EVALUATION OF STARS: INCREDIBLE YEARS® TEACHER CLASSROOM MANAGEMENT

Evaluation Report

September 2022

Sarah Tang, Kerry Martin, Eleanor Bradley, Ruth Staunton, Chris Morton, Ben Styles and Caroline Sharp

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education. Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

Jonathan Kay
Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

0207 802 1653

jonathan.kay@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Department for Education

What Works Network

THE SUTTON TRUST

Impetus Private Equity Foundation

Contents

About the evaluator 2

Acknowledgements 2

Executive summary 3

Introduction 5

Theory of Change 9

Methods 14

Impact evaluation results 28

Implementation and process evaluation results 38

Cost evaluation 61

References 71

Appendix A: EEF cost rating 74

Appendix B: Trial documents 75

Further appendices 82

About the evaluator

This project was independently evaluated by a team from the National Foundation for Educational Research (NFER): Sarah Tang, Kinnery Koria, Kerry Martin, Eleanor Bradley, Chris Morton, Caroline Sharp, Ben Styles, Kathryn Hurd, Kathryn Gallop, Julie Thompson, Jishi Jose, Kirstie Ramsay, Gemma Stone, and Connie Rennie.

Contact details:

National Foundation for Educational Research
The Mere Upton Park
Slough
Berkshire
SL1 2DQ
t: +44 (0)1753 574123

e: **s.tang@nfer.ac.uk**

Acknowledgements

We are grateful to all the primary schools that participated in this evaluation. The intervention, fieldwork and data collection took place during an extremely challenging and uncertain time for schools due to the Covid-19 pandemic and we appreciate the commitment participating schools showed. We are especially thankful to the school staff members who took case-study interviews and observations as well as those who responded to the surveys and completed the assessments.

We would like to thank the team at the EEF including Diotima Rapp, Arnaud Vaganay, Florentina Taylor, Katie Luxton and Maria Pomoni.

Executive summary

The project

Incredible Years® Teacher Classroom Management (IY®-TCM) is a continuing professional development programme which aims to develop teachers' classroom management techniques, improve behaviour in the classroom, reduce disruption to learning, and increase attainment.

IY®-TCM is targeted at practitioners who teach three- to eight-year-olds. It is part of a complementary suite of interventions created by Incredible Years® designed to support children to develop their emotional regulation and social competence. The programme involves multiple strategies for reducing inappropriate behaviour and encourages teachers to support students to regulate their own emotions. Teachers integrate these strategies into their everyday classroom practice. The strategies are expected to be particularly helpful for students with behavioural difficulties.

In IY®-TCM, teachers are trained in six whole-day workshops, delivered over a period of six months, by behaviour support workers. The focus is on collaborative learning, discussions of teachers' own experiences and working as a group to find solutions to problems encountered in the classroom. Supporting materials include a handbook, handouts and key learning summaries.

This evaluation sought to evaluate IY®-TCM at efficacy level. It involved 139 schools and 5,623 pupils from six regions, engaged in a two-arm, cluster randomised controlled trial (RCT). Schools were randomly allocated to either receive training in IY®-TCM for all Year 1 and Year 2 teachers, or act as a control group. Teachers in the intervention group received training from October 2019 – March 2020, and pupils were followed from Year 1 (aged 5-6 years) into Year 2 (aged 6-7 years). Year 2 teachers received training a year prior to teaching the pupils in the trial.

This programme and its evaluation were affected by the 2020 partial school closures caused by the Covid-19 pandemic, and the cancellation of KS1 SATs in 2021. As a result, the evaluators were not able to use maths attainment at KS1 to estimate the impact of IY®-TCM on attainment, and it was therefore not possible to rate the security of impact estimates.

Secondary outcomes evaluating pupil wellbeing, pupil behaviour in class, and relationships between teachers and pupils could not be collected in the first year (summer 2020) of the evaluation but were collected in the second year (summer 2021). Mixed methods research exploring training attendance, delivery in the classroom, teacher stress, and self-efficacy were also partially impacted due to COVID-19 measures; some, but not all, of this information could be collected.

The programme was delivered from October 2019 to July 2021. During the trial, exposure of pupils to the programme and time spent in the classroom with the teacher was reduced. The final teacher training session was also cancelled for two out of 16 training groups. The evaluation was conducted by the National Foundation for Educational Research (NFER), and delivery was managed by the University of Exeter.

Figure 1: Key Conclusions

Key conclusions
1. Due to Covid-19, the primary outcome for this evaluation was not collected, so no measure of impact on KS1 maths attainment is reported and no padlocks are assigned. Key conclusions are based on secondary outcome measures and qualitative data from the implementation and process evaluation.
2. The secondary outcome measures found no evidence that the IY®-TCM programme influences behaviour, social and emotional wellbeing, concentration, prosocial behaviour or student-teacher relationships for pupils receiving the intervention compared to pupils in the control group. These results are based on surveys completed by Year 2 teachers at the end of the project, comparing the intervention and control group. However, due to low response rates, they are based on 55 per cent of pupils in the trial. Furthermore, 21 per cent of the pupils in the second year of the trial were taught by teachers who were not on the training register for the IY®-TCM training, so these pupils did not have any exposure to the trained teachers when in Year 2[1].

[1] The proportion of pupils in Year 2 taught by a teacher who had not been trained in IY®-TCM is likely to be slightly higher than 21 per cent as some teachers were on the training register but did not attend any workshops.

3. Only 60 per cent of the intervention teachers attended four or more of the training sessions. Pupils of those teachers demonstrated better behaviour, compared to the control group, using surveys completed by classroom teachers at the end of the trial. This impact was small and there may be other differences between teachers that attended sessions and those that did not which could contribute to the difference in impact.

4. Teachers who attended the training showed larger improvement in perceived self-efficacy compared to those in the control schools, when measured through surveys at the start and end of the trial. However, due to low response rates, this is based on fewer than half of teachers in the trial.

5. Findings from interviews with teachers, the developer team, and trainers suggest that the combination of the costs to schools in releasing all Year 1 and Year 2 teachers and covering for their absence from the classroom was a factor in the low attendance of training sessions, when considered alongside the perceived value of the programme in addressing schools' needs.

Additional findings

Due to Covid-19, the primary outcome for this evaluation was not collected, so no measure of impact on KS1 maths attainment is reported. Key conclusions are based on secondary outcome measures for behaviour, social and emotional wellbeing, concentration, prosocial behaviour and student-teacher relationships, and qualitative data from the implementation and process evaluation.

There is no evidence that the IY®-TCM programme influences behaviour, social and emotional wellbeing, concentration, prosocial behaviour or student-teacher relationships for pupils receiving the intervention compared to pupils in the control group. These results are based on surveys completed by Year 2 teachers, in the intervention and control groups, at the end of the project. However, these surveys had very low response rates, which means that the results are based on 55% of pupils. Furthermore, at least 21 per cent of pupils analysed did not receive the intervention from an IY®-TCM trained teacher in Year 2.

Further analysis, looking at the impact when teachers in Year 1 and Year 2 attend at least four out of the six training sessions, indicates pupils demonstrated better behaviour, when comparing surveys completed by classroom teachers in the intervention and control group, at the end of the trial.

The results from this evaluation partially support previous trials of IY®-TCM, which have used behavioural outcome measures. Results from an RCT undertaken by the delivery team (University of Exeter) in 80 schools in England, with pupils aged four to nine years, indicated improvements in low-level behaviour at nine months, 18 months, and 30 months (Ford *et al.,* 2019). In this trial, where pupils were taught by teachers who attended at least four training sessions, there was a small positive impact on behaviour.

Although it was not possible to evaluate the impact of the intervention on attainment, exploration of the Theory of Change through mixed methods research does support the initial stages of the causal chain. Teachers who attended training saw larger improvements in their perceived self-efficacy compared to the control group, when comparing survey results at the start and end of the trial. Similarly, there was also some evidence that behaviour management had improved more for the intervention teachers than those in the control group over the trial period, drawing on self-reported surveys.

In general, surveys indicated teachers were very satisfied with quality of IY®-TCM training, however, these were completed by less than half of teachers in the intervention group, and so results may not reflect those who chose not to respond. Only 60 per cent of the intervention teachers attended four or more of the training sessions. Interviews with teachers, the delivery team, and trainers suggest that the burden on schools to release all Year 1 and Year 2 teachers was a key factor in attendance. This potential barrier to implementation, is particularly important given the differing impact results depending on training attendance. Ensuring that senior leadership and teachers had buy-in and understood the aims of the intervention - to address low-level behaviour disruption - were also identified as potential barriers to implementation.

Cost

The cost per pupil per year over three years is estimated as £27, which is very low cost. This is based on an estimated cost for delivery of IY®-TCM equivalent to £14 per pupil per year, and additional costs borne by the school estimated at £12 per pupil per year, which includes cover and travel costs to attend training. Schools were not charged for the intervention in this trial.

Introduction

Background

The extent of disruptive behaviour in English schools is much debated not least as it is hard to measure consistently and objectively. Haydn (2014) brings together evidence from four studies (interviews and surveys) and concludes that behaviour is a larger problem in schools that is often reported by Ofsted and government commissioned reviews (e.g., Steer, 2009) concluding that there appear to be few schools where the levels of behaviour are considered not to disrupt the class in any way.

The Covid-19 pandemic has led to heightened concern about the social and emotional wellbeing of children as they have been living with tremendous uncertainty, home-schooling and lack of socialisation with peers and extended family. Research from NFER showed that 81 per cent of schools surveyed had a focus on the emotional and social wellbeing on the return to school in September 2020 (Sharp *et al.*, 2020). The government has produced a Covid-19 mental health surveillance report which brings together evidence on levels of mental health over the period of Covid-19 and lockdown (OHID, 2021). A key source for younger children is from the Co-Space study which surveyed 9161 parents/carers across the UK (Skripkauskaite et al., 2021). There were increases in behavioural and attention difficulties at times of peak restrictions and school closures.

While it is perhaps self-evident that some control of a classroom is necessary for effective teaching and learning quantitative evidence is relatively rare. The Hay McBer Report commissioned by government to develop a model of teacher effectiveness found that lack of disruption correlated with pupil progress at primary level (Hay McBer, 2000). Goodman and Gregg (2010) again using data from ALSPAC find that pupil behaviour is one of the factors that explains widening of the attainment gap between those from richer and poorer backgrounds at primary school.

Levels of poor behaviour in pupils are also associated with higher levels of teacher stress and wellbeing as well as higher levels of staff turnover (e.g., Ofsted, 2019; Burge, Lu, & Phillips, 2019). Given teachers are a key determinant in pupil attainment, it seems that the recruitment and retention of effective staff provides another hypothesised link between pupil behaviour and attainment.

The Incredible Years® Teacher Classroom Management (IY®-TCM) programme aims to improve pupil behaviour and readiness to learn in the classroom by equipping teachers to support pupils' social and emotional wellbeing. The programme was developed in the United States and is one of a series of programmes by Incredible Years®' founder Carolyn Webster-Stratton. Incredible Years® series of programmes aim to address children's social and emotional wellbeing aged from birth to 12 years. There are separate but interlocking suites of programmes for teachers, parents and children. IY®-TCM is one of three programmes for teachers[2] and focusses on developing teachers' classroom behaviour management techniques and methods to foster children's prosocial behaviour.

IY®-TCM is for teachers of children aged three to eight years. There are six full-day training sessions which are delivered over an extended period of five/six months in order to allow participating teachers to try out techniques they have learnt between sessions. The strategies promoted through the training sessions include the use of increased use of praise, coaching and boundary setting with reduced use of timeout and consequences (see 'Teaching Pyramid®' in Appendix D). During the training sessions, behaviour management strategies are discussed and observed through video examples of classroom practice and role-playing scenarios. For more detail on the programme see *Intervention* section.

The Early Intervention Foundation guidebook[3] rates the evidence of the effectiveness of IY®-TCM as level 3+ (out of 4 possible levels) based on three RCTs undertaken in Ireland, USA and Wales. The Welsh trial (Hutchings *et al.*, 2013) was run in 11 primary schools in Wales with randomisation at class level. The outcome measure used was the Teacher-Pupil Observation Tool (TPOT) and independent assessors were used who were blinded to allocation. Eight or nine pupils were selected from each class with a range of scores using the Teacher version of the Strengths and Difficulties Questionnaire (TSDQ). None of the class-level outcomes from the TPOT measure showed a significant difference

[2] **https://incredibleyears.com/about/incredible-years-series/**
[3] **https://guidebook.eif.org.uk/programme/incredible-years-teacher-classroom-management**

between control and intervention groups except for the 'child off task' component where classes where the teachers had received IY®-TCM were shown to have less time off task then those whose teachers had not.

A more recent trial in Irish schools involving random allocation of 22 teachers across 11 schools found a significant effect (Hickey *et al.*, 2017) on self-reported teacher behaviours with an increase in positive classroom management practices (0.56) and a decrease in negative strategies (-0.43). A larger study in the US (Reinke, Herman and Dong, 2018), found statistically significant impacts of the programme on a number of teachers reported pupil-level measures including prosocial behaviour (0.13), emotional dysregulation (-0.14), and overall social competence (0.13) as measured using the Teacher Observation of Classroom Adaption Checklist (TOCA-C) and the Revised Social Competence Scale – Teacher version (T-COMP).

The delivery team (University of Exeter) have also completed an RCT in 80 schools in England with pupils aged four to nine years (Ford *et al.*, 2019). This trial was the first in England using the IY®-TCM programme and considered mental health (SDQ) and behaviour outcomes (Pupil Behaviour Questionnaire (PBQ)). Researchers found evidence of a positive impact on mental health of pupils at nine months after the beginning of the IY®-TCM training, but this improvement was not sustained at later follow-ups (*ibid*, 2019). There was however evidence of improvements in low-level behaviour (as measured by the PBQ) at all follow-ups after the intervention training began – at nine months, 18 months and 30 months.

Existing trials of IY®-TCM have used behavioural outcome measures for the primary outcome and existing evidence suggests children's well-being is linked to academic attainment (Gutman and Feinstein, 2008; Zins *et al.*, 2004). This evaluation is an efficacy trial and was planned to examine whether there is an impact of the programme on attainment through looking at maths attainment at the end of Key Stage 1. Maths and literacy were considered as the attainment outcome as they are collected routinely at the end of Key Stage 1. Maths was chosen over literacy based on evidence suggesting that maths attainment is linked to self-regulation interventions and executive functioning (e.g., Ng, Bull and Khng, 2021). Due to the Covid-19 pandemic Key Stage 1 assessments were cancelled in 2021 which meant the collection of the primary outcome was not possible. After consideration by the delivery team, the evaluation team and the EEF, it was decided that the trial would continue with the secondary measures which looked at behaviour and other measures associated with mental health. The PBQ was the main secondary outcome as it was considered to be the first link in the causal chain in achieving impacts in attainment outcomes.

The trial is a two-armed cluster randomised trial with randomisation at school level. A total of 157 schools were recruited and 139 were randomised to take part in the trial (as some schools withdrew prior to randomisation). Randomisation took place in summer 2019 after baseline data collection and intervention schools then were invited to attend six days of training from October 2019 to March 2020. Intervention schools were asked to send all of their Year 1 and Year 2 teachers to attend the training. If Year 2 teachers remained teaching a Year 2 class in the following academic year in the same school, then the trial cohort of pupils would have exposure to an IY®-TCM trained teacher for two consecutive academic years over the trial period.

Intervention

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 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 from a meta-analysis suggests that the prevalence of attention-deficit hyperactivity disorder for children and adolescents worldwide is 3.4 per cent and for any disruptive disorder is 5.7 per cent (Polanczyk *et al.*, 2015). 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 systematic review evaluated teacher training interventions aimed at improving children's social and emotional wellbeing through classroom management. The review included 14 studies looking at eight interventions and found that IY®-TCM had one of the stronger evidence bases (along with the Good Behaviour Game) (Whear *et al.*, 2013). Some studies have found early evidence of the impact of IY®-TCM on pupil social-emotional development. Ford *et al.* (2019) report a small short-term improvement to children's mental health, particularly for those who are already struggling. Nye (2017) also shows an overall effect on pupil mental health, with a significant effect on children with behaviour problems.

This intervention has been trialled recently with 80 schools (Ford *et al*., 2019), in order to assess, amongst other elements, impact on pupil mental health and wellbeing. This independent evaluation is 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) was 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. Years 1 and 2 were chosen as this facilitated use of Key Stage 1 assessments as the primary outcome measure reducing the burden on schools. Key Stage 2 assessments take place when pupils are older than the target age range for this IY®-TCM intervention.

The target regions were Reading, Cornwall, Bristol, Southampton (Hampshire), Dorset and Liverpool. The training was offered to all 70 schools (and 230 teachers) in the intervention group. To be eligible to receive the training, teachers must have had at least four days of teaching per week. In cases where teachers were on a job share, both teachers sharing the job were trained (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 was delivered to both Year 1 and Year 2 teachers, so that when the Year 1 pupils move up to Year 2, they were taught by teachers who had 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 received the intervention. Though these Year 2 pupils are not part of the trial, both Year 1 and Year 2 teachers were part of the trial. Only pupils who started the intervention in Year 1 and continued on to Year 2 had their data analysed for the impact evaluation.

What: Materials

The University of Exeter recruited 'group leaders' to deliver the training to teachers. These group leaders attended a consecutive three-day Incredible Years® accredited training course, where they were provided with a training manual (Leader's Manual) that included suggested scripts, videos, a day-by-day programme, posters, stickers and worksheets. These manuals were used by group leaders to deliver the IY®-TCM training to the study teachers.

Teachers in the intervention schools received six one-day training sessions, delivered by the group leaders once a month for six months. Group sizes were determined by the number of teachers being trained per school but the target number is to have four schools per session i.e., twelve teachers. At commencement teachers were provided with the *Incredible Teachers* handbook, which contains information relating to the whole course. At each training session, each teacher received handouts for the session and to take away, as well as 'buzz' documents (summaries of key learning points).

What: 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, pupils' 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

Twelve group leaders (trainers) all completed the three-day IY®-TCM basic training led by a certified IY®-TCM trainer. These group leaders delivered the training of the intervention as described above. Six of the group leaders were experienced in delivering IY® training and the remaining six had behavioural support related roles. Experienced group leaders ran half of the IY®-TCM groups alone. The other groups were delivered in pairs with experienced group leaders delivering their sessions jointly with a less experienced trainer, this enabled the more experienced group leader to provide mentorship and training the less experienced trainer and ensured that all groups were run with a high degree of fidelity.

Teaching assistants were not directly trained as part of the intervention, but since teaching assistants spent a large amount of time with the study children, teachers were encouraged to share information and training documents with them. This decision was pragmatic in relation to the resources and time available to the trial.

How: Modes of delivery

The University of Exeter recruited schools to the trial in the target areas. The group leader planned to train groups of minimum eight and maximum 15 teachers; the minimum is set to retain a group dynamic. Some groups were larger than 15 due to demand for places but group leaders were asked that they were happy to do this. Ideally, each group should include a mix of teachers from at least four different schools from the geographical hub.

Group leaders and their group stayed together throughout the training period (other than in emergencies when a teacher was unable to attend due to unforeseen circumstances). Given that the course is experiential, involves creating a collaborative and safe environment and builds on each successive session, it is important that teachers do not miss sessions or send a replacement as each session builds on previous sessions with participants trialling approaches in between sessions and feeding back to the group at the next session. In the previous trial of IY® in England (Ford, *et al*., 2019) 90 per cent of teachers attended at least four sessions with 58 per cent attending all six. Complier Average Causal Effect (CACE) analysis was undertaken to see if there was a difference in effectiveness if all six sessions were attended. The results were almost identical to the main results suggesting that attending all six sessions was not necessary to have an impact on pupils' mental health (as measured by the SDQ). For this trial, it was agreed that if a teacher attended at least four of the six sessions they would be considered to be compliant with the intervention.

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 should be encouraged/supported throughout the full school day.

Where: Location and setting

The intervention took place in Reading, Cornwall, Bristol, Southampton/Hampshire, Dorset and Liverpool. The developer team chose these regions by considering areas that the trainers could travel to relatively easily to deliver the workshops and that had higher levels of deprivation in terms of pupils eligible for free school meals. Training was delivered at a regional location, away from the participating schools. The school settings are described above ('Who; recipients of the intervention').

When and how much: Duration and dosage of the intervention

The training consisted 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). The expectation is that teachers use the strategies and behaviours every day.

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'.

