 2021	Teachers undertake SDQ and PBQ online end-point surveys, revised Student Teacher Relationship Scale (STRS), and endpoint Practitioner survey Collect KS1 raw scores from schools	GS, KH
Aug-Dec 2021	Analysis of all outcomes for Cohort 1 Year 2 pupils, report writing.	CR, BS, GS, SS
Dec 2021	First draft of the report to EEF	GS
Jan-Mar 2022	Peer review, comments and adjustments to the report	All
Sept 2025	KS2 results (Y6 SATs) avaiable	-
Dec 2025	Possible addendum report (TBC): Year 6 SATs follow-up, attendance, exclusions	(NFER, commissioned separately)

References for Theory of Change

i. Patterson GR. Coercive family process. Eugene, OR: Castalia; 1982.
ii. Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 2, pp 191-215.
iii. Piaget, J., & Inhelder, B. Le developpement des quantites physiques chez l'enfant (2nd ed.). Neuchatel: Delachaux and Neistel, 1962
iv. Bowlby, J. (1951) Maternal Care and Mental Health. WHO, Monograph Series No. 2. Geneva: WHO.