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.*, 2019). The only change to the delivery of the intervention in this trial as compared to the previous one ran by the developer team was that in the original trial all groups were led by two experienced group leaders. The main changes in this trial in terms of design were the addition of an attainment outcome (although ultimately this was not possible due to cancellation of Key Stage 1 assessments) and the size of the sample. It is also worth noting that in the original trial (Ford *et al*., 2019) only one teacher and one class from each school took part in the trial. 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.

Techniques to ensure success in the implementation included selecting highly experienced group leaders. Half of these were individuals who have previously delivered the IY®-TCM training, while some were new to the programme. New group leaders were carefully selected based on their previous experience, which included experience of behavioural support or the delivery of other IY® programmes. They received the mandatory three-day training recommended by the IY® team, completed the recommended checklists after every session to standardise practice. The less experienced trainers were paired with highly experienced trainers to deliver the workshops, to maximise fidelity of delivery and to ensure that the less experienced could learn from and be supported by their more experienced colleagues.

The developer team reported that the relationship with schools was a top priority to ensure successful delivery. Behavioural management is something that causes teachers a great deal of stress and distress. In the feasibility study for the original STARS trial some teachers felt that their skills were being questioned by being sent on the course (Marlow *et al*. 2015). Similarly, good communication about the research process in terms of the timing and purpose of classroom observations and surveys was perceived as essential in the original STARS feasibility work (Marlow *et al*, 2015). To aid effective relationships with schools in the current trial a named contact was established as soon as a school was allocated to the intervention group and throughout the first year of the programme. There were also named and easily contactable individuals from the developer team available to respond to queries and comments from participants regarding the intervention. Schools were allocated into groups or 'hubs' and had an allocated trainer/group leader; this established a small network of support between schools and the trainer which aimed to support positive relationships.

Theory of Change

The Theory of Change for the IY®-TCM programme was developed for this trial with the University of Exeter. The Logic Model developed by the designers of the programme is in Appendix E.

Figure 1: Theory of Change

Assumptions

Evidence suggests that some types of negative or coercive interaction between children and adults reinforces disruptive behaviour (Patterson, 1982). In previous trials, the STARS intervention has successfully impacted upon social and emotional development (Ford *et al.*, 2019) 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); and developmental interactive learning methods (Piaget and Inhelder, 1962), and incorporates cognitive behavioural approaches and Bowlby's attachment theory on the importance of positive relationships (Bowlby, 1951). 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.*, 2019)

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/Hampshire, Liverpool, Dorset
Pupils: Year 1 (Cohort 1) with follow up at the end of Year 2 (i.e., 2-year intervention)

Evaluation objectives

The original research questions are below. Due to the cancellation of Key Stage 1 assessments in summer 2021 it was not possible to examine questions 1, 5, 6 and 7.

The revised protocol and statistical analysis plan are available on the **EEF website**.

Primary research 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 research 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/inattention 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 SDQ, compared to 'business as usual'?

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 greater 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 RQ6, however the SDQ responses will be included as a continuous variable, not a categorical variable.

Ethics and trial registration

Ethical agreement for a school's participation within the trial was provided by the headteacher of the school. Parents were provided with full details about the intervention and were given the opportunity to withdraw their child from data processing if they had any objections to this. Participant opt-in consent was sought for participants (teachers) in the interviews/case studies that form the IPE. All data gathered during the trial were held in accordance with the General Data Protection Regulation (GDPR - 2016, applicable in the UK from May 2018), and were treated in the strictest confidence by the NFER, EEF and the University of Exeter. Our legal basis for gathering and using personal data was legitimate interest, through our work as a research organisation. Our legal basis for gathering special data was covered by GDPR Article 9 (2) (j) (see below). For qualitative data collected as part of the IPE, opt-in consent was sought from participants.

Copies of the information sheet and privacy notices are included in Appendix B. The trial is registered on the ISRCTN registry at **https://doi.org/10.1186/ISRCTN77306152**.

Data protection

For the purpose of the research, NFER collected and processed both personal and special data.

Personal data:

The legal basis for processing personal data was 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 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. Personal pupil data was collected from schools. This data included name, date of birth, gender, FSM eligibility, UPN and teachers' questionnaire responses about each child's behaviour (Pupil Behaviour Questionnaire and Student Teacher Relationship Scale).

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 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 was not likely to cause substantial damage or distress to the data subjects and 2) the purpose of the research was not to make decisions about particular data subjects. We considered conducting a Data Protection Impact Assessment but concluded that it did not meet the criteria specified. As such, we carried out a Data Protection Advice Record to explore issues. Special data collected relating to pupils were teachers' questionnaire responses about pupils' mental health and wellbeing, gathered through the Strengths and Difficulties Questionnaire.

Personal and special data relating to pupils were collected from schools and transferred between schools and NFER via an online secure portal (personal data) or via Questback - secure online survey software. Data was then downloaded and saved to a secure server. Data collected will be deleted a year after the final report is published.

In setting out the roles and responsibilities for this trial, the three parties (NFER, the University of Exeter and EEF) signed 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, provided 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).

Under the original design (pre-Covid-19) there were plans for NFER to link data for all pupils in the trial with background and assessment information from the National Pupil Database (NPD), held by the Department for Education (DfE). This was no longer necessary as the primary outcome of the trial was not possible to collect due to Covid-19 related disruptions. The pupil-level data will however be deposited in the EEF's data archive.

Project team

The delivery team was based at Exeter University and the evaluating organisation was the National Foundation for Educational Research (NFER). The names and roles of key team members are included in the Table 3 below.

Table 3

Name	Affiliation	Role in project
Rachel Hayes	University of Exeter	Lead developer, responsible for delivery of the intervention
Tamsin Ford	University of Cambridge (was University of Exeter)	Lead developer, responsible for delivery of the intervention
Bryony Longdon	University of Exeter	Developer, responsible for the day-to-day delivery of the intervention
Ben Styles	NFER	Project Director, responsible for leading the evaluation team and project delivery
Gemma Stone/Sarah Tang	NFER	Project Leader, responsible for overseeing the day to day running of the trial and process evaluation
Constance Rennie/Chris Morton	NFER	Statistician, responsible for statistical analysis and design
Caroline Sharp	NFER	Project Director (IPE), responsible for overseeing the process evaluation
Kerry Martin	NFER	IPE Lead, responsible for IPE instrument design and analysis
Eleanor Bradley	NFER	IPE researcher, IPE fieldwork and analysis
Kathryn Hurd	NFER	Test and Schools administration lead, responsible for overseeing school contact and testing
Jishi Jose/Kinnery Koria	NFER	Project manager (operations), responsible for the day-to-day contact with schools
Lisa Kuhn	NFER	Assessment advisor, responsible for guiding the team on selection and marking of appropriate assessments

Methods

Trial design

Table 4: Trial design

Trial design, including number of arms		Two-arm, cluster randomised
Unit of randomization		School
Stratification variable (s) (if applicable)		Region, number of reception teachers (school size proxy)
Primary outcome	Variable	Not measured*
	Measure (instrument, scale, source)	Not measured*
Secondary outcome(s)	Variable(s)	Pupil classroom behaviour Pupil emotional and social well-being Pupil concentration Pupil prosocial behaviour Student teacher relationship
	Measure(s) (instrument, scale, source)	Pupil Behaviour Questionnaire, 0-12, (Allwood *et al*., 2018) SDQ Total Difficulties Score, 0-40, SDQ SDQ Hyperactivity Scale, 0-10, SDQ SDQ Prosocial Scale, 0-10, SDQ Student Teacher Relationship Scale (revised version), 25-75, adapted from Pianta (2001)
Baseline for primary outcome	Variable	N/A
	Measure (instrument, scale, source)	N/A
Baseline for secondary outcome(s)	Variable	Pupil classroom behaviour Pupil emotional and social well-being Pupil concentration Pupil prosocial behaviour
	Measure (instrument, scale, source)	Pupil Behaviour Questionnaire, 0-12, (Allwood *et al*., 2018) SDQ Total Difficulties Score, 0-40, SDQ SDQ Hyperactivity Scale, 0-10, SDQ SDQ Prosocial Scale, 0-10, SDQ

*Planned primary outcome was mathematics attainment in Key Stage 1 assessments but these assessments were cancelled in 2021 due to school disruptions relating to Covid-19.

The trial is an efficacy trial as although IY®TCM has previously been trialled across 80 English schools, that trial did not include attainment outcomes (Ford *et al.*, 2019). The design is a two-armed RCT with randomisation at the school level as the intervention is a CPD programme for teaching staff. School-level randomisation was considered appropriate as the programme is concerned with how teachers approach behaviour challenges, and it is likely to be related to school context and the school's broader behaviour policy. It was felt that contamination between intervention and control groups would have been considerably more likely if randomisation was at class level as teachers within the same school (and year group) are more likely to plan together and share approaches than with those in other schools. The trial was run over two academic years (2019-20 and 2020-21) and with a cohort of primary pupils who were in Key Stage 1 (Year 1 and then Year 2). In the summer term 2019 all schools who were participating on the trial completed baseline pupil-level assessments for their pupils in the reception classes. The baseline measures were the SDQ and the PBQ. Reception

teachers completed these measures for the pupils in their classes. Once baseline measures had been completed the school was put forward for randomisation which occurred in July 2019.

The IY®-TCM training began in October 2019 for all Year 1 and Year 2 teachers in schools which had been allocated to the intervention group. For schools allocated to the control group teachers continued with business as usual which may include CPD activities (though not IY®-TCM). Six training days took place over the following five/six months in six different regions. Each region ran between one and four groups. Due to Covid-19 two of the 16 planned groups' final training days were cancelled[4]. There were 18 teachers in these two cancelled groups. On 18 March 2020, around the same time as the final workshops, school buildings closed for the majority of pupils in English schools (Timmins, 2021). Schools opened again for Year 1 pupils on 1 June 2020 although this took a range of forms with some schools only opening for part time attendance for the rest of this summer term. Under the original design a midpoint data collection using the SDQ and PBQ was due to take place in the summer term 2020 but this was cancelled due to school closures.

Endpoint assessments took place in summer 2021. In addition to the SDQ and PBQ a revised version of the Student Teacher Relationship Scale (STRS) was included at endpoint. In the planned trial design, the primary outcome measure was the maths element of the Key Stage 1 assessments. Due to school closures and the pressures on schools relating to Covid-19, Key Stage 1 assessments were cancelled for summer 2021[5] and therefore no primary outcome was collected.

The programme was delivered at no cost to the schools other than the costs implicit with teaching staff attending the training days such as travel and supply cover. Incentives were paid to both control and intervention schools on completion of the endpoint assessments. The incentives were £200 for completion of 75 per cent of the pupil-level assessments requested for each full-time teacher at end point. The same amount was also available for schools once the midpoint assessments had been completed. The decision was taken between EEF and University of Exeter to pay schools the midpoint incentive payments as some schools reported having made budgetary decisions on the basis of receiving these extra funds (i.e., some schools had planned to use this toward the costs of arranging supply teachers to cover the training days).

Participant selection

Primary schools were recruited to the trial. To be eligible for the trial, schools had to be mainstream state-maintained primary schools who would be able to attend training sessions in one of the six regions. Different regions had different recruitment targets depending on the number of groups that were planned for each area[6]. Estimates for the number of schools needed to fill a group were based on three teachers in Key Stage 1 in each school though these estimates were revised once schools were recruited, and the actual size of the schools was known. Schools in special measures were excluded from participating. All Year 1 and Year 2 teachers in the intervention schools were eligible and expected to take part provided that they had classroom responsibility with their Year 1 or Year 2 class for at least four days a week. Where classes were split more evenly between teachers (i.e., where teachers taught the Year 1 or Year 2 class for 3 days or fewer) both teachers were expected to attend.

The University of Exeter undertook the recruitment of schools to the trial. Over 4,000 schools were approached through unsolicited emails or letters during the academic year 2018-19 outlining the trial. If a school expressed an interest in taking part, Exeter sent them further information including the school privacy notice, memorandum of understanding and information sheet (see Appendix B for copies of these documents). A total of 158 schools were recruited and their contact information was shared with NFER. NFER then shared the parent information sheet, parent privacy notice and 'parent withdrawal from data processing' letter with schools via secure school portals. Schools were asked to share these documents with parents/carers (see **Data protection section**). NFER then shared an empty data template via the portals which schools completed to provide information about their reception pupils (provided that their parents had not requested that their child be withdrawn from data collection[7]). NFER used this pupil-level data to prepopulate the

[4] There were 18 groups initially but two were cancelled after the first training day due to low numbers. Teachers due to attend these cancelled sessions were offered sessions with other groups.
[5] **https://www.gov.uk/government/publications/2021-key-stage-1-assessment-and-reporting-arrangements-ara**
[6] Initial recruitment targets per region were as follows: Bristol (4 groups) – 28 schools, Cornwall (2 groups) – 20 schools, Dorset (3 groups) – 28 schools, Reading (4 groups - went down to 3 groups after first workshop) – 28 schools, Liverpool (3 groups) – 28 schools, Hampshire (2 groups - went down to 1 group after first workshop) – 16 schools.
[7] 28 pupils (across 19 schools) withdrew from the data collection.

baseline measures (SDQ and PBQ) into online forms for the reception teachers to complete. Reception teachers across all schools were asked to complete the baseline measures via online forms. Once the baseline measures were completed the school was put forward to randomisation. Of the 158 schools initially recruited 139 completed the baseline outcome measures and went on to be randomised in July 2019.

Outcome measures

Baseline measures

Under the original protocol selected elements of the Early Years Foundation Stage Profile (EYFSP) were planned to be used as a baseline measure for the primary analysis. The EYFSP is a measure of a child's attainment in relation to 17 early learning goals[8] prior to the age of five. It is intended to provide a reliable, valid and accurate assessment of a child at the end of Foundation Stage. As the Maths attainment primary outcome was no longer possible it was not necessary to collect the EYFSP from the NPD to be used as a baseline covariate in this analysis.

For all the secondary analysis the baseline measures of the same instrument were used other than for the Student-Teacher Relationship Scale which was only collected at endpoint. For the STRS the baseline measurement of the PBQ was used instead.

Originally, we had planned to take a midpoint measurement of the PBQ and the SDQ (at the end of the first year of the intervention). However, due to disruptions caused by Covid-19, this was not possible. As such, measurements of the SDQ, PBQ and STRS were collected at one follow-up time-point only: at the end of the second year of the intervention.

Primary outcome (to answer RQs1, 5, 6, 7)

The primary outcome as originally planned was maths attainment as measured by the arithmetic and reasoning papers of the Key Stage 1 assessments. Key Stage 1 was chosen as it is a statutory assessment completed by all eligible Year 2 pupils so there is a lower burden on schools as compared to using an alternative assessment. Maths attainment was chosen as the primary outcome (rather than literacy which is also assessed at the end of Key Stage 1) as it has been linked to self-regulation interventions and executive functioning (e.g., Ng, Bull and Khng, 2021). Schools were to provide the raw score of these scales to NFER and the total of the two papers was the planned outcome measure. However, due to the cancellation of Key Stage 1 assessments this measure was not possible and therefore there is no primary outcome for this trial. Alternative assessment options to the Key Stage 1 assessments were explored and discussed within the evaluation team and EEF but it was decided that there was no viable option that would not add burden to schools at an already challenging time.

Secondary outcomes

All secondary outcomes are at pupil level but are completed by the pupil's main classroom teacher.

1. Pupil Behaviour Questionnaire (PBQ) (to answer RQ3)

The PBQ is an instrument that was developed by the University of Exeter. It was validated in 2018 and was used in the only previous English trial of this programme run by Exeter (Allwood *et al.*, 2018; Ford *et al.*, 2018). Whereas the SDQ focuses upon indicators of pupil mental health and individual behaviours, this scale captures behaviours that have a negative impact on the classroom. As such, it relates to an important aspect of the programme: managing behaviour within the classroom. It is a six-item scale, each rated on a three-point scale (never happens, occasionally happens, frequently happens). The scores range from 0-12 with a higher score reflecting worse behaviour. It has a Cronbach's alpha value of 0.85 (Allwood *et al*., 2018). Factor analysis completed with data from Ford *et al.* (2018) found that the PBQ measures a single construct and correlates in the expected direction with the teacher SDQ.

2. Pupil emotional and social well-being (to answer RQ2a)

Emotional and social wellbeing was measured using the Total Difficulties score of the SDQ (Goodman, 2001). The SDQ was chosen as it is a reliable measure of pupils' emotional and social wellbeing (Goodman, 1997). It consists of 25

[8]https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/790580/EYFSP_Handbook_2019.pdf

items, split into five subscales with five items each (emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, prosocial behaviour, each scored 0-10). The Total Difficulties score (TDS) is the sum of four of the five subscale scores - emotional symptoms, conduct problems, hyperactivity/inattention, and peer relationship problems. It has a possible range of between 0 and 40, with a higher score representing a higher level of difficulty. The TDS was coded using the syntax published by Youth in Mind[9] and was analysed as a binary outcome, (in line with categorization used in the previous trial, Ford, *et al.*, 2019), as well as a continuous outcome. A score of 12 or above represents the 80th percentile of the British school-age population, and those above the score of 12 have been classified as 'strugglers' in previous analysis of the IY®-TCM programme (Ford *et al.*, 2019). This cut-off point was used for the purpose of comparing to previous research.

3. Pupil concentration (to answer RQ2b)

Pupil concentration was measured using the hyperactivity/inattention subscale of the SDQ. This is a five-item scale that reliably measures levels of restlessness and distractibility. Each item is scored between 0-2 giving a range of scores between 0-10. Higher scores reflect lower levels of concentration. This subscale was chosen as it fits closely with the theory of change and improvements were found in previous investigations of the programme (Ford *et al.*, 2019).

4. Pupil prosocial behaviour (to answer RQ2c)

Pupil prosocial behaviour was measured using the prosocial subscale of the SDQ. This is a five-item scale that measures levels of helpfulness and kindness. Each item is scored between 0-2 giving a range of scores between 0-10. A higher score reflects higher levels of prosocial behaviour. This subscale was chosen as it again fits closely with the theory of change and improvements were found in previous investigations of the programme (Ford *et al.*, 2019).

5. Student Teacher Relationship Scale (to answer RQ4)

The STRS is an instrument designed for teachers of children aged between three and 12 which measures a teacher's perception of conflict, closeness and dependency with a specific child (Pianta, 2001). The short version of the instrument was suggested for this trial as it allows measurement of pupil-teacher relationship, a focus of the intervention that is not captured in other outcome measures. Following internal discussion, NFER proposed that some questions, through emphasis and phrasing, were inappropriate for a UK context. Following discussion with EEF, a revised version was produced jointly by NFER and the University of Exeter developer team. This scale consists of 15 items, scored between 1 ('definitely does not apply') and 5 ('definitely applies') (see Appendix F). Colleagues at the University of Exeter undertook some work to assess the validity of the adapted measure. A total of 247 UK teachers completed the measure for two of their pupils, one with whom they had a good relationship, and one with whom their relationship was more strained. Responses were compared with validated measures that vary according to the quality of student-teacher relationship. The adapted measure demonstrated good external and discriminant validity. Some items are negatively worded and are therefore reverse scored. The total score has a range of 25-75, with higher scores reflecting a better relationship between the pupil and teacher.

Sample size

When the trial was designed and the protocol written, the sample size calculations suggested that an achieved sample of 140 schools would achieve a minimum detectible effect size (MDES) of 0.17 as shown in **Table 7**. The MDES and sample size calculations were calculated using a bespoke Excel spreadsheet developed by NFER. The ICC and pre and post correlation assumptions were drawn from a previous NFER trial using KS1 maths scores (Lord *et al.*, 2018). The recruitment target of 140 schools was only just missed (139), and the assumption of an average cluster size (42) was close to the actual average cluster size at randomisation (40.6) – see **Table 7**. Therefore, at randomisation the trial was still powered to detect an effect of 0.17 in the primary outcome. It is worth highlighting here that this MDES is based on the original primary outcome which was not possible to collect due to Covid-19. Due to the cancellation of the KS1 tests, the PBQ is being used as the main secondary outcome. It is theorised that attainment will be causally impacted by a change in pupil behaviour, therefore we were confident that, for the same number of schools, the trial was also powered to detect an effect as measured by the PBQ, a reliable measure of pupil behaviour (Allwood *et al*., 2018).

[9] http://www.sdqinfo.org/py/sdqinfo/c0.py

Randomisation

Randomisation was carried out by a statistician at NFER using a full SPSS syntax audit trail. Randomisation was conducted at school level, with region and school size used as stratifiers. Six target regions were used to facilitate training administration: The number of reception teachers per school was used as a proxy of school size. The schools were randomised 1:1 to intervention/control (business as usual). The analysts were not blinded to the randomisation result, however the code used is included in Appendix G. Baseline measurements were received from schools before they would be put forward to randomisation. Due to delays in receiving baseline data from some schools, randomisation took place in two batches. The same code was used for both batches. The follow-up measures were completed by the pupils' classroom teachers and therefore they were not blind to the group allocation.

Statistical analysis

Primary analysis

Due to the cancellation of Key Stage 1 assessments in summer 2021 there is no primary analysis for this trial. As described in the revised statistical analysis plan (SAP), once it was clear that collecting data on the primary outcome was not going to be possible, the evaluation team and the delivery team met to discuss possible options. It was agreed that the PBQ would be considered the main secondary outcome and any additional analyses planned for the primary outcome would be undertaken using this secondary outcome. After considering the causal chains implicit in the Theory of Change, it was felt that changes in low-level behaviour that should be picked up by the PBQ would be one of the initial possible changes observed in the classroom after IY®-TCM strategies are implemented

Secondary analysis

All secondary outcome analysis used an 'intention-to-treat' (ITT) approach. This means that pupils who had measurements at baseline and follow-up were included in the models, regardless of whether their teacher attended the IY®-TCM programme. The analysis for all multilevel models in this investigation was run in R using the 'lme4' package.

For the main secondary outcome of pupil behaviour, the model was a linear multilevel model with two levels (school and pupil). The dependent variable for the model was the PBQ total raw score at follow-up with the following covariates:

- Prior behaviour as measured by the baseline PBQ
- A series of dummy geographical variables (randomisation stratifiers)
- A series of dummy school size variables (proxy measure: number of reception teachers per school (randomisation stratifiers))
- Intervention allocation dummy variable

Two models were run on the secondary outcome of pupil emotional and social well-being as measured by the Total Difficulties Score (TDS) from the SDQ – one on a binary variable and the other using the TDS as a continuous scale as laid out in the SAP. A logistic regression was run using the cut-off score of 12 or above as a binary outcome, and a linear regression was run on the TDS raw score as a continuous outcome. The models used were multilevel models with two levels (school and pupil).

The dependent variable for the logistic regression model was a dummy variable representing a score of 12 or above on the TDS at follow-up with the following covariates:

- Baseline raw score of the TDS
- A series of dummy geographical variables (randomisation stratifiers)
- A series of dummy school size variables (proxy measure: number of reception teachers per school (randomisation stratifiers))
- Intervention allocation dummy variable

The dependent variable for the linear regression was the raw score of the TDS, with the same covariates as the above model.

The same linear regression models were run on the following secondary outcome measures: pupil concentration, prosocial behaviour and the STRS. Each model was assessed using a two-level multilevel model (school, pupil) with the following covariates:

- Baseline measurements of the respective scales (with the exception of the STRS where the baseline PBQ was used as the baseline measure)
- A series of dummy geographical variables (randomisation stratifiers)
- A series of dummy school size variables (proxy measure: number of reception teachers per school (randomisation stratifiers)
- Intervention allocation dummy variable

Analysis in the presence of non-compliance

The University of Exeter collected attendance data of teachers at each training session, using attendance logs. Compliance was agreed during the set-up period of the trial and described in the protocol (at teacher level) and the SAP (at pupil level). Compliance is a dichotomous variable whereby a teacher was considered compliant to the programme if they attended at least four of the six training sessions. At pupil level this translates into a pupil's Year 1 and Year 2 teacher both being compliant for the pupil to be considered compliant. In the case where the pupils have been taught by the same teacher across the two years, the compliance will apply to the one teacher only. If the students are taught by different teachers between Year 1 and Year 2, both teachers will have had to have attended at least four of the six training sessions.

As per EEF guidance a two-stage least squares model was used to calculate the Complier Average Causal Effect (CACE) estimate (Angrist and Imbens, 1995). This analysis seeks to compare outcomes for 'compliant' pupils in the intervention group with pupils in the control group who would have been 'compliant' if given the opportunity. The first stage of the model was compliance regressed on all covariates that are used in the main secondary outcome model and the group allocation variable. The second stage of the model regressed the main secondary outcome on the covariates used in the main model and included a covariate representing the teacher's estimated level of compliance from the first stage of the model. The coefficient of the estimated compliance measure is the CACE estimate of the compliance effect The R package 'ivpack' was used to perform the CACE analysis on the main secondary outcome only.

Missing data analysis

We assessed the level and pattern of missing data from the primary model. To assess the missingness mechanism, we ran a multilevel logistic regression model on whether a case had follow-up data for the main secondary outcome, regressed on the covariates of the same model plus other school level and available pupil level variables. As one covariate was found to significantly predict missingness, we ran the main secondary model with this extra variable included. As per EEF guidance, if the model estimates with and without the covariates had differed substantially then the outcome was probably missing not at random and sensitivity analysis would have been required using multiple imputation. As it was, the substantive model estimates with and without the extra covariate were similar, so this further stage did not take place.

Sub-group analyses

As FSM-eligible pupils represent a particularly important subgroup, a separate analysis of FSM-eligible pupils was carried out as per standard EEF practice. Sample size calculations run prior to the Covid-19 disruptions indicated that a sample of 1,411 FSM eligible pupils (the number randomised) would ensure enough power to detect an MDES of 0.2 on the primary outcome. At analysis stage we conducted ES calculations for the main secondary outcome instead due to cancellation of Key Stage 1 assessments in summer 2021. The subgroup model mimicked the model used to assess the main secondary outcome (PBQ), however with FSM pupils only. The MDES for the FSM subgroup at analysis stage was 0.31.

A further subgroup analyses was run on the whole sample, again using a two-level multilevel model (school and pupil) with PBQ scores as the outcome. In this model a dummy variable indicating pupil FSM eligibility was included as a covariate, as well as an interaction term between FSM eligibility and the intervention allocation dummy. The coefficient

of this interaction term therefore represents the additional (or reduced) benefit of the intervention for FSM pupils over non-FSM pupils. Otherwise, covariates in the model were the same as the main secondary analysis:

- Prior behaviour as measured by the PBQ
- A series of dummy geographical variables
- A series of dummy school size variables (proxy measure: number of reception teachers per school)
- Intervention allocation dummy variable
- Additionally, subgroup analysis was run for those pupils that had a SDQ Total Difficulties Score of 12 or more at baseline. Similarly, to the procedure for FSM above, first the main secondary model was rerun for this subgroup of pupils only. Then a multilevel model was run with PBQ score as the outcome and an interaction included between the intervention dummy and a dummy variable indicating whether pupils had a SDQ Total Difficulties Score of 12 or more at baseline. The other covariates in this model were the same as those listed above for FSM. This subgroup analysis was not specified in the statistical analysis plan and should be considered exploratory.

Additional analyses and robustness checks

While performing the secondary analysis it was noted that all follow-up measures had a distinctly non-Gaussian distribution, while model diagnostics indicated heteroscedasticity and non-Gaussian residuals (see Secondary Analysis section of results). In response to this a sensitivity analysis was conducted for each linear model in which the Huber-White standard errors (White, 1980) for the intervention coefficient were calculated and compared with the regular multilevel standard errors[10] used in the secondary analysis. Confidence intervals for standardised effect sizes, using both the multilevel and Huber-White standard errors, were also compared. Huber-White standard errors are often employed instead of multilevel models to correct for clustering of observations, but it is also possible to use them together with a multilevel framework to correct for other types of model misspecification (Wang and Merkle, 2018). Huber-White standard errors were calculated using the 'sandwich' and 'merDeriv' packages in R. As it transpired results were similar using the two methods, so no further analysis (e.g., Poisson regression) was required. This analysis was not specified in the statistical analysis plan and should be considered exploratory.

Estimation of effect sizes

Standardised effect sizes and their confidence intervals were calculated for all linear secondary outcome models. The numerator for the effect size calculation was the coefficient of the intervention group from the model. As these were two-level multilevel models, the denominator was the total variance from each multilevel model without covariates:

$$\mathrm{ES} = \frac{(\bar{Y}_T - \bar{Y}_C)_{\mathrm{adjusted}}}{\sqrt{\sigma_S^2 + \sigma_{\mathrm{error}}^2}}$$

In this equation σ_s^2 and $\sigma_{\mathrm{error}}^2$ represent the between-school and residual variance respectively. Confidence intervals for each effect size were derived by multiplying the standard error of the intervention group model coefficient by 1.96. These were converted to effect size confidence intervals using the same formula as the effect size itself. In the case where the outcome was a binary variable, an odds ratio was instead calculated.

Estimation of ICC

A two-level multilevel model of PBQ scores regressed on the intervention dummy variable was run without any covariates to estimate the ICC, which in this case was the proportion of the total variance accounted for by between-school differences. The ICC for the main model was also calculated from the output of the main secondary model, (including the covariates).

[10] That is, standard errors obtained from the covariance matrix provided as part of the default output of 'lme4'.

Longitudinal analysis

No longitudinal analysis was planned as part of this trial however there is the possibility of linking the data to the NPD and using assessment data from Key Stage 2 as a follow up measure. Using data in the NPD relating to exclusions and attendance may also be considered.

Implementation and process evaluation

Research methods

The implementation and process evaluation (IPE) is intended to complement the impact evaluation by providing information on how implementation affected the outcomes of the trial.

The IPE investigated the following research questions:

- IPE RQ1: To what extent was fidelity to the intervention maintained?
- IPE RQ2: How much of the intended intervention has been delivered?
- IPE RQ3: To what extent did teachers and pupils engage with the intervention?
- IPE RQ4: What does 'business as usual' consist of for the comparison group?
- IPE RQ5: What level and type of support did the developer team provide to intervention schools?
- IPE RQ6: Does the intervention have a perceived impact on teacher self-efficacy?
- IPE RQ7: Does the intervention have a perceived impact on teacher stress?

A range of data collection methods were utilised in order to answer the IPE research questions. An overview of the IPE methods presented in Table 5 below. See Appendix H for an overview showing how the IPE research questions and sources map on to different IPE dimensions.

Table 5. IPE methods overview

Research methods	Data collection methods	Participants/ data sources	Data analysis methods	Intended research questions addressed as planned	Intended research questions not addressed/ addressed partially	Implementation/ Theory of Change relevance
Surveys (pre / post)	Online baseline questionnaire	Y1 and Y2 teachers in intervention schools (N = 213) and control schools (N = 202)	Exploratory analysis	IPE RQ3 IPE RQ4 IPE RQ6 IPE RQ7		Context; usual practice
	Online endpoint questionnaire	Y1 and Y2 teachers in intervention schools (N = 101) and control schools (N = 69)	Exploratory analysis	IPE RQ3 IPE RQ4 IPE RQ6 IPE RQ7		Context; usual practice; responsiveness; fidelity, quality, cost
Document analysis	Attendance register	Training attendance registers	Descriptive analysis	IPE RQ2		Compliance; reach
Observations	Structured observation of IY®-TCM training	Observation of IY®-TCM training sessions (N=3)	Qualitative analysis	IPE RQ1 IPE RQ3 IPE RQ5		Fidelity; adaptation; responsiveness
	Structured CLASS observation of teaching practice	Y1 (n=4) and Y2 teachers (N=4) in 4 intervention case-study schools Baseline - October 2019 – completed Follow-up Y1 teachers Summer 2020 – not completed Follow-up Y2 teachers Summer 2021 - not completed	Descriptive analysis		IPE RQ2 IPE RQ3	Context; fidelity; adaptation; dosage; responsiveness; programme differentiation; quality

Interviews	In depth, semi-structured interview with the developers	Development team members (N=2)	Qualitative analysis	IPE RQ1 IPE RQ2 IPE RQ3 IPE RQ5		Context; fidelity; adaptation; responsiveness, cost
	In depth, semi-structured interviews with IY®-TCM trainers	IY®-TCM trainers (N=3)	Qualitative analysis	IPE RQ1 IPE RQ2 IPE RQ3 IPE RQ5		Context; fidelity; adaptation; responsiveness
	In depth, semi-structured interviews with Y1 and Y2 teachers in intervention case-study schools	Baseline: Y1 and Y2 teachers from 10 schools (N = 19) Midpoint: Y1 and Y2 teachers from 9 schools (N = 15) Interviews with Y2 teachers from 5 schools (N = 5)	Qualitative analysis	IPE RQ1 IPE RQ5 IPE RQ6 IPE RQ7	IPE RQ2 IPE RQ3	Context; fidelity; dosage; adaptation; responsiveness; programme differentiation; quality; cost

Data collection

Surveys

In September 2019, Year 1 and Year 2 teachers from control and intervention schools were asked to complete a baseline survey (online, see Appendix I). The survey explored: their school's existing policy and practice for behaviour management and current strategies for dealing with challenging behaviour; teacher's confidence in behaviour management; any previous CPD they had received related to behaviour management; teacher's self-efficacy and work-related stress. Overall, 415 Year 1 and Year 2 teachers (intervention N = 213 and control N = 202), completed a baseline survey.

In June 2021, all Year 1 and Year 2 teachers from control and intervention schools were asked to complete an endpoint survey (online). Most of the survey questions asked at baseline were repeated in the endpoint survey (see Appendix J). New questions were added for intervention group teachers to explore their engagement and responsiveness to the intervention; particularly take up of the techniques during and after the training and the associated costs of the intervention. The endpoint survey was also adapted slightly to capture the impacts of the Covid-19 pandemic on implementation. See Appendices J and K for the full survey instruments used. Overall, 170 Year 1 and Year 2 teachers (intervention N = 101 and control N = 69), completed an endpoint survey. In total, 83 intervention teachers and 57 control group teachers responded to both baseline and endpoint surveys.

Training attendance register

The IY®-TCM team offered training to all intervention schools between October 2019 and March 2020. This training was planned to be delivered as six whole-day sessions[11] (one per month) in six regional hub locations (18 TCM training courses were planned in total). All Year 1 and Year 2 teachers who had at least four days of classroom responsibility each week or were in a job share with both teachers were required to attend from each school. The developer collected a record of attendance at each training session; this data was collated and shared securely with NFER. Attendance data was analysed to measure compliance.

[11] The final training sessions in two areas had to be cancelled as the country had gone into a national lockdown due to the Covid-19 pandemic. We also noticed that whilst the other 14 courses were completed before the national lockdown, attendance at the last workshops was reduced, with many teachers saying they were unable to attend the final session due to having to cover staff absence.

Case studies

A series of case-studies were planned in order to explore implementation factors and the assumptions of the Theory of Change in greater depth. The research team adopted a purposive sampling approach to select ten case-study schools from the intervention group. The case studies were selected to provide some variation in characteristics such as: geographical location and school size (based on the number of Year 1 and 2 classes).

Six case-study schools were invited to take part in 'light-touch' telephone interviews with Year 1 and Year 2 teachers and four schools were invited to take part in school visits. The visits were intended to involve an observation of teaching in Year 1 and 2 classes (to provide rich, descriptive data on classroom climate, teacher efficacy and relationships between teachers and pupils) and interviews with the Year 1 and 2 teachers. The research team intended to conduct data collection with case-study schools at three different time points (baseline: before Year 1 and 2 teachers had received any training; midpoint: with Year 1 teachers in June 2020 while they were teaching the intervention pupils and endpoint: with Year 2 teachers in February 2021 when they were teaching the intervention pupils).

To gain a measure of any change over the two years of the evaluation and to bring the study in line with previous evaluations of IY®-TCM (for example, Murray *et al.*, 2017), researchers used the CLASS – the Classroom Assessment Scoring System[12] (Pianta *et al.*, 2008) to code the observations they conducted. Baseline data collection from case-study schools was completed as intended. However, due to Covid-19 restrictions, we were unable to carry out mid-point and endpoint school visits and interviews. Instead, we attempted to carry out follow-up interviews with all Year 1 and 2 teachers in the autumn term 2020, and with Year 2 teachers in summer term 2021. Some teachers were slow to respond to requests for follow-up interviews. We adapted our reminder strategy and increased the number of reminder phone calls and emails. We also offered greater flexibility around the timeline for interviews as well as an opportunity to provide written feedback to interview questions via email to facilitate participation. Despite this, we were unable to carry out follow-up interviews with all those involved at baseline.

Observations of IY®-TCM training

Observations of three training workshops (sessions 1, 4 and 6) in three different hub locations took place between November 2019 and March 2020. The three training sessions were selected to achieve some variation in terms of hub location, session number/focus and whether the trainer was a new or existing IY® trainer (i.e., delivering the training in pairs or individually). The observations focused on how IY®-TCM was delivered, whether delivery appeared to be consistent across different hub areas, what was delivered and how closely it aligned with the recommended practices of IY®-TCM, as well as indications of quality, effectiveness and teacher response.

Interviews with IY®-TCM trainers

Interviews with three IY®-TCM trainers who delivered in three hub locations took place in April 2020 (after the IY®-TCM training had been completed). In order to provide coverage across hub areas, the selected trainers were from hub locations which did not take part in the observations. The trainer interviews explored perceptions of the effectiveness of the training/intervention; perceptions of school or hub related differences; challenges encountered, and how these were overcome.

Interviews with developers

In May 2021 (towards the end of the trial), the research team conducted online interviews with two members of the University of Exeter developer team to support our understanding of delivery and the support provided to teachers as well as the costs of the intervention as it was delivered in the evaluation.

Data collection instruments

In addition to the use of CLASS (Pianta *et al.*, 2008) see details above, the research team developed the following data collection instruments: baseline and endpoint teacher survey, teacher, developer, and trainer interview schedules and IY®-TCM training observation schedule. The teacher survey included some standardised items on teacher self-efficacy

[12] 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.

(not the full scale) from the Teacher Self Efficacy Scale (Bandura, 1997) and on work-related stress (not the full scale) from the Teacher Concerns Inventory (Kyriacou, 2001). The developer also contributed to the design of the endpoint survey to ensure it captured data on the key components of the intervention. The research instruments were developed to meet the specific requirements of the trial, although they follow a similar format to others used in EEF trials conducted by NFER. The research team collected all data from surveys, interviews, and observations. The University of Exeter collected and collated the attendance register data. The registers contained pre-specified fields for the purposes of the trial. This information was shared securely with NFER.

Rationale for the data collection methods utilised and implications of Covid-19 related adaptions

NFER researchers chose the range of data collection methods outlined above as they offered both breadth and depth to the implementation and process evaluation. The ToC was used to help prioritise data collection to focus on the key features of the intervention and the assumptions underpinning it. The surveys provide an efficient way of measuring implementation and fidelity and inform our understanding of usual practice across a large number of schools. However, as these are self-reported measures, we also sought to gain observational evidence of the delivery. Combined with additional qualitative information gained from interviews with teachers, these data collection sources yielded further insight into the perceived impacts of IY®-TCM. As noted in the *Limitations* section, the IPE is limited by the small number of observations and interviews conducted and the relatively low response rate to the teacher endpoint survey and any insights should be viewed as tentative.

Analysis

The IPE data available for analysis is set out below.

- Observations of three IY®-TCM workshop sessions
- Interviews with three trainers after they had delivered a workshop session
- Observations of Year 1 and Year 2 teachers in four case-study intervention schools (prior to IY®-TCM training)
- Baseline Interviews with 19 Year 1 and Year 2 teachers from ten case-study schools (prior to IY®-TCM training)
- Follow-up interviews with 15 Year 1 and Year 2 teachers from nine case-study schools (around 6 months after the IY® training)
- End point interviews with five Year 2 teachers from five case-study schools (around 16 months after the IY® training)
- Baseline surveys from 415 Year 1 and Year 2 teachers in intervention schools (n=213) and control (n=202) schools.
- Endpoint survey from 170 Year 1 and Year 2 teachers in Intervention (n=101) and control (n=69) schools.
- Interviews with two members of the developer team towards the end of the trial.

The research team summarised qualitative data from observations and interviews with developers, trainers, and case study schools, in a grid representing data sources mapped against research questions (e.g., fidelity, quality of delivery, adaptation). The analysis used both inductive and deductive approaches (i.e., seeking to identify patterns in the data as well as to test the assumptions in the ToC to enable detailed analysis of the available data.

Survey responses were analysed to explore the school and teacher-level impacts specified in the theory of change, namely any differences between control and intervention group at endpoint in relation to teachers' work-related stress and any changes to school policies for behaviour management. Further analysis was carried out to explore the assumptions underpinning the intervention and the mechanisms of change related to three areas: perceived self-efficacy, causes of work-related stress and frustration with behaviour management among teachers. For each area individual item scores were summed to form a composite score, then improvement in these composite scores since baseline was compared between control and intervention teachers using Mann-Whitney U tests. Additional analysis was also conducted to compare the frequency with which the IY®-TCM training was used for different lengths of teaching

experience and for different levels of experience with difficult classroom behaviour. These comparisons were made using Fishers' exact tests rather than chi-squared tests due to low cell counts in many of the resulting cross-classification tables. Further details of how each analysis was conducted can be found with the relevant figures in the results section.

This additional analysis was exploratory, observational, and limited to participating schools, which is not a representative sample of the population. It was not appropriate to break the responses down by sub-group (e.g., by year group taught or training hub area) due to the small numbers involved. Findings are reported at overall level only, to avoid the risk of identifying specific schools or individuals.

Quantitative data from teacher surveys was analysed using R. This analysis was conducted with a full audit trail and quality-assured by a senior statistician at NFER. The attendance data was collated and summarised using Microsoft Excel.

The research team triangulated the IPE qualitative and qualitative data sources in order to cross-validate the results and support the interpretation of the findings.

Costs

The cost of programme delivery was determined from the perspective of the school. Information was 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.

Data relating to costs borne by school (in terms of time as well as financial costs) were collected via questions included in the endpoint practitioner survey and supplemented with interviews with intervention teachers and the development team. As the programme was funded for intervention schools by the EEF and the University of Exeter, further cost information was sought from the developer team. Business as usual cost data from control schools were collected via the endpoint practitioner survey.

Costs were calculated as a cost per pupil from the school's perspective, as if schools were paying for the intervention, based on additional costs – i.e., those costs above normal spend (in terms of time and financial cost). We considered additional 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 also report 'time' in terms of the amount of hours spent by staff and any re-allocation of existing resources. Costs per pupil were 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 was estimated per school year, and then over multiple years (up to three years).

Timeline

The timeline is included in Table 6. Several instruments/activities were not possible due to Covid-19 and the related disruption to school. These elements were:

- midpoint data collection of PBQ and SDQ (due summer 2020),
- midpoint and endpoint CLASS observations, and
- Key Stage 1 (and EYFSP) data collection from NPD.

Table 6: 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, CS
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	CLASS observations (baseline – Year 1 and Year 2 intervention group teachers)	GS
Oct 2019 - Mar 2020	Teacher training workshops (six in total, one per month) IPE case studies and telephone interviews	RH/TF KM, EB
Mar-Jun 2020	Schools closed to the majority of pupils due to Covid-19	(schools)
Jun 2020	Teachers complete midpoint online SDQ and PBQ for all pupils [UPDATE: cancelled due to Covid-19] CLASS observations (midpoint - Year 1 teachers only) [UPDATE: cancelled due to Covid-19]	SR, KH KM, EB
Sep 2020	Year two of intervention commences (trial pupils begin Year 2) IPE telephone interviews begin	KM, EB
Nov 2020	Data collection from school regarding teacher names for trial pupils (to match pupils to their teacher)	KH
Jan-Feb 2021	Schools closed to the majority of pupils due to Covid-19	(schools)
Feb 2021	CLASS observations (endpoint - Year 2 teachers only) [UPDATE: cancelled due to Covid-19]	KM, EB
Jun 2021	Pupils sit Key Stage 1 assessments [UPDATE: cancelled due to Covid-19]	(schools)
Jun-Aug 2021	Teachers undertake SDQ and PBQ online endpoint surveys, revised Student Teacher Relationship Scale (STRS), and endpoint practitioner survey IPE telephone interviews	ST, KH
Sept-Nov 2021	Analysis of all outcomes for Cohort 1 Year 2 pupils, report writing.	CR, BS, ST, CS
Dec 2021	First draft of the report to EEF	ST
Jan-Mar 2022	Peer review, comments, and adjustments to the report	All
Summer 2022	Final report published	EEF

Impact evaluation results

Participant flow including losses and exclusions

Figure 2: Participant flow diagram (2 arms)

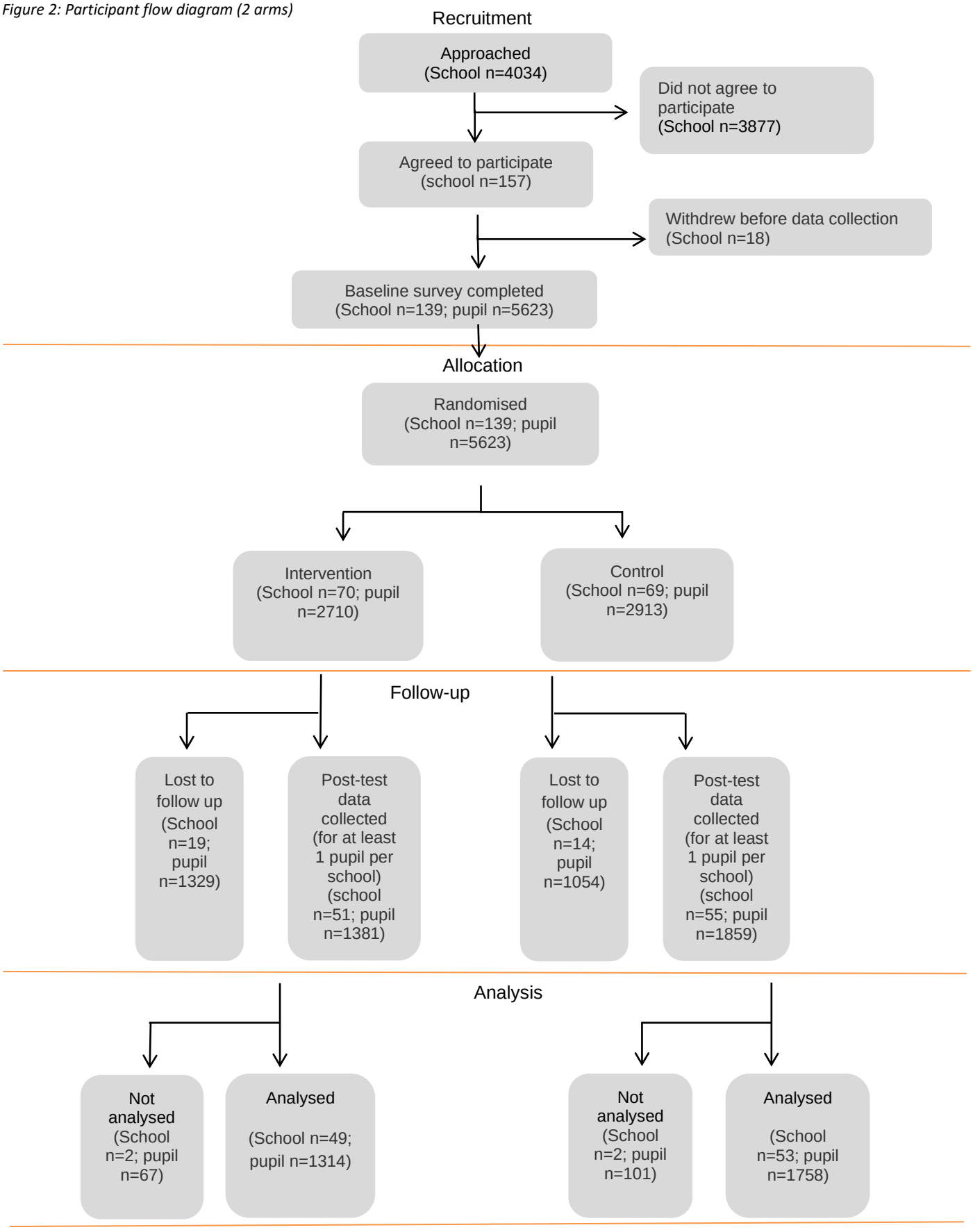


Figure 2 shows the participant flow for the main secondary outcome (PBQ). The developer team approached 4,034 schools to take part in the evaluation (as described in **Participant selection** section). Of the 33 schools lost between randomisation and endpoint data collection (19 intervention, 14 control), nine withdrew during the delivery of the intervention training (nine intervention, two control). Reasons were not given for the majority of withdrawals, but time pressures and staff changes were stated as reasons provided for a handful of schools. The proportion of endpoint surveys returned by control and intervention schools can be viewed graphically in Appendix K.

Pupils could only be included in the main secondary analysis (those 'analysed' in the study flow diagram) if teachers answered the survey questions that allowed their PBQ outcome to be calculated. N=67 intervention pupils and N=101 control pupils, despite returning their follow-up surveys, did not answer these questions and so their data was not included in the main secondary analysis. A further two control and two intervention schools had no pupils included in the main secondary analysis for this reason, despite returning at least one follow-up survey.

Table 7: Minimum detectable effect size at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
Outcome measure		KS1 maths score		KS1 maths score		Pupil Behaviour Questionnaire (PBQ) score	
MDES		0.17	0.20	0.17	0.20	0.21	0.30
Pre-test/post-test correlations	Level 1 (pupil)	0.51	0.51	0.51	0.51	0.51	0.53
	Level 2 (school)	-	-	-	-	-	-
Intracluster correlations (ICCs)	Level 2 (class)	-	-	-	-	-	-
	Level 2 (school)	0.15	0.15	0.15	0.15	0.16	0.16
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		Two	Two	Two	Two	Two	Two
Average cluster size		42	10.5	40.6	10.2	30.1	5.8
Number of schools	Intervention	70	70	70	70	49	39
	Control	70	70	69	69	53	42
	Total:	140	140	139	139	102	81*
Number of pupils	Intervention	2940	735	2710	680	1314	208
	Control	2940	735	2913	731	1758	258
	Total:	5880	1470	5623	1411	3072	466

Note: * This is the number of schools that reported having at least one FSM eligible pupil.

The parameters involved in the calculation of the minimum detectable effect size (MDES) at the protocol, randomisation and analysis stages are displayed in Table 7. Pre-test/post-test correlations remain similar to those anticipated at the protocol and randomisation stages, although due to the change in main outcome at the analysis stage, there is no reason in principle why this should be the case. More importantly, there has been substantial loss to follow-up at both the school and pupil level, leading to an increase in the MDES that can be detected with a power of 80 per cent from 0.17 to 0.21. The proportion of pupils eligible for FSM has also been lower than anticipated at the protocol and randomisation stages and accordingly the MDES for this group has increased by more than for pupils overall.

It should be noted that due to Covid-19 the main study outcome has changed from KS1 maths scores to PBQ scores and how likely a true effect size of 0.21 or more is should be considered in this context. It is perhaps reasonable to assume a larger effect size for PBQ as it is a more proximal measure to the intervention. However, it is less easy to predict how much larger it might be.

Attrition

There was a high level of attrition in this trial with 45 per cent of the pupils randomised not being analysed with control pupils having a lower rate of attrition compared to intervention. The attrition at the level of the school was lower (27%) as can be seen from the **participant flow diagram**. Some school-level attrition happened during the intervention training phase particularly in the intervention group. Much of the attrition was due to lack of response at the endpoint follow up. Between the autumn term data collection (autumn 2020)[13] and endpoint data submissions there was a loss of 28 per cent of pupils in the intervention group and 25 per cent of pupils in the control group. At school level there were 12 more schools in the control group who provided at least 76 per cent of their pupil surveys at endpoint as compared to intervention (see Appendix K). It is worth noting that the majority of the project was run during a period of uncertainty for schools as they responded to the challenges of operating during the Covid-19 pandemic. The response rate at endpoint was lower than expected before the Covid-19 pandemic and appears to be largely due to the timing of the survey and the rise in Covid cases. Schools reported that they were struggling to complete the survey due to teachers being in isolation or the school closing early.

Table 8: Pupil level attrition from the trial (main secondary outcome)

		Intervention	Control	Total
Number of pupils	Randomised	2710	2913	5623
	Analysed	1314	1758	3072
Pupil attrition (from randomisation to analysis)	Number	1396	1155	2551
	Percentage	51.5%	39.6%	45.4%

Pupil and school characteristics

Table 9 below displays the baseline school and pupil-level characteristics of the intervention and control groups at the randomisation stage. School-level characteristics were obtained from NFER's own Register of Schools dataset, while pupil-level characteristics were provided directly by the schools. The population used for calculating the national-level Key Stage 2 average maths scores was English primary schools, not including special or independent schools.

The distribution of baseline characteristics is generally similar between the groups. The possible exception is baseline PBQ and TDS values, which are higher in the intervention group, with effect sizes of 0.10 (95% CI: -0.01, 0.22) and 0.12 (-0.01, 0.25) respectively. Due to schools being randomised, these differences will have occurred due to chance. The impact of the difference in baseline PBQ and TDS scores was minimised by their inclusion as covariates in the secondary

[13] It was necessary to include an additional data collection point in autumn term 2020 after the cancellation of the midpoint assessment, in order to find out from schools which teacher was teaching the trial pupils in Year 2 (and confirm who taught them in Year 1).

analysis models. There was no evidence of imbalance in baseline PBQ values at the analysis stage: the effect size for this difference was -0.04 (-0.18, 0.09). Plots displaying the distribution of baseline PBQ scores in the intervention and control groups can be found in Appendix L.

Table 9: Baseline characteristics of groups as randomised

School-level (categorical)	National-level mean	Intervention group		Control group		
		n/N (missing)	%	n/N (missing)	%	
Training region						
Bristol		13/70 (0)	18.6%	14/69 (0)	20.3%	
Cornwall		11/70 (0)	15.7%	11/69 (0)	15.9%	
Dorset		13/70 (0)	18.6%	12/69 (0)	17.4%	
Liverpool		14/70 (0)	20%	14/69 (0)	20.3%	
Reading		13/70 (0)	18.6%	12/69 (0)	17.4%	
Hampshire		6/70 (0)	8.6%	6/69 (0)	8.7%	
N reception teachers						
1		36/70 (0)	51.4%	34/69 (0)	49.3%	
2		30/70 (0)	42.9%	26/69 (0)	37.7%	
3		3/70 (0)	4.3%	7/69 (0)	10.1%	
4		0/70 (0)	0%	1/69 (0)	1.4%	
5		1/70 (0)	1.4%	1/69 (0)	1.4%	
% FSM eligibility national quintile 2018/19						
Lowest 20%		13/70 (0)	18.6%	9/69 (1)	13%	
2nd lowest 20%		11/70 (0)	15.7%	15/69 (1)	21.7%	
Middle 20%		17/70 (0)	24.3%	19/69 (1)	27.5%	
2nd highest 20%		18/70 (0)	25.7%	12/69 (1)	17.4%	
Highest 20%		11/70 (0)	15.7%	13/69 (1)	18.8%	
School type						
Academy/free school		34/70 (0)	48.6%	23/69 (0)	33.3%	
Maintained		36/70 (0)	51.4%	46/69 (0)	66.7%	
Overall Ofsted rating 2018/19						
Outstanding		12/70 (3)	17.1%	8/69 (6)	11.6%	
Good		41/70 (3)	58.6%	47/69 (6)	68.1%	
Requires improvement		10/70 (3)	14.3%	7/69 (6)	10.1%	
Inadequate		4/70 (3)	5.7%	1/69 (6)	1.4%	
School-level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	
KS2 average scaled score in maths 2018/19	104.4	58/70 (12)	103.7 (3)	53/69 (16)	103.8 (2.9)	
Pupil-level (categorical)		n/N (missing)	%	n/N (missing)	%	
FSM eligibility						
No		2215/2687 (0)	82.4%	2484/2896 (0)	85.8%	
Yes		472/2687 (0)	17.6%	412/2896 (0)	14.2%	
Pupil-level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Effect size
Baseline PBQ score		2687/2687 (0)	1.7 (2.6)	2896/2896 (0)	1.4 (2.3)	0.10 (-0.01, 0.22)
Baseline TDS		2687/2687 (0)	7.1 (6.1)	2896/2896 (0)	6.2 (5.7)	0.12 (-0.01, 0.25)

Notes: 'Percentages of total' not 'valid percentages' are displayed, so where there is missing data, percentages do not sum to 100. Due to rounding some percentages may not sum to exactly 100.

Outcomes and analysis

Secondary analysis

The distribution of each secondary outcome for the intervention and control group is plotted in Appendix M. It can be seen that these distributions are distinctly non-Gaussian: in particular, they have many zeros (or the maximum value for scales where larger numbers are positive). It is not itself a modelling assumption that the dependent variables have a Gaussian distribution, but substantial deviations from normality can cause lead to violation of model assumptions. This can be seen in Appendix N, where residuals for the linear models are plotted against fitted values. Although these plots can be harder to interpret for what is ordered categorical data, there is evidence of heteroscedasticity (unequal variance across variable levels): the vertical spread of the points widens while moving along the x axis. There is also evidence that model residuals are non-Gaussian, as shown by an asymmetric vertical spread of point on each side of the x axis. In response to this a sensitivity analysis was performed by calculating Huber-White standard errors (see additional analysis section below).

The results for the main secondary analysis can be seen in Table 10. The effect size of the intervention for the PBQ outcome was -0.04 (95% CI: -0.20, 0.12). As zero is contained within the confidence interval for this estimate, it is not possible to reject the null hypothesis here. That is, this study provides no evidence that the IY®-TCM training programme improves PBQ scores among KS1 pupils within an intention to treat (ITT) framework. The proportion of variance in the PBQ outcome attributable to differences between schools (the ICC) is 0.16 for this model. For the model with only the intervention variable with no covariates the ICC is 0.09.

Results for other secondary analysis are also displayed in Table 10. For all outcomes zero is contained within the CI so there is too much uncertainty around the estimates to conclude that the intervention has an effect. Recalling that higher scores can be positive (SDQ Prosocial, STRS) or negative (PBQ, SDQ TDS, SDQ Hyperactivity) depending on the measure, only the estimates for PBQ and SDQ Hyperactivity are in the direction corresponding to a beneficial effect (had the confidence intervals been narrower). The components used to calculate the effect size for all secondary outcomes can be found in Appendix C.

As well as being treated as continuous (Table 10), SDQ TDS was also dichotomised at a score of 12 or more and included as the outcome in a multi-level logistic regression, the results for which are in Table 11. An odds ratio of 1.10 (0.62, 1.94) means that the odds of a pupil having a SDQ TDS score of 12 or more is 10 per cent greater in the intervention group. Again, this is probably due to chance rather than a true effect, as the confidence interval for the odds ratio contains one.

Table 10: Main secondary analysis (PBQ outcome) and other secondary analyses

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	N (missing)	Mean (95% CI)	Total n (intervention; control)	Standardised effect size (95% CI)	p-value
Pupil Behaviour Questionnaire	1314 (1373)	1.42 (1.29, 1.55)	1758 (1138)	1.58 (1.46, 1.70)	3072 (1314; 1758)	-0.04 (-0.20, 0.12)	0.60
SDQ Total Difficulties Score	1315 (1372)	6.45 (6.09, 6.80)	1759 (1137)	6.35 (6.07, 6.63)	3074 (1315; 1759)	0.05 (-0.16, 0.26)	0.66
SDQ Hyperactivity Scale	1315 (1372)	2.77 (2.61, 2.93)	1759 (1137)	2.98 (2.84, 3.13)	3074 (1315; 1759)	-0.07 (-0.22, 0.09)	0.40
SDQ Prosocial Scale	1315 (1372)	8.18 (8.05, 8.30)	1759 (1137)	8.19 (8.08, 8.30)	3074 (1315; 1759)	-0.06 (-0.26, 0.14)	0.56

Student Teacher Relationship Scale (Revised version)	1310 (1377)	67.14 (66.68, 67.59)	1756 (1140)	66.69 (66.29, 67.10)	3066 (1310; 1756)	-0.06 (-0.27, 0.15)	0.60

Table 11: Dichotomising SDQ TDS secondary outcome at a score of 12 or greater

	Unadjusted proportion with 12+ score				Intervention/control odds ratio		
	Intervention group		Control group				
Outcome	N (missing)	Proportion (95% CI)	N (missing)	Proportion (95% CI)	Total n (intervention; control)	Odds ratio (95% CI)	p-value
SDQ Total Difficulties Score of 12+	1315 (1372)	0.190 (0.169, 0.211)	1759 (1137)	0.173 (0.155, 0.190)	3074 (1315; 1759)	1.10 (0.62, 1.94)	0.75

Analysis in the presence of non-compliance

The compliance requirement for pupils in this study was that both their Year 1 and Year 2 teacher attended a minimum of four out of six IY®-TCM training sessions. Of the 16 training groups being run, two had their final session cancelled due to Covid-19, which was treated as non-attendance of that session by all teachers in that training group. Establishing whether each pupil's Year 1 and Year 2 teacher had attended four training sessions relied on Year 1 and Year 2 teacher name information provided by schools. Where this information was missing or illegible, that pupil's compliance variable was treated as missing. Where more than one teacher name was given for a pupil for a particular year then both teachers had to attend four sessions for that year's compliance requirement to be met (if only one teacher attended four sessions this was treated as a missing compliance variable)[14]. Twenty-five intervention pupils from two classes had missing compliance information due to one of the factors described above. All control pupils were treated as not meeting the compliance requirements, regardless of missing or unclear teacher names.

There was a mean of 1.63 different Year 1 teachers (range 1-5) amongst all schools; 1.57 (1-5) at intervention schools and 1.70 (1-5) at control schools. For Year 2 teachers the mean was 1.60; 1.49 (1-3) at intervention schools and 1.69 (1-5) at control schools[15].

In Year 1, N=1066 (78%) of intervention pupils included in the analysis had a teacher that attended four or more sessions; in Year 2 the figure was N=798 (58%). Of the Year 1 pupils, N=106 (8%) had a teacher that was not listed in the training register at all; for Year 2 pupils this was higher at N=292 (21%). The percentage for Year 2 teachers who are not on the training register is quite high and this is likely to be at least in part due to the training register being compiled using the teacher names provided by the schools based on their prediction of which teachers would be taking Year 2 classes the following academic year, which could be incorrect due to teachers leaving the school or moving year groups. Altogether N=731 pupils (54%) met both the Year 1 and Year 2 requirements and therefore the overall compliance requirement.

A CACE analysis was performed using complete cases only (no missing PBQ outcome or compliance information), the results of which are below (Table 12). The F-statistic for stage one of the two-stage least squares regression was 289 (p<0.001) and the correlation between intervention status and the compliance indicator was 0.63. It can be seen that pupil compliance has an effect size of -0.13 (-0.24, -0.01) for the outcome of PBQ score, with the entire confidence interval lying below zero. This indicates that the IY®-TCM programme may have a small beneficial effect on PBQ scores amongst those pupils whose teachers attend most of the training.

Table 12: CACE analysis for the main secondary outcome

[14] This only occurred for one class at baseline and one class at endpoint. It would be preferable to consider how many days of teaching each teacher did per week when deciding overall compliance, but this level of detail was not present in the data.
[15] The Year 2 figures are based on pupils with returned endpoint surveys only.

Model stage	Total n (intervention; control)	Predictor	Standardised effect size (95% CI)	p-value
Stage 1: compliance indicator regressed on intervention status*	3050 (1292; 1758)	Intervention status	1.23 (1.18, 1.28)	<0.01
Stage 2: PBQ score regressed on compliance indicator*	3050 (1292; 1758)	Compliance indicator	-0.13 (-0.24, -0.01)	0.03

*Additionally baseline PBQ score, and the geographical and school size variables were included at both stages.

Missing data analysis

There was a high level of missing data for the secondary outcomes of this study: PBQ (N=2511, 45%), all SDQ-based scores (N=2509, 45%) and STRS (N=2517, 45%). This was largely due to complete loss of follow-up, rather than particular survey items being missing: Year 2 surveys were not returned for 1306 intervention and 1037 control pupils. All of the covariates included in the secondary outcome models (baseline PBQ and SDQ TDS, training region, number of reception teachers) had complete data.

In order to assess which variables might be associated with the missing data mechanism for the PBQ outcome, a multilevel logistic regression was performed in which the outcome was PBQ follow-up (1= 'yes', 0= 'no'). Results for covariates that were included in the regression, but which were not included in the main secondary outcome model, are shown in Table 13. There was evidence that baseline SDQ TDS scores may be associated with the missing data mechanism for the PBQ outcome (p=0.02), so this variable was taken forward to the sensitivity analysis.

As a sensitivity analysis, baseline SDQ TDS scores were included as a covariate in the main secondary outcome model, which was left otherwise unchanged. This resulted in a slight increase in the absolute effect size estimate, from -0.04 (-0.20, 0.12) to -0.06 (-0.21, 0.10). However, the substantive conclusion of the analysis remains the same: no evidence of an effect of the IY®-TCM training on PBQ scores. This suggests that the absence of the variables in Table 13 from the main secondary analysis has not been a cause of substantial bias in the estimation of the intervention's effect on PBQ. However, given the quantity of missing outcome data, as well as the imbalance between control and intervention pupils, study results should still be interpreted with caution.

Table 13: Multilevel logistic regression where the outcome is follow-up for the PBQ variable (1= 'yes', 0= 'no')

Covariate[16]	N (%) covariate missing	Odds ratio (95% CI)	p-value
Male gender	0 (0)	1.14 (0.93, 1.40)	0.19
Pupil FSM eligibility	0 (0)	0.98 (0.75, 1.29)	0.90
Baseline SDQ Total Difficulties Score	0 (0)	0.97 (0.95, 0.99)	0.02
School % FSM eligibility 2018/19	32 (0.6)	0.95 (0.89, 1.01)	0.10
School Ofsted rating of 'good' or 'outstanding' 2018/19	358 (6.4)	1.88 (0.33, 10.65)	0.48
All Y1 teachers at school attended 4+ training sessions	448 (8.0)	3.01 (0.41, 22.12)	0.28

[16] Only covariates that were not included in the main secondary outcome analysis are displayed here, although intervention status, baseline PBQ score, training region and number of reception teachers were included in this regression. Number of reception teachers was treated as a continuous variable due to model convergence issues caused by leaving it as categorical.

Sub-group analyses

The main secondary analysis was performed again, restricted this time to the subset of pupils that were eligible for FSM. The results for this analysis are shown in Table 14 and they are similar to those for the main secondary analysis, except that confidence intervals are wider due to the smaller sample size. Whether there is a differential effect of the IY®-TCM training on PBQ scores for FSM pupils was tested by including an interaction between the intervention indicator variable (1= 'intervention', 0= 'control') and a pupil-level FSM indicator (1= 'yes', 0= 'no') in the main secondary outcome model. The estimated adjusted mean difference for a FSM pupil taking part in the intervention versus a non-FSM pupil is -0.03 (-0.45, 0.39; p=0.90), showing no evidence of a differential effect for FSM pupils.

The main secondary analysis was also performed for the subgroup of pupils with a SDQ Total Difficulties Score of 12 or more at baseline, that is, pupils struggling with their emotional and social wellbeing before the intervention began. This analysis was not specified in the statistical analysis plan and should be considered exploratory. The effect size estimate associated with the intervention in this subgroup was again fairly similar to the main secondary analysis, with no evidence of an effect (see Table 14). Similarly, to the FSM analysis above, another model was run in which an interaction between the intervention indicator and an indicator for a SDQ Total Difficulties Score of 12 or more was added to the main secondary analysis model. In this case the adjusted mean difference between a pupil at an intervention school with a baseline Total Difficulties Score of 12 or above versus one with a score below 12 is 0.05 (-0.34, 0.45, p=0.79).

Table 14: The main secondary analysis model (PBQ outcome) restricted to particular subgroups of pupils

	Unadjusted means				Effect size		
	Intervention group		Control group				
Subgroup	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Standardised effect size (95% CI)	p-value
FSM-eligible pupils	208 (264)	2.06 (1.68, 2.44)	258 (154)	2.10 (1.75, 2.45)	466 (208; 258)	-0.06 (-0.29, 0.18)	0.62
Pupils with a SDQ Total Difficulties Score of 12+ at baseline	239 (309)	3.38 (2.94, 3.83)	289 (190)	3.58 (3.17, 4.00)	528 (239, 289)	-0.10 (-0.36, 0.16)	0.44

Additional analyses and robustness checks

In order to assess whether inferences made in the secondary analysis section are reliable in the presence of model assumption violations (non-Gaussian residuals, heteroscedasticity), Huber-White standard errors were calculated for each model (Table 15). Standard errors shrink slightly for all secondary outcomes when using the Huber-White method[17] and consequently the confidence intervals for the effect sizes become narrower. However, the substantive findings remain the same for all secondary outcomes: that there is no evidence of an effect of the IY®-TCM training on non-cognitive pupil outcomes.

[17] Generally, the Huber-White standard errors will be higher than conventional estimates, but it is possible for them to be lower.

Table 15: standard errors and resulting confidence intervals for the intervention coefficient in all linear secondary analysis models. Standard errors from the secondary analysis are compared with those obtained by the Huber-White method.

Outcome	Standard error of intervention coefficient		Standardised effect size 95% CI	
	Multilevel estimate	Huber-White estimate	Multilevel estimate	Huber-White estimate
Pupil Behaviour Questionnaire	0.203	0.187	(-0.200, 0.115)	(-0.188, 0.103)
SDQ Total Difficulties Score	0.695	0.675	(-0.164, 0.261)	(-0.158, 0.255)
SDQ Hyperactivity Scale	0.242	0.228	(-0.224, 0.088)	(-0.215, 0.079)
SDQ Prosocial Scale	0.238	0.235	(-0.261, 0.142)	(-0.258, 0.138)
Student Teacher Relationship Scale (revised version)	0.930	0.908	(-0.266, 0.153)	(-0.261, 0.148)

Implementation and process evaluation results

Key findings

Reflections on compliance

The IY®-TCM training consisted of six full-day workshop sessions spread across October 2019 – March 2020. Compliance to the intervention is indicated by attendance of at least at four out of the six sessions. Attendance registers from each of the 18 training-group sessions were used to measure how much of the intervention Year 1 and Year 2 teachers participated in.

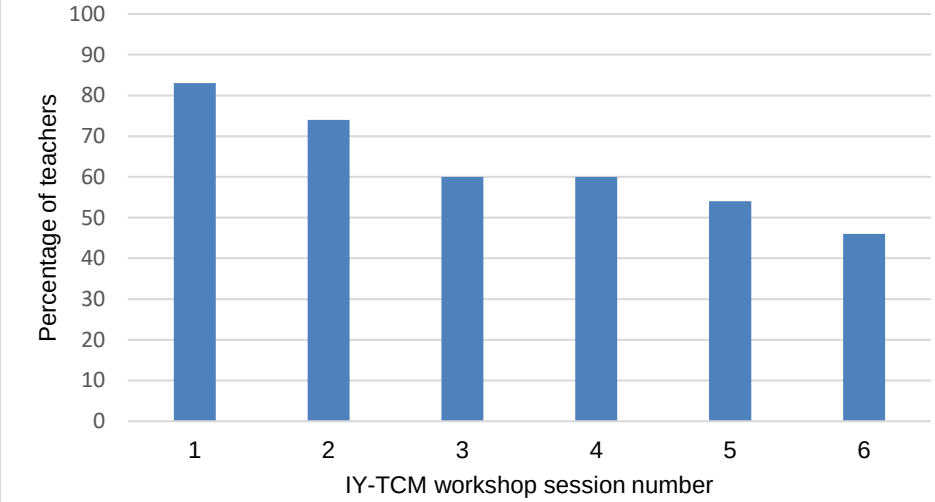
Teacher attendance at the IY®-TCM training

At the teacher level, compliance was achieved by 60 per cent[18] of teachers randomised to the intervention group (attending at least four of the six sessions). Thirty-five per cent of teachers attended all six sessions and twelve per cent did not attend any sessions[19].

Two of the IY®-TCM training groups were cancelled after the first workshop due to low numbers of attendees. The remaining teachers were given places at other groups.

Figure 3 shows the distribution of teachers assigned to the intervention who attended each of the workshop sessions. Attendance reduced gradually over time. By workshop three, teacher attendance had reduced by around one quarter. Just under half of teachers (46%) attended the final workshop session.

Figure 3: Distribution of teachers who attended each IY-TCM workshop session



Interviews with the IY®-TCM developer team, IY®-TCM trainers and with a selection of Year 1 and Year 2 teachers from case-study schools offer insight into the reasons why teachers did not attend the training.

One of the key issues was the cost to schools of releasing so many teachers to attend the training. Schools were required to release all Year 1 and Year 2 teachers for six full-day training sessions over the training period (October 2019 – March 2020) however they did not receive any funding for supply cover as part of the trial. The developer said

[18] The percentage reported here is different to the pupil-level compliance measure (54%) reported in the outcomes and analysis section as this is the teacher-level compliance. Sixty per cent of those teachers expected to attend the training, attended at least four out of the six workshops. Pupil-level compliance is measured at the level of the pupil rather than the teacher and requires that the pupil was taught in Year 1 and Year 2 by a teacher who attended at least four workshops.
[19] Teachers who did not attend any of the training are included in the compliance results as part of the 'intention to treat-treat' analysis.

that these costs could be onerous for schools: *'It's six days out of class and* [involves] *all Year 1 and Year 2 teachers – that is a massive amount of money.'*

Some schools found the practicalities of releasing all Year 1 and Year 2 teachers for six days over the course of the training period a barrier to attending, meaning that along with the financial challenges, they considered the cost-benefit of all their eligible teachers attending the sessions. Although schools were able to send teachers to courses run on different days to reduce the burden on the school, some schools that could not release all (eligible) teachers attempted to address this by nominating one teacher to attend the training on behalf of the school or through teachers attending alternate sessions. The developer team and trainers discouraged this, highlighting the importance of teachers attending the full course. Trainers emphasised to participants that the IY®-TCM training is incremental in nature, building on the fundamental principles of the approach and providing teachers with more knowledge and strategies each session. They said that teachers who attended all six sessions had a better understanding of how the programme content fitted together and gained more from their attendance, compared to those who were unable to participate fully in the training. However, in cases where teachers missed a session, the trainers permitted attending teachers from the same school to feed back to their colleagues ahead of their attendance at the next workshop.

Another issue was the relationship between the approaches advocated in the training and schools' existing behaviour policies. Trainers reported that during the workshops, a few of the participating teachers disclosed that the headteacher of their school was unwilling to change current policies and practice on behaviour management. In some instances, this had negatively impacted upon attendance because teachers did not feel they would be able to put the training and strategies learned into practice back in the classroom.

A very small number of teachers who participated in the case-study interviews said that they did not find the training they had attended a valuable use of their time, which therefore impacted their attendance at later training sessions. Their reasons for feeling this way are described in the '*Usual practice'* and *'Quality of delivery'* section of this report.

Implications of Covid-19 upon teachers' attendance at the training

The final IY®-TCM workshop (session 6) took place in March 2020 when, towards the end of the month (23rd March), schools closed to the majority of their pupils as England went into national lockdown. Two of the 16 training groups had their sixth and final training session cancelled due to partial school closures. One of these sessions was due to take place on 19th March and the other on 24th March. However, the developer and trainers reported lower attendance at the sixth sessions in areas where they did take place, due pressures on school staff capacity at that time.

Fidelity

To what extent does the delivery of the training workshops adhere to the intended approach/practices advocated by IY®-TCM?

To ascertain fidelity, the evaluation team carried out observations of three training sessions (in different hub areas) and interviews with trainers and the developer. The findings from these observations and interviews suggest that trainers delivered the programme content with high fidelity to the recommended approach.

The IY®-TCM programme is highly manualised with clear criteria for training and the materials and manuals provided as part of the intervention are very structured, which supports delivery with fidelity at all levels. Nevertheless, it allows for 'adaptation with fidelity', meaning that trainers can give greater focus to some elements of the training over others, depending upon where teachers require the most support. The developer team reported that they had some initial concerns that in some cases delivery was 'too rigid'. For example, they highlighted that even though the trainers' manual[20] details the vignettes they should present for each topic covered in a session, if teachers demonstrate they understand the concept, it would not be necessary to present them all – yet some trainers had continued to do so. The developer reported having conversations with these trainers to resolve this issue and said that overall, trainers kept the sessions on track and ensured that teachers received all of the training content, for example by addressing any other issues teachers raised in the breaks. After each workshop, trainers completed standardised agendas and checklists of

[20] This is the Leader's Manual as described on the Incredible Years website: **https://incredibleyears.com/programs/teacher/classroom-mgt-curriculum/**.

the expected curriculum that they should have covered during the session. The developer reviewed these and confirmed that there were no major differences in the coverage of content across the different training groups, further supporting that the training was delivered consistently and with high fidelity across the six hub areas. Any changes that trainers reported making to their delivery of the training sessions can be found in the *'Adaptations'* section of this report.

One of the strategies for ensuring successful implementation of the training related to the recruitment and selection of experienced trainers to deliver the sessions and providing new trainers with intensive training. The evaluation team asked the developer to comment upon their recruitment strategy and how they prepared new trainers to deliver the programme, to ensure standardised delivery. Of the 12 trainers involved in the trial, six had delivered IY®-TCM previously, three had delivered another IY® programme in the past (e.g., the parenting programme) and three trainers were new to IY®. The developer team reported that they carefully selected new trainers based on their levels of prior training and classroom experience because this was key to engaging with teachers and facilitating the training. The six new trainers, and three experienced trainers who requested refresher training, received training over three days from an accredited IY® trainer. This training provided them with an understanding of the theories behind IY®-TCM, detail on the content of the six sessions and an opportunity to practice delivering in small groups. The developer also reported that trainers received a script to follow for each session, although it was not a specific requirement that they used it. In addition to this training support, new trainers delivered the programme alongside an experienced trainer who provided supervision.

Dosage

Which techniques, and how frequently, were teachers implementing the IY®-TCM in the classroom?

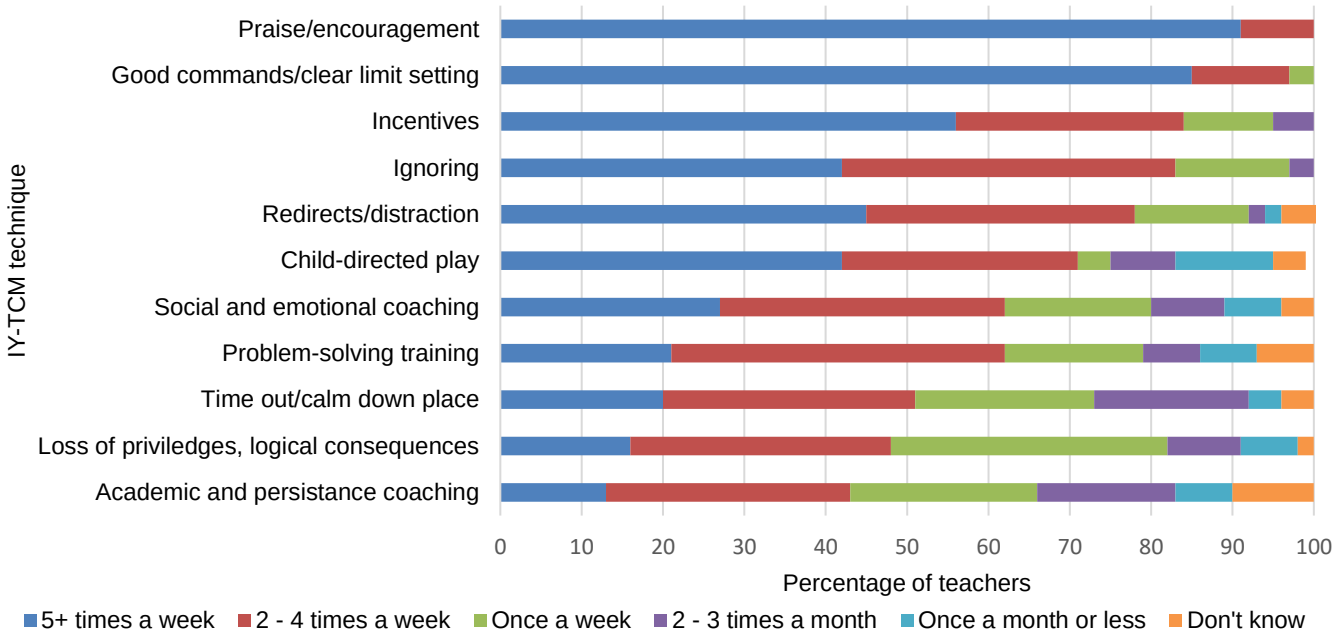
The programme Theory of Change (ToC) expects that teachers will embed IY®-TCM techniques into their practice, meaning that pupils receive a daily dosage of the intervention. Findings from the survey and interviews are largely positive and support the ToC because they suggest that most teachers who responded did embed the intervention into their practice, although the extent to which they used the techniques differed.

The endpoint survey in summer 2021 asked teachers in the intervention group which techniques from the IY®-TCM training they used, followed by how frequently they used these techniques. Use of the techniques amongst responding teachers was high, with most teachers reporting that they had used most of the techniques promoted in the training. However, the endpoint survey was completed by less than half of teachers in the intervention group[21], and so results may not reflect those who chose not to respond. Most teachers reported using praise/encouragement (85%), ignoring (78%), good commands/clear limit setting (72%) and incentives (70%). Smaller proportions of teachers said they used academic and persistence coaching (37%), problem-solving training (35%) and child-directed play (29%).

The endpoint survey also asked teachers in the intervention group to indicate how frequently they used each IY®-TCM technique on a scale from 'five or more times a week' to 'once a month or less'. The findings suggest that most of the responding teachers used at least one technique on a daily basis ('five or more times a week'). The majority of teachers (91%) used praise/encouragement on a daily basis and most (85%) reported using good commands/clear limit setting daily. Teachers reported using 'time out', loss of privileges and academic persistence coaching techniques less frequently (see Figure 4). The findings from the case-study interviews with teachers offer an explanation for these trends.

[21] A total of 230 teachers were assigned to the intervention group and 101 intervention group teachers completed the endpoint survey, of these, 82 attended the IY®-TCM training.

Figure 4: How often teachers used techniques from IY-TCM training



Notes: Data from the endpoint survey: *How often, on average, do you use the techniques from STARS: Incredible Years® Teacher Classroom Management training in the classroom?* (N = 74). Items from the survey re-ordered to be presented in descending order. Percentages have been round to integers so may not sum to 100.

Interviews with 15 teachers in nine case-study schools provided further information on teachers' use of the IY®-TCM techniques. Despite the partial school closures that occurred straight after the training workshops and therefore impacted pupils' exposure to the intervention, when teachers reflected on their first year of delivering IY®-TCM, 12 Year 1 and 2 teachers in seven schools reported integrating the techniques into their classroom practice. All five Year 2 teachers interviewed at the end of the trial in summer 2021 reported that they had embedded the programme into their practice, using the techniques on a daily basis. Some teachers also commented that they used the techniques whilst teaching remotely, described further in the section below. The teachers interviewed said that techniques which reinforced rules and focused on positive behaviour through positive reinforcement were most effective and easiest to implement in their classrooms. In addition, teachers were already using these, or similar, techniques meaning they did not have to make significant adaptations to their practice to embed the techniques. This provides an explanation for why survey respondents reported using such techniques most frequently. In comparison, teachers who were interviewed reported using punitive techniques (such as loss of privileges) less often, mostly because this did not align with their school's behaviour policy and because IY®-TCM encourages teachers to use such techniques selectively. Not all teachers could implement the 'time out'/calm down place for the practical reason of not having appropriate space in their classroom or school.

The evaluation team intended to conduct CLASS observations at baseline and endpoint (see methodology for details) within three case-study schools to observe implementation of the IY®-TCM techniques in practice. Unfortunately, the endpoint observations, which were scheduled to take place after all six training sessions had been completed, could not take place due to government guidance on limiting travel and maintaining social distancing during the pandemic.

The evaluation team carried out further analysis to explore any differences in teachers' use of the techniques, related to their teaching experience and the level of challenging behaviour they reported.

As will be mentioned in the *'Quality of delivery'* and *'Responsiveness'* sections of this report, interviews with IY®-TCM trainers suggest that some experienced teachers did not find the training valuable because they were already aware of the techniques IY®-TCM promoted. To explore whether there was evidence of this in the survey responses, the team compared the frequency with which each behaviour management technique was reportedly used across three categories of teaching experience: four years or less, five to 19 years and 20 years or more. Pairwise Fisher's exact tests were used to determine whether there was a statistically significant difference in teacher's use of techniques

advocated by IY®-TCM between each pair of experience categories. The analysis by teaching experience only showed three statistically significant differences which, given the risk of type I error (false positive tests) entailed by multiple comparisons, provided little or no evidence that teachers' use of the techniques differed based upon their level of experience. The three differences showed that more teachers with four years or less teaching experience reported using ignoring as a technique (N=14, 100%), compared to teachers with five to 19 years' experience (N=28, 65%) (p = 0.012). More teachers with four years or less teaching experience reported using good commands/clear limit setting (N=13, 93%) compared to teachers with five to 19 years' experience (N=26, 60%) (p = 0.043). Finally, more teachers with 20+ years' experience reported using time out/calm down place (N=20, 80%) compared to teachers with five to 19 years' experience (N=24, 56%) (p = 0.034).

To investigate if use of the intervention was dependent upon the level of challenging behaviour in the classroom, the five response options for the statement 'I frequently deal with difficult behaviour in the classroom' were merged into agreement ('agree'/'strongly agree') or lack of agreement ('strong disagree'/'disagree'/'neither agree nor disagree'). The evaluation team investigated whether more teachers who reported frequently dealing with challenging behaviour reported use of the eleven IY®-TCM techniques using Fisher's exact tests. The analysis only showed one statistically significant difference, providing little or no evidence that teachers' use of the techniques differed based upon the level of challenging behaviour they experienced. More teachers who reported that they did not frequently deal with challenging behaviour reported using problem solving training (N=22, 46%) than those teachers who did report frequently dealing with such behaviour (N=7, 21%) (p = 0.035). Given the number of significance tests performed, it is important to note, again, that this statistically significant finding may be due to chance.

How was dosage of the intervention impacted by Covid-19?

The endpoint survey in summer 2021 asked teachers to indicate how often they had used the IY®-TCM techniques when teaching remotely. Teachers said they used praise/encouragement, incentives and good commands/clear limit setting most frequently, with over 80 per cent of teachers using these techniques at least once a week. The case-study interviews with teachers offered further insights into this trend, with teachers reporting that these positive reinforcement techniques could be easily adapted for the online classroom. Some teachers had also adapted their use of emotional coaching and relationship building techniques for online delivery. Teachers gave examples of providing emotional coaching through individual weekly phone calls to pupils and holding class assemblies to maintain the pupil-teacher and peer relationships. These findings suggest that although IY®-TCM was not specifically designed to be used as part of remote teaching delivery, teachers were able to adapt and implement some techniques more readily than others, meaning some pupils were still exposed to some of the IY®-TCM intervention whilst being taught by trained teachers as part of remote learning.

Why did teachers not implement IY®-TCM techniques?

The majority of teachers in the intervention group responding to the survey did use the IY®-TCM techniques (93%). The main reasons for teachers not using the techniques mentioned in a small number of case-study interviews were because teachers did not recognise the added value of IY®-TCM to their practice as they were already using the techniques and approaches promoted by the programme or because the recommended approached did not align with their school's behaviour policy.

Usual practice

Programme differentiation

What was usual practice in intervention schools before the intervention?

Interviews with teachers in the nine case-study schools and classroom observations at the beginning of the trial in three of these schools provide insight into the types of behaviour management approaches intervention schools used before participating in the IY®-TCM programme. These included a range of visual behaviour displays to communicate to pupils the outcomes of their behaviour choices and the use of reward systems (e.g., ClassDojo points) implemented at both the level of the individual pupil and the whole class. There was variation in the extent to which these schools used sanctions as part of their behaviour management.

Before they started the IY®-TCM programme, all teachers who took part in the interviews said were unsure about the specifics of the programme, but they hoped that the techniques it provided would align with their existing school

behaviour policies – in their view this would be necessary in order to facilitate the integration of IY®-TCM into their practice.

What are the implications for the results (attribution) if usual practice is similar to the intervention?

Ten teachers in seven case-study schools involved in the interviews reported that (as described in the *'Quality of delivery'* section of this report) the IY®-TCM approach was similar to their practice before the training and that they were already familiar with using many of the techniques. For example, one teacher commented: *'STARS didn't really teach us anything new, but if did remind us about focussing on the positives.'* (Year 1 teacher)

The similarities between the IY®-TCM approach and the practices teachers already used have two main implications for the intervention. First, this could provide reassurance to teachers that aspects of their existing practice were evidence-based and they should continue using them, in some cases to a greater extent. On the other hand, it could lead to teachers feeling that their ongoing investment of time in the training (i.e., six full-day sessions) was not worthwhile. Both of these responses were present in the interview data (see also the *'Quality of delivery'* section). For example, several teachers from the case-study schools said they had increased their use of praise and reinforcement of positive behaviour and rewarding minor positive behaviours. However, three experienced teachers from different schools said that their usual practice was so similar to that promoted by IY®-TCM that they were unable to identify any impacts for themselves, their pupils or the wider school resulting directly from their participation in programme.

What practices and approaches did IY®-TCM replace in participating schools?

The greatest difference to the practice of teachers interviewed was the removal of punitive approaches for dealing with non-compliant behaviour. Some teachers reported that they no longer used visual behaviour displays to record negative pupil behaviours, for example by substituting this for the recommended IY®-TCM approach to give children praise and rewards for positive behaviours. One teacher who previously used visual behaviour displays in this way commented:

'*Because of STARS, we just don't believe that is right any more. Children shouldn't be named and shamed, they shouldn't be having to sit and see their name on a thunder cloud all day. So we have moved fully away from that, it can be soul destroying for them. We picked up a lot in STARS about impacts on their* [pupils'] *confidence and self-esteem.'* (Year 1 teacher)

Further detail on the changes made to behaviour management following participation in IY®-TCM can be found under the section below on: '*What are the impacts of the IY®-TCM programme on school behaviour policies'?*

Monitoring of the control group and of additional behaviour management CPD received by the intervention group

What were the business-as-usual practices in control schools and how distinctive are these from IY®-TCM practices?

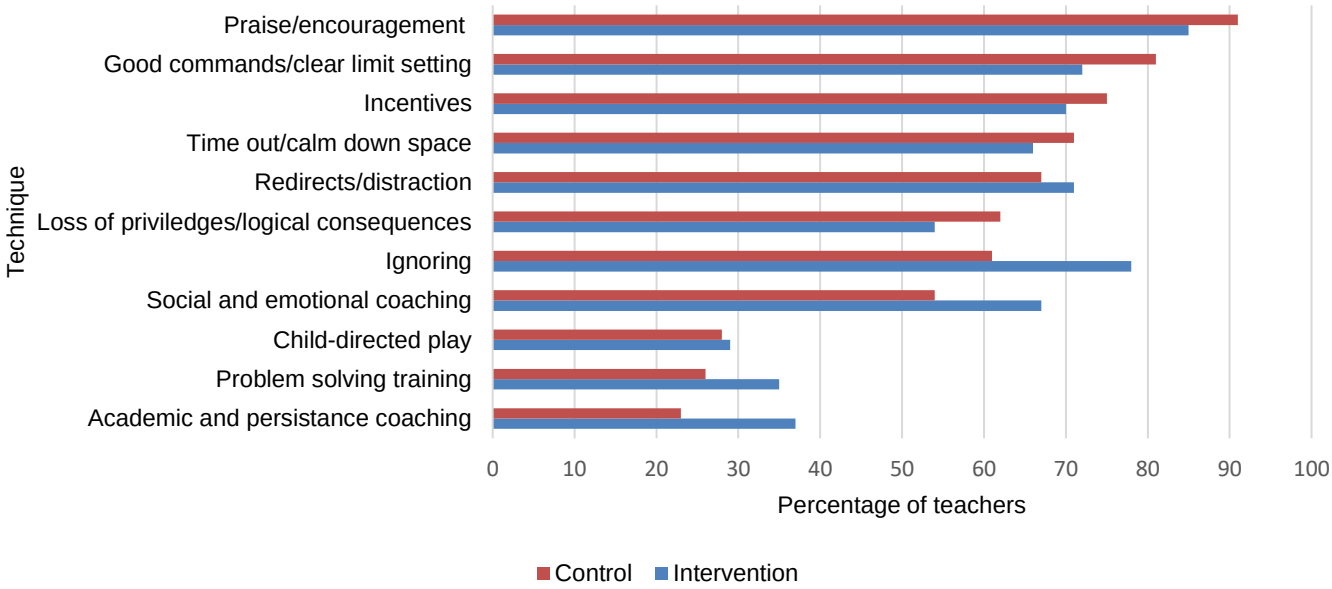
To understand the business-as-usual practices in control schools, the baseline and endpoint surveys asked teachers randomised to the control group to comment upon their school-wide behaviour management strategy, as well as any specific behaviour management techniques they used. This information was used to understand how schools' normal practices differed from IY®-TCM.

Almost all teachers in control schools reported that they had a school-wide behaviour management policy and most felt that the policy had adequate strategies for managing challenging behaviour.

Survey findings suggested that the IY®-TCM practices were similar to those used in control schools, but there was a difference of emphasis on particular techniques. The survey presented control group teachers with a list of behaviour management techniques that the IY®-TCM promoted and asked them to indicate which of the techniques they used in their practice. Control group teachers reported that they used the same techniques promoted by the IY®-TCM programme. This might be expected, because the finding that intervention teachers were already familiar with the techniques promoted by IY®-TCM before attending the training is likely to be applicable to control group teachers. However, despite this, there were some differences between intervention and control groups teachers' use of behaviour management techniques. Fewer teachers in the control group used the more specific strategies promoted by IY®-TCM for example ignoring or redirecting negative behaviours and providing pupils with social-emotional and academic coaching.

Figure 5 shows the percentage of teachers within each group who reported using certain behaviour management techniques.

Figure 5: Behaviour management techniques used by teachers in the control compared to intervention group at endpoint



Notes: Data from the endpoint survey. Control group: *Which of the following techniques do you use for classroom management purposes?* (N = 69). Intervention group: *Which techniques have you used as a result of the STARS: Incredible Years® Teacher Classroom Management training?* (N = 82). Items from the survey re-ordered to be presented in descending order.

The largest proportion of control and intervention group teachers reported using praise/encouragement in their practice. This finding might be expected as all teachers will be familiar with using this technique and can easily implement it into their practice. Teachers' use of techniques in response to negative behaviour differed between the control and intervention group. Fewer teachers in the control group reported redirecting or ignoring negative behaviour compared to those in the intervention group while more teachers in the control group reported using loss of privileges. The IY®-TCM programme encourages teachers to ignore or redirect poor behaviour and discourages use of punitive approaches, which could explain these trends. Furthermore, fewer teachers in the control group reported using coaching techniques. These are likely to be techniques that teachers require training to implement, so their differential use between teachers in the control and intervention group again could suggest that this is an area where the IY®-TCM training impacted upon teachers' practice.

How did the business-as-usual practices in control schools change over time?

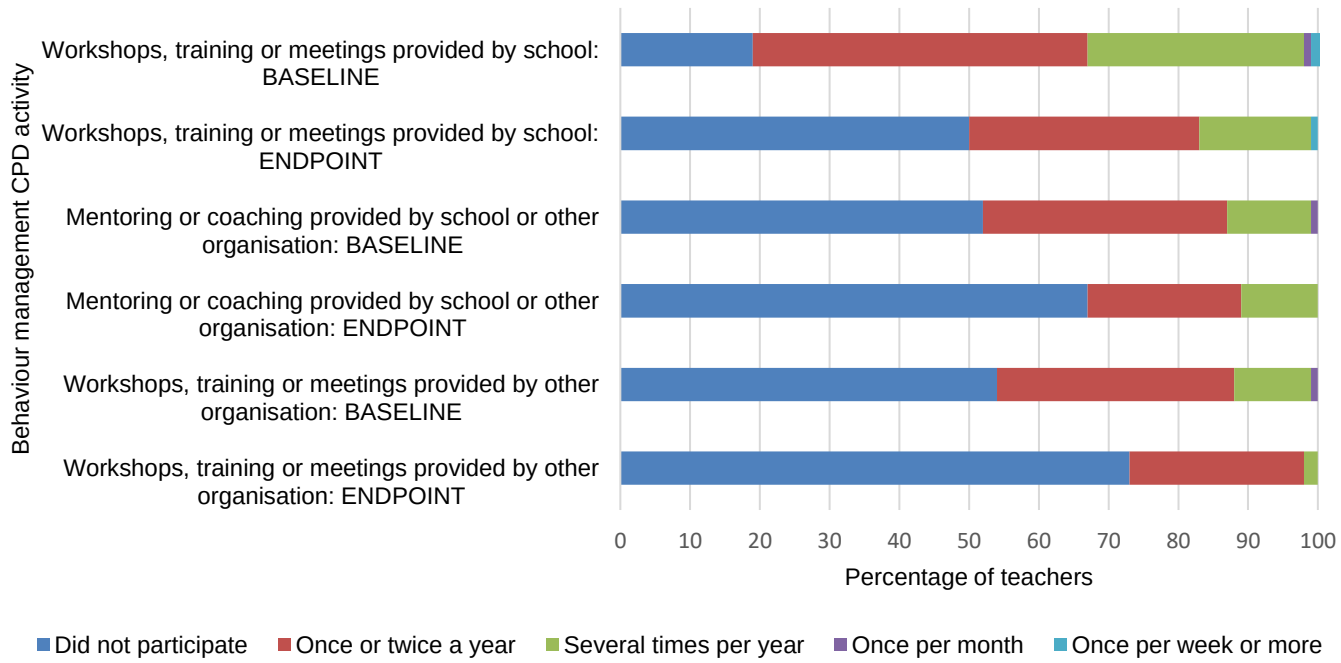
The survey asked teachers in the control group at baseline and endpoint to indicate how frequently they had participated in behaviour management related CPD activities. This question aimed to understand firstly, the type of behaviour management CPD teachers received and secondly, how their participation in these activities changed over the course of the trial. Control group teachers reported participating more frequently in behaviour management CPD provided by their school compared to that provided by external organisations. Teachers reported lower participation in all forms of behaviour management related CPD activities at endpoint which may be at least in part related to changes in CPD foci due to the Covid-19 pandemic.

At baseline, training on behaviour management was relatively common. Most teachers in the control group (81%) reported that over the last year, they had participated in 'workshops, training or meetings about behaviour management provided by my school'. Around half of teachers (48%) had participated once or twice a year and around a third said they participated more often (at least several times a year). Teachers reported participating in 'workshops, seminars or other CPD about behaviour management provided by another organisation', or 'mentoring or coaching provided by your school or another organisation' less frequently. Around a third indicated that they had participated in these activities once or twice a year, however around half said they had not participated in these activities at all.

Participation in all activities reduced over the course of the trial. The survey asked control group teachers again at endpoint to indicate how frequently they had participated in these activities since November 2019. Half of the 64 responding teachers (50%) had participated in workshops, training or meetings in school, but smaller proportions had been involved with mentoring or coaching on behaviour management (33%) or workshops, training or meetings provided by another organisation (27%).

Figure 6 shows how frequently teachers reported participating in these CPD activities at both time points.

Figure 6: Control group teachers' participation in behaviour management CPD activities



Notes: Data from the control group from the baseline and endpoint survey. Baseline: *In the past year, have you participated in any of the following activities? If yes, how frequently?* (N = 202). Endpoint: *Since November 2019, have you participated in any of the following activities? If yes, how frequently?* (N = 69. Five teachers did not respond so percentages were recalculated based on the 64 teachers who did respond). Items from the survey re-ordered to be presented in descending order. Percentages have been round to integers so may not sum to 100.

What (other) behaviour management related CPD activities have intervention group teachers taken part in?

The endpoint survey asked teachers in the intervention group what other behaviour management CPD they had participated in over the course of the trial. Teachers who were assigned to the intervention and who reported that they *did* attend the IY®-TCM training responded to this question (N = 77), as did teachers who were assigned to the intervention but reported that they *did not* attend (N = 19[22]). The findings showed that intervention group teachers continued to participate in other behaviour management activities during the trial, but the extent of teachers' participation differed by activity and between the two groups of teachers (those who attended the training and those who did not).

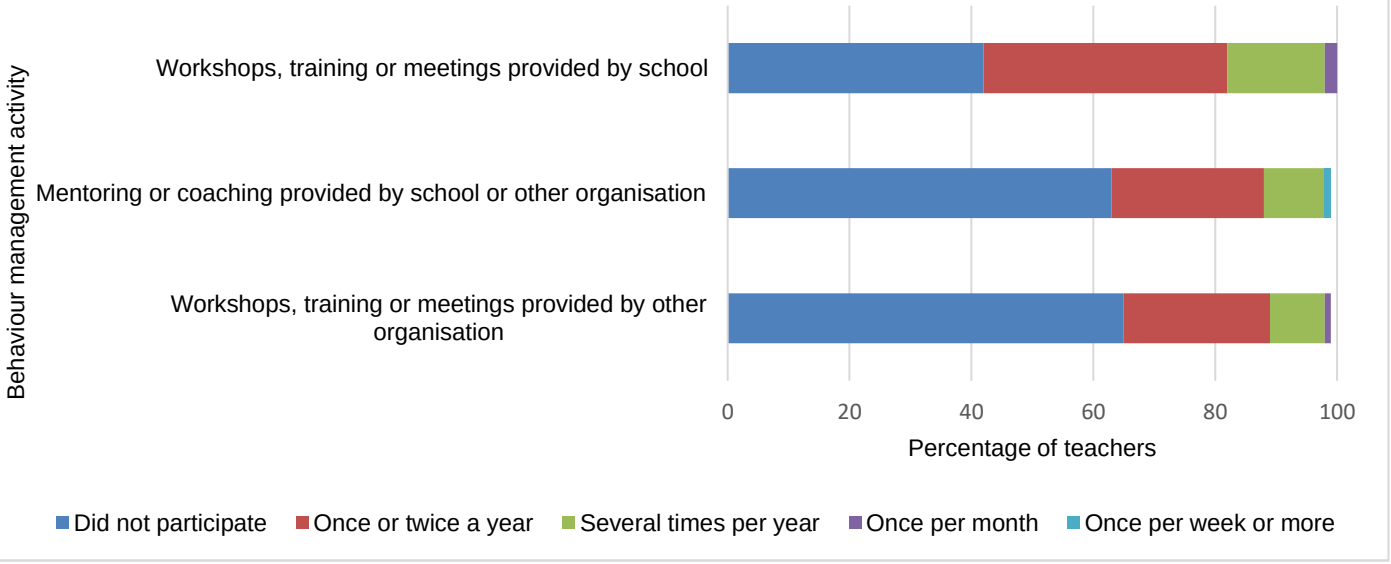
To ascertain what other behaviour management CPD activities teachers who attended the training received in addition to the IY®-TCM programme, the endpoint survey asked teachers to indicate the type training activities they had participated in since November 2019 and the frequency of these activities. The interviews with teachers provided further insights.

Over half of teachers (58%) in the treatment group who responded to the survey had also participated in workshops, training or meetings about behaviour management provided by their school during the trial period. Smaller proportions of teachers had participated in mentoring or coaching (36%) and/or workshops, seminars or other CPD on behaviour provided by another organisation (34%).

The frequency that teachers participated in these activities is shown in Figure 7.

[22] This includes one respondent who did not know if they attended the training.

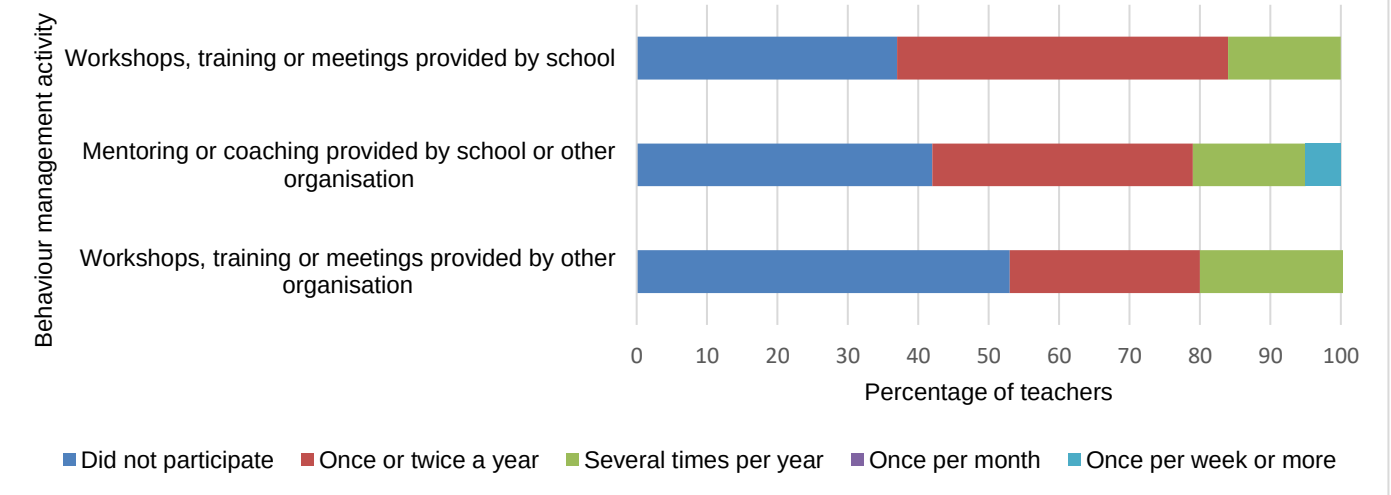
Figure 7: Teachers' participation in behavour management CPD alongside IY®-TCM



Notes: Data from the endpoint survey: *Since November 2019, have you participated in any of the following activities in addition to the STARS: Incredible Years® Teacher Classroom Management programme? If yes, how frequently?* (N = 82. Five teachers did not respond so percentages were recalculated based on the 77 teachers who did respond). Items from the survey re-ordered to be presented in descending order. Percentages have been round to integers so may not sum to 100.

Teachers who were part of the intervention but did not attend the IY®-TCM training (N = 19) reported participating in other behaviour management activities slightly more frequently than those who did attend the training. Almost two-thirds of teachers (63%) said they had participated in workshops, training or meetings about behaviour management provided by their school. Just over half of teachers (53%) had participated in mentoring or coaching on behaviour management and just less than half (48%) had participated in workshops, seminars or other CPD provided by another organisations. The frequency that these teachers participated in other behaviour management activities in addition to IY®-TCM is shown in Figure 8.

Figure 8: Teachers' participation in behavour management CPD instead of IY-TCM



Notes: Data from the endpoint survey: *Since November 2019, have you participated in any of the following activities? If yes, how frequently?* (N = 19). Items from the survey re-ordered to be presented in descending order. Percentages have been round to integers so may not sum to 100.

As part of the case-study interviews with teachers in the intervention group, the evaluation team asked whether they had received any other behaviour management CPD, alongside IY®-TCM. Some had done so but reported that the content either aligned with IY®-TCM, or was for a different purpose, for example to address high-level classroom disruption and to support the behaviours of pupils with SEND. These findings suggest that despite some teachers

receiving other behaviour management CPD alongside IY®-TCM, this did not interfere with their ability to implement the IY®-TCM approach.

Quality of delivery

This section includes the findings from intervention group teachers and trainers on their perceptions of the quality and effectiveness of the IY®-TCM training. The impacts that teachers recognised from implementing the programme for themselves, their pupils and their school are also considered.

What do teachers think about the quality and effectiveness of delivery by IY®-TCM?

Quality of the IY®-TCM trainers

Teachers were very satisfied with the quality of the IY®-TCM trainers. The endpoint survey asked teachers to indicate how useful they found elements of the IY®-TCM training and 79 per cent of the 82[23] responding teachers said that the support from the trainer was a useful aspect. Some of the teachers who participated in a case study interview said their IY®-TCM trainer had teaching experience, which they valued because it meant the trainer could relate to the behavioural challenges teachers discussed. Teachers also commented on the trainers' clear knowledge and experience of the IY®-TCM programme.

'I thought it was brilliant. They [the trainers] *were both really knowledgeable – any questions they had thrown at them, they had a response. They came across really confident.'* (Year 1 teacher)

Quality of the IY®-TCM content

Over three-quarters of the 82[24] survey respondents (77%) said that they had found the information presented by the trainer to be useful. The teachers who participated in the case-study interviews provided some additional views on the quality of the IY®-TCM content.

Eleven of the 15 teachers interviewed said they had been satisfied with the content of the IY®-TCM training, describing it as 'helpful, 'informative' and 'comprehensive'. The four who said they were less satisfied with the content of the course felt that they did not gain much from it because: they were already using the recommended techniques; behaviour management was not an issue for them; the training did not provide support for the high-level disruption and more challenging behaviours they were facing in their classrooms, (e.g. in relation to pupils' Special Educational Needs); and/or they felt the video-clip examples used in the sessions were outdated and not sufficiently related to the English classroom context.

The evaluation team asked teachers who took part in case-study interviews whether they felt they were any techniques or situations where they needed additional support. For the most part, teachers felt they received all the training they needed to implement IY®-TCM confidently in their classrooms. However, some teachers did identify that additional refresher training would have been useful for them at the beginning of the 2020/21 academic year, given the disruption to classroom teaching immediately after the training which was then followed by the school summer holidays. In addition, teachers said that they would have found it valuable to visit other schools where teachers successfully implement IY®-TCM, to observe examples of best practice.

Elements of the IY®-TCM training that worked well and less well

The evaluation team asked teachers from case-study schools what they felt had worked particularly well about how the IY®-TCM training was delivered. Interviewed teachers particularly welcomed the time provided in the sessions for reflection and discussions between teachers from different schools with different experiences. As reported in the 'Responsiveness' section of this report, 89 per cent of the survey respondents said this was a useful aspect of the IY®-

[23] There were 82 teachers who were shown the survey questions around delivery of the training. Two of these did not respond to this question. There were 101 respondents to the survey out of a potential 230 teachers in the intervention group although 19 of these stated that they had not attended any IY®-TCM training so they would not have been shown the questions concerning the delivery of the training.
[24] See footnote 24. Again, there were two teachers who offered no response to the question around usefulness of the information presented by the trainer.

TCM training. Teachers also valued having time between training sessions to go back into the classroom, try out the new techniques then feedback on the success of these at the next training session.

Teachers also commented on aspects of the training they felt worked less well. A few teachers involved in the case-study interviews disliked the role-play tasks. Teachers suggested this element of the training could be improved either by reducing the amount of role-play, or through the trainers doing this themselves, until teachers become more comfortable and familiar with one another. On the other hand, other teachers found the role-play exercises helpful practice for dealing with the challenging situations they encountered in the classroom.

The evaluation team asked teachers whether they felt the training had prepared them to implement IY®-TCM. Most teachers felt that the training had prepared them well to implement it in their classroom and they felt confident in using the techniques taught after each training session. Teachers said that practicing and discussing strategies during the training had helped them to understand how to implement them.

What do the IY®-TCM delivery team think about the quality and effectiveness of delivery?

The evaluation team interviewed two members of the developer team towards the end of the trial period about their perceptions of the quality and effectiveness of the training delivery. One of the developers had been involved in the joint delivery of a IY®-TCM training session themselves due to staff absence so was able to comment on this experience as well as their feedback from ongoing interactions with the trainers during the delivery period. We also interviewed a small number of trainers about their experiences of IY®-TCM delivery. Both the developers and the trainers themselves reported feeling mostly positive about the how the training was delivered yet did highlight challenges to effective delivery and aspects of the programme that require improvement.

The developer reported being confident that trainers across all six hub areas had delivered high quality training sessions. As mentioned in the *'Fidelity'* section of this report, the developer and trainers felt that trainers' knowledge and experience of the school system facilitated this quality.

Trainers reported aspects of the programme content and delivery that they felt required improvement. Teachers raised the issue of the content being outdated, too basic and the need for it to be updated to reflect that teachers' knowledge, understanding and practice is more advanced than when the programme was developed. They further commented that some of the training venues did not have the necessary facilities to deliver sessions to best effect.

What factors are affecting the quality of delivery?

The evaluation team observed three IY®-TCM training sessions in total (sessions 1, 4 and 6), delivered in three different hub areas. Researchers identified that although there were differences in how the three trainers delivered these sessions based on the topic areas being covered and the individual experiences they drew upon, they did not identify any differences in the quality of delivery, which they observed to be high across all three training groups.

Trainers highlighted that the relationship and rapport they built with the teachers was a key factor for effective delivery and researchers observed that trainers created a safe, respectful environment for the groups, which facilitated collaboration and discussions between teachers. Trainers also described how they developed a network amongst the teachers. They commented that they facilitated discussions and reflections on elements of the training that teachers enjoyed – through questioning and prompting teachers to think deeper about how to build upon their practice.

Trainers reported on the difficulty of delivering the sessions effectively when met with resistance from a few teachers. One trainer also highlighted the particular challenge of delivering the content effectively to teachers whose classes had very high levels of pupil need. They identified the impact this was having upon teacher's confidence and well-being and found it challenging to respond to these needs, alongside delivering the programme content as intended.

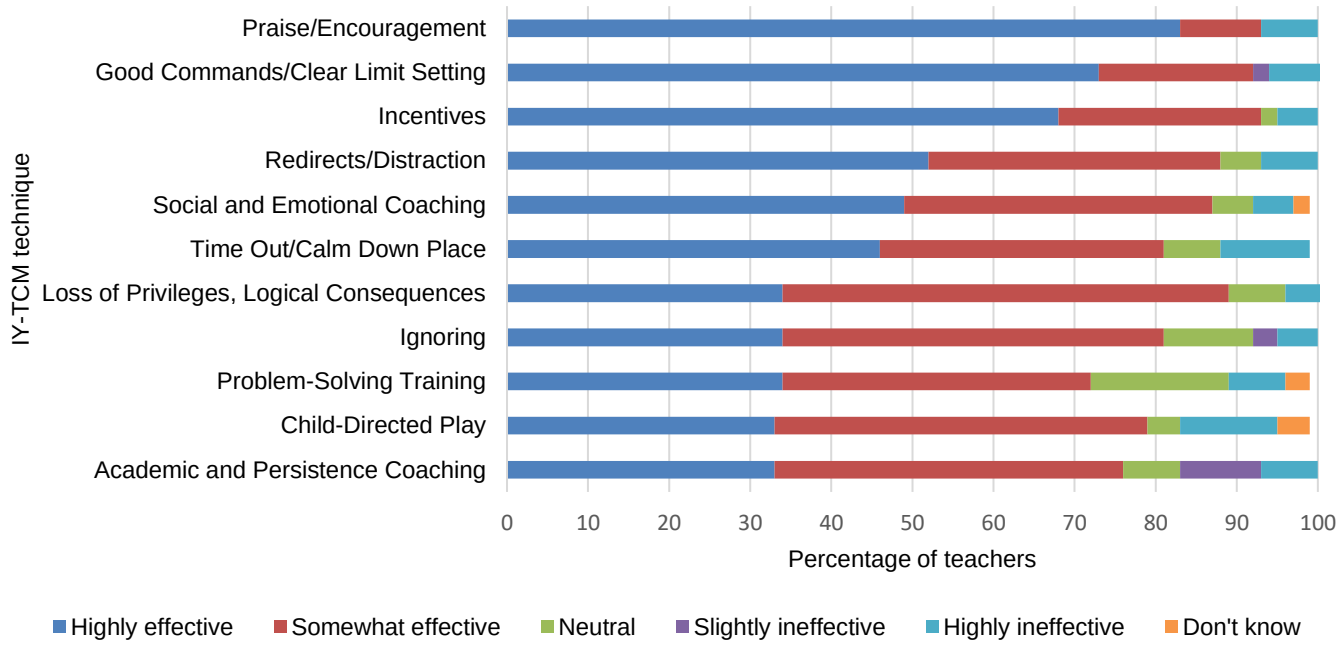
What challenges did deliverers face in providing the programme?

The challenges trainers faced in delivering the sessions mostly related to teachers having a negative response to the programme and therefore trainers found it difficult to engage them in the sessions. These challenges are described earlier in this section and in the *'Responsiveness'* section of this report.

What were teachers' perceptions of the effectiveness of individual elements of the programme?

The survey asked teachers in the intervention group to rate how effective they found the techniques from the IY®-TCM training. With the exception of problem-solving training, over three-quarters of teachers rated all techniques as somewhat effective or highly effective. Figure 9 shows how effective teachers found the IY®-TCM techniques.

Figure 9: Teachers' perception of the effectiveness of IY-TCM techniques



Notes: Data from the endpoint survey: *How effective have you found the techniques from STARS: Incredible Years® Teacher Classroom Management?* (N = 74). Items from the survey re-ordered to be presented in descending order. Percentages have been round to integers so may not sum to 100.

Teachers involved in case-study interviews offered further insights into the techniques they found to be effective and the reasons for this. Overall teachers said they found techniques that supported positive relationships and helped pupils to regulate their own behaviour most effective. Teachers reported that techniques such as ensuring they greeted all children individually each morning and activities where pupils and teachers shared something about themselves had been particularly effective for building positive relationships. These techniques had helped pupils and teachers to get know one another, and to build an environment where there was trust and respect.

Some teachers said that techniques, such as setting up a 'calm corner' and practicing emotional coaching, had been particularly effective for supporting pupils' emotional regulation. These interviewees reported that pupils in their classes had started to identify when both themselves and their peers were experiencing heightened emotions. They said pupils now recognised when they needed to take 'time out'. This typically prevented their behaviours from escalating.

Some teachers found that changes to their use of rewards systems had been particularly effective because they now incentivised pupils to display the expected, positive behaviours. They had noticed that pupils who displayed less compliant behaviour began to model pupils who received the incentives and verbal praise so they too could receive the same rewards and positive attention. These teachers reported that use of these positive reinforcements had helped to build pupils' self-esteem, had increased their confidence and had encouraged a positive attitude towards their work and behaviour. On the other hand, pupils who did display negative behaviours began to realise their actions would be ignored, leading to a reduction in such actions.

Two teachers also commented on the value of scripts – teachers had written these down to ensure that when dealing with unacceptable behaviour, they entered in the situation knowing what they would say, meaning they could remain calm and unflustered. Teachers also reported sharing these scripts with Year 3 teachers so the message pupils received would remain consistent when they moved to the next year group.

What are the perceived impacts on pupils' engagement in learning?

The programme ToC states that in classrooms where teachers implement IY®-TCM, inappropriate behaviours will be effectively managed and reduced, therefore reducing the disruption to teaching. This will increase pupils' capacity to engage with their learning. The ToC also states that the techniques teachers use will lead to pupils feeling more positive and confident about actively participating in learning. These outcomes will ultimately lead to improved engagement in learning and improved academic outcomes.

The findings largely supported the intended outcomes stated in the ToC. Data for the endpoint survey and interviews with teachers in case-study schools was largely positive and suggested that they found IY®-TCM effective. The majority of survey respondents (94%) felt the impact on pupil engagement was positive – over half (58%) of the 74 teachers who responded reported a large positive effect and just over a third (36%) reported a slightly positive effect. No teacher reported that the techniques had a negative effect upon pupils' engagement in learning and only a small minority of teachers felt the techniques had no effect (3%) or were unsure (3%). The interviews with teachers provide further insight into the impact IY®-TCM had upon pupils.

Some of the teachers interviewed reported improvements in pupils' behaviour had led to calmer classrooms. Noting improved pupil concentration and increased capacity to engage with learning. They felt this contributed to pupils making better progress and achievement. Although not all teachers reported an effect on increased academic outcomes, some teachers reported that pupils showed more confidence in their abilities, and more perseverance, resilience and pride in their work.

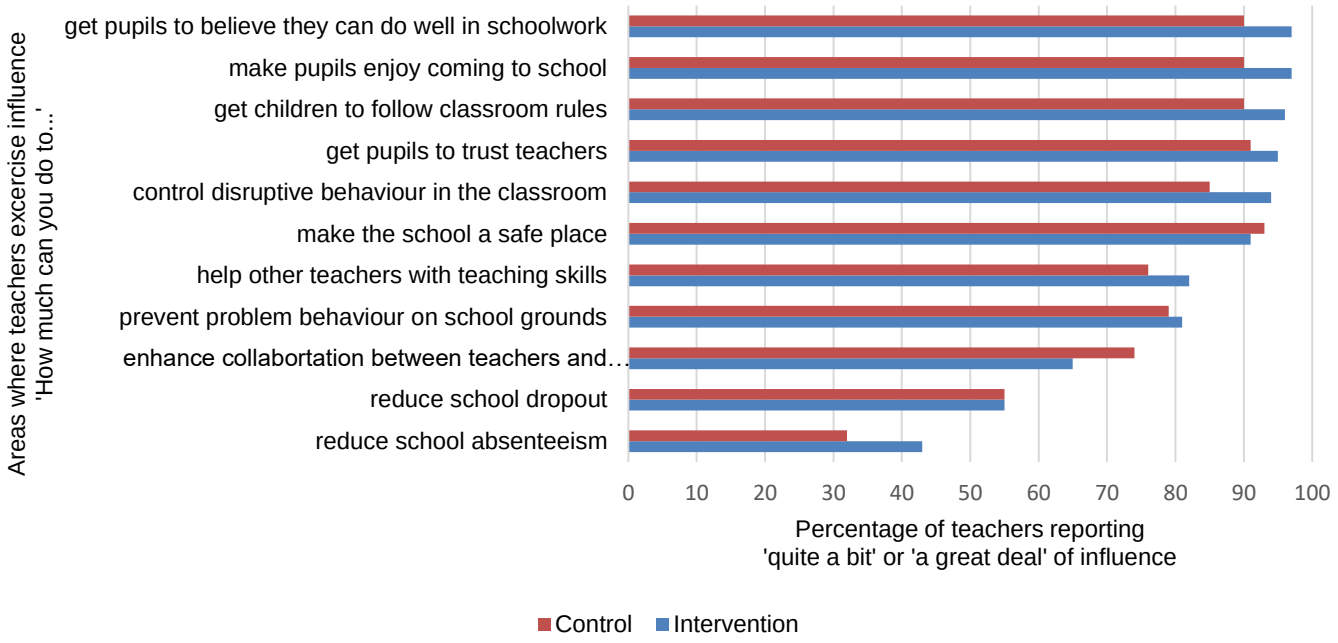
Three teachers who were interviewed did not identify any impacts upon their pupils' engagement in learning. Reasons for this included: limited time to fully implement the IY®-TCM approaches due to partial school closures (see '*How was dosage of the intervention impacted by Covid-19?'* for further details); pupils' academic and social-emotional skills regressing during partial school closures and that they were already implementing similar techniques to IY®-TCM, so they were unable to identify any distinct impacts of the approach.

What are the perceived impacts on the IY®-TCM programme on teacher practice and self-efficacy?

The programme ToC assumes that teacher self-efficacy underpins the mechanisms of change for IY®-TCM. To investigate this assumption, the survey measured perceived self-efficacy by asking intervention and control group teachers at baseline and endpoint the amount of influence they felt they exercised in their job roles. Figure 10 shows the comparison between intervention and control teachers perceived self-efficacy at endpoint.

At endpoint, teachers in both groups (intervention and control) felt they had the most influence in areas relating to classroom level decisions/actions, such as getting pupils to: 'believe they can do well in their schoolwork'; 'follow classroom rules'; 'trust teachers'. In comparison, intervention and control group teachers appeared to believe they had less influence over whole school level decisions and issues, such as reducing school dropout or reducing school absenteeism. However, teachers in the intervention group reported higher self-efficacy in all but three areas the survey asked teachers to comment upon. This finding supports the ToC as it suggests that, compared to teachers in the control group, teachers in the intervention group felt better able to influence decisions relating to classroom and whole-school level challenges.

Figure 10: Comparison of intervention and control group teachers' self-efficacy at endpoint



Notes: Data from the endpoint survey. *Please indicate your opinions about each of the statements below by selecting the appropriate number.* (Control: N = 69; Intervention: N = 101). Items from the survey reordered to be presented in descending order.

Analysis was conducted on 'matched' teachers (those that responded to both baseline and endpoint surveys) that responded to the items on efficacy (intervention N=53, control N=82) to explore any differences in teachers' growth in self-efficacy over the trial period. The evaluation team converted agreement with each of the 11 items to a numeric scale from –4 to 4, with higher numbers corresponding to greater agreement with the (positively worded) items. The items were then summed to form composite scores and each respondent's baseline composite score was deducted from their endpoint composite score, so that the resulting measure had a theoretical range of -88 to 88, with positive numbers indicating a general improvement in self-efficacy from baseline. The improvement from baseline amongst control and intervention respondents was then compared using a Mann-Whitney U test. This analysis showed that on average teachers who took part in the IY®-TCM training saw a greater improvement in terms of their overall self-reported efficacy as a teacher than those assigned to the control group (mean = 3.26 points; p = 0.035), suggesting that participation in IY®-TCM positively impacts upon teachers perceived self-efficacy.

The survey also asked intervention and control group teachers at baseline and endpoint to comment on their experience and confidence with behaviour management. This was done via three items; agreement with each item statement was measured using a 5-point Likert scale ('strongly disagree' to 'strongly agree'). For each of these three items the Likert scale was also converted to a numeric scale from -2 to 2, with higher numbers corresponding to greater agreement with the item statements. The respondent's baseline number was then deducted from their endpoint number, so that the resulting measures had a theoretical range of -4 to 4, with larger numbers corresponding to greater agreement with the item statements since baseline. This allowed comparisons between control and intervention teacher's change from baseline, described for each item below.

The first item statement was 'I rarely deal with difficult behaviour in the classroom'. At endpoint, more teachers in the intervention group (39%) reported frequently dealing with challenging behaviour (either 'disagreeing' or 'strongly disagreeing with the statement) compared to teachers in the control group (35%). Amongst matched teachers that responded to this item (N = 139), those who reported in the survey that they took part in the IY®-TCM training saw a greater improvement from baseline than those assigned to the control (mean = 0.4 points) in terms of how frequently they dealt with difficult behaviour. These findings suggest that implementation of IY®-TCM was associated with teachers' perceived frequency of dealing with challenging pupil behaviour over the trial period. However, it must be noted that at baseline and endpoint, fewer teachers in the control group reported frequently dealing with challenging behaviour suggesting overall better perceived behaviour amongst control group pupils.

The second item statement was 'I usually feel confident when dealing with difficult behaviour'. At endpoint, more teachers in the intervention group (84%) reported feeling confident with dealing with difficult behaviour ('agreeing' or 'strongly agreeing' with the statement) compared to the control group (79%). Amongst matched teachers that responded to this item (N = 139), those teachers who reported in the survey that they took part in the IY®-TCM training reported a greater improvement from baseline than those assigned to the control (mean = 0.33 points) in terms of their confidence in dealing with difficult behaviour. These findings suggest that participation in IY®-TCM may have supported teachers to feel more confident when dealing with difficult behaviours in the classroom.

The third item statement was 'I do not worry about pupil/class behaviour when I am not at school'. At endpoint, fewer teachers in the intervention group (34%) reported that they worried about pupil/class behaviour when not in school ('agreeing' or 'strongly agreeing' with the statement) compared to teachers in the control group (39%). Amongst matched teachers that responded to this item (N = 139), teachers who reported in the survey that they took part in the IY®-TCM training saw a greater improvement from baseline than those assigned to the control (mean = 0.68 points), in terms of how much they worried about pupil behaviour outside of school. These findings suggest that participation in IY®-TCM was associated with a reduction in the amount teachers worry about pupils' behaviour when not at school.

The evaluation team asked teachers who participated in interviews what impact the IY®-TCM training had upon their self-efficacy relating to managing challenging behaviour and their confidence in dealing with such behaviour. Eleven of the 15 teachers reported that the range of techniques the training provided them with meant they felt better able to deal with challenging behaviours. Teachers further commented that their confidence in managing challenges behaviours and situations had increased resulting from the training.

What are the perceived impacts of the IY®-TCM programme on teachers' work-related stress?

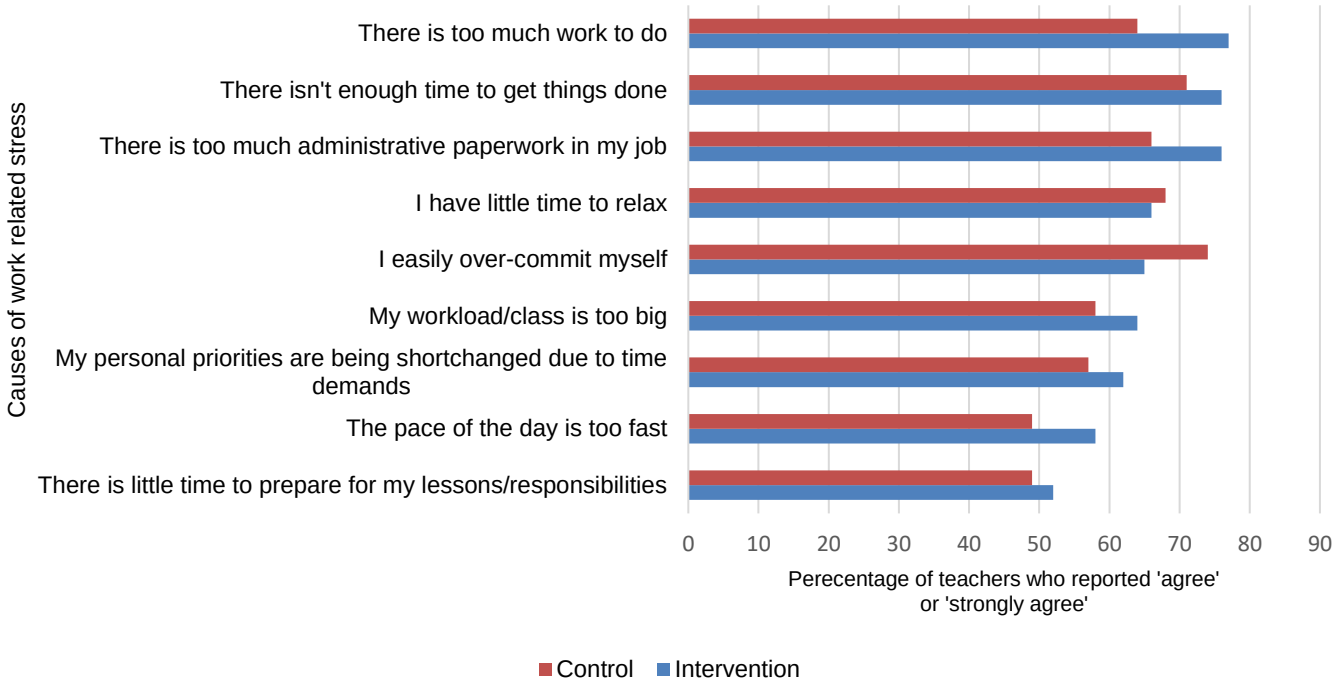
The programme ToC intends that the improvement in pupils' behaviour and teachers' increased ability to manage challenging behaviour will lead to a reduction in their work-related stress. The findings from the endpoint survey and interviews with teachers go some way to supporting the ToC.

Causes of work-related stress

The survey measured control and intervention group teachers' perceptions of work-related stress through asking them, at baseline and endpoint, to indicate the extent to which they agreed with a series of statements about their experiences at work. Figure 11 shows the comparison between the percentages of teachers in both groups who agreed or strongly agreed at endpoint that the situations were a cause of stress for them. For all but two causes of stress ('I have little time to relax' and 'I easily over-commit myself'), more teachers in the intervention group than control group reported that the situations were a cause of stress for them. Additionally, the proportion of intervention group teachers who agreed or strongly agreed with each statement was over 50 per cent, suggesting that work related stress amongst the respondents was high.

These findings do not support the ToC as they do not suggest that participation in IY®-TCM helped to reduce work-related stress in the intervention group.

Figure 11: Comparison of the causes of work related stress between intervention and control group teachers at endpoint



Data from the endpoint survey. *To what extent do you agree with each of these statements about your experiences at work?* (Control: N = 69; Intervention N: = 101). Items from the survey re-ordered to be presented in descending order.

Analysis was conducted on 'matched' teachers (those that responded to both baseline and endpoint surveys) that responded to the items on efficacy (intervention N = 55, control N = 83) to explore any differences in teachers' growth in self-efficacy over the trial period. The evaluation team converted agreement with each of the 11 items to a numeric scale from –4 to 4, with higher numbers corresponding to greater agreement with the (positively worded) items. The items were then summed to form composite scores and each respondent's baseline composite score was deducted from their endpoint composite score, so that the resulting measure had a theoretical range of -88 to 88, with positive numbers indicating a general improvement in self-efficacy from baseline.

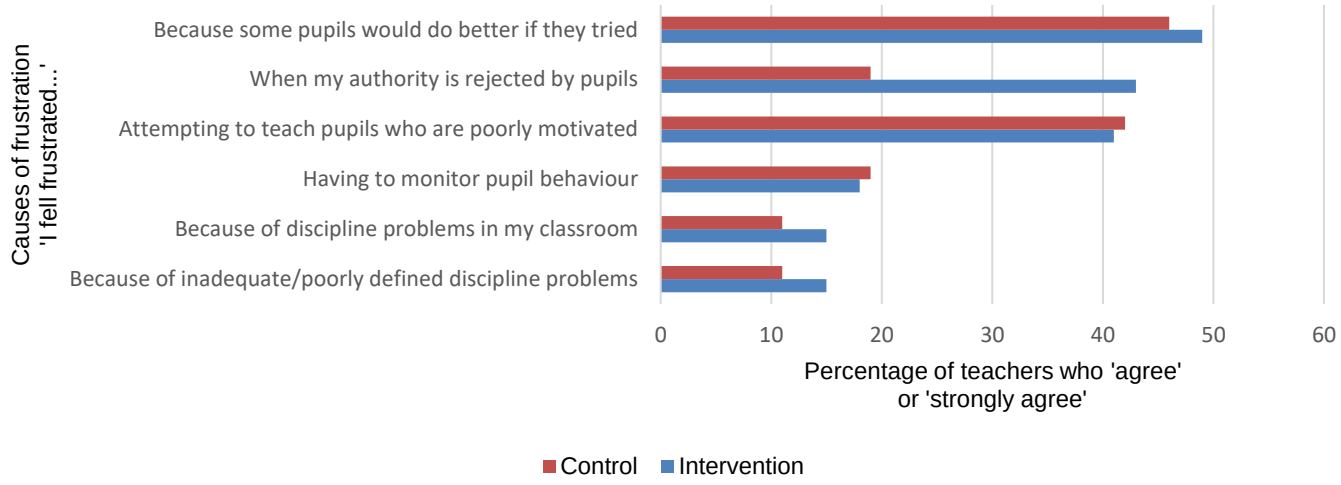
As with the analysis conducted for teacher self-efficacy, additional analysis was conducted on matched teachers that responded to the items on work-related stress (intervention N = 55, control N = 83). Each item on the Teacher Concerns Inventory was converted to a numeric scale from -2 to 2, with higher numbers corresponding to disagreement with the item statement (i.e., a higher score is good as the original item statements were negatively worded so a higher score indicates the item is not a cause of stress for the teacher). The items were then summed to form composite scores and each respondent's baseline composite score was deducted from their endpoint composite score so that the resulting measure had a theoretical range of -36 to 36, with positive numbers indicating a general improvement in work-related stress from baseline. The improvement from baseline amongst control and intervention respondents was then compared using a Mann-Whitney U test. This analysis showed that on average, teachers who participated in the IY®-TCM training saw greater improvement from baseline than those assigned to the control (mean 0.67 points) in terms of their ability to cope with time management and work-related stress. However, this finding was not statistically significant (p = 0.520) and so does not provide evidence that implementation of IY®-TCM in the classroom had a positive impact on teachers' self-reported time management and work-related stress. This finding could indicate some intervention bias in self-reporting.

Frustration related to behaviour management

The baseline and endpoint surveys asked teachers in the intervention and control groups to indicate the extent to which six statements relating to behaviour management were a cause of frustration for them. Figure 12 shows the comparison between intervention and control groups teachers' causes of frustration at endpoint.

As shown in the figure, for all but two causes of frustration ('attempting to teach pupils who are poorly motivated' and 'having to monitor pupil behaviour'), more teachers in the intervention group reported feeling frustrated with the situations presented compared to teachers in the control group.

Figure 12: Comparison of the causes of intervention and control teachers' causes of frustration at endpoint



Notes: Data from the endpoint survey: *To what extent do you agree with each of these statements? I feel frustrated…'* (Control: N = 69; Intervention: N = 101). Items from the survey re-ordered to be presented in descending order.

Analysis was conducted on matched teachers that responded to the items on behaviour management (intervention N=53, control N=82). Each item was converted to a numeric scale from -2 to 2, with higher numbers corresponding to disagreement with the item statement (i.e., a higher score is good as the original item statements were negatively worded). The items were then summed to form composite scores and each respondent's baseline composite score was deducted from their endpoint composite score so that the resulting measure had a theoretical range of -24 to 24, with positive numbers indicating a general improvement in behaviour management. The improvement from baseline amongst control and intervention respondents was then compared using a Mann-Whitney U test. On average, across all six sources of frustration, teachers in the intervention group saw greater improvement from baseline than those assigned to the control group (mean 1.05 points) in terms of decreasing frustration with student discipline and motivation. However, this difference was not statistically significant (p = 0.379) meaning that the evaluation team cannot firmly conclude that participation in IY®-TCM was associated with teachers feeling less frustrated with pupils' behaviour.

Teachers who took part in the interviews were asked about the perceived impact that implementation of IY®-TCM had on their work-related stress. Seven of the 15 teachers reported reduced work-related stress resulting from improved behaviour management in the classroom. These teachers reported that when poor behaviour did arise, they could now deal with the situation calmly through having constructive conversations with children to de-escalate the situation. They reported that they felt calmer and had more patience, which had two results – teachers reported that children picked up on their calmness which overall led to a calmer and more positive classroom climate. Teachers also reported that because they felt less frustrated and did not raise their voice, their wellbeing improved. In addition, these teachers reported feeling less overwhelmed by their job overall and when confronted with challenging behaviour.

What are the impacts of the IY®-TCM programme on school behaviour policies?

The programme ToC states that schools' participation in IY®-TCM will lead to improved school-wide behaviour policies, however only a small number of survey respondents reported that their school behaviour policy was updated following their participation in the programme.

Just over half of the 98 survey respondents in the intervention group (54%) reported that their school's behaviour management policy had changed since November 2019. In total, twenty-five teachers reported that this change resulted from their school's involvement in the IY®-TCM programme, suggesting that this intended outcome stated on the ToC was largely not met.

Teachers and trainers highlighted that support from headteachers was a vital factor in facilitating teachers to implement the IY®-TCM approach within their classrooms. They commented that, for the programme to have an impact, headteachers would need to embed the approach as part of the whole school's behaviour management practice so that all pupils receive the same message about behavioural expectations. Trainers reported that some teachers went back to their schools feeling empowered to present the approach to their headteacher and support with implementing the techniques across their schools, however they did not receive the buy-in from their headteacher or the support to alter the behaviour policy to incorporate IY®-TCM. One trainer suggested that inviting headteachers to some of the training sessions would help overcome this barrier, as they would better understand the premise of the IY®-TCM approach and when back in school, support teachers with implementation of the techniques. The ToC for this trial does not explicitly highlight the role of senior leaders however the school leadership structure, existing school policies and influences from local and national policy were considered contextual factors in the original trial led by the developer team (Ford *et al.*, 2019; see Appendix E for the logic model for original STARS trial). These findings suggest that direct involvement of head teachers or other senior leaders in the intervention, at some level, is an important mechanism for achieving change and should have been a factor included in the ToC. The ToC also did not anticipate the issue of current school policy preventing teachers fully implementing the approach however it was highlighted in the STARS logic model.

The survey asked teachers to comment upon how the changes to the school policy (because of IY®-TCM, Covid-19 or another reason) impacted their ability to implement the IY®-TCM approach. 27 of the 45 teachers who responded said that the change positively impacted their ability to implement the approach and 15 teachers said the change had no impact upon this. These findings suggest that teachers were able to continue implementing the programme within their classrooms despite changes to their school's behaviour policy.

Of the 38 teachers who said their behaviour policy had not changed since November 2019, 15 said the policy had no impact on their ability to implement IY®-TCM and 13 said their school policy positively impacted their ability to implement IY®-TCM. Only a small proportion of teachers (three) said their school policy negatively impacted their ability to implement IY®-TCM.

Interviews with teachers offer insight into the changes schools made to incorporate the IY®-TCM approach and the impact this had upon the wider school. Eight teachers from five schools said that their school behaviour policies had been updated to incorporate the principles of IY®-TCM. As described in the *'Responsiveness'* section of this report, these eight teachers reported that their headteacher was receptive to the IY®-TCM techniques and welcomed the chance to update their policy. This meant that all pupils throughout the school would receive the IY®-TCM approach.

These changes in line with the IY®-TCM approach manifested in different ways across the schools. Teachers reported that the policies now emphasised the use of positive reward systems, which featured in every classroom, reflecting the IY®-TCM focus on positive behaviours. Teachers also reported that they no longer used some systems of dealing with non-compliant behaviour, because IY® encourages teachers to ignore or redirect challenging behaviour. This meant schools discarded strategies such as sending children to the headteacher or another member of SLT or writing pupils' names on the board if they misbehaved. Teachers also reported that visual displays where they moved children's names down for negative behaviour no longer featured in their classrooms. (Further detail on the behaviour management systems teachers used prior to the training can be found in the *'Programme differentiation'* section of this report.)

Two teachers from different schools reported that their schools coached children to display positive behaviours through talking through with children the trigger of negative behaviours and what they can do to deal with the trigger in a constructive way in the future, for example walking away from a peer who is frustrating them.

Teachers reported that these changes had led to improved behaviour across their schools and that children received a consistent message of behaviour from all teachers and support staff.

'It [the school is] a calmer place to be. The behaviour is better and across the school, we are more consistent with our behaviour management approaches and supportive of other classes.' (Year 2 teacher)

How have the impacts been affected by Covid-19?

During the interviews, most teachers felt that, had implementation of IY®-TCM not suffered the disruptions of partial school closures, the impacts they recognised for themselves, and pupils would have been greater.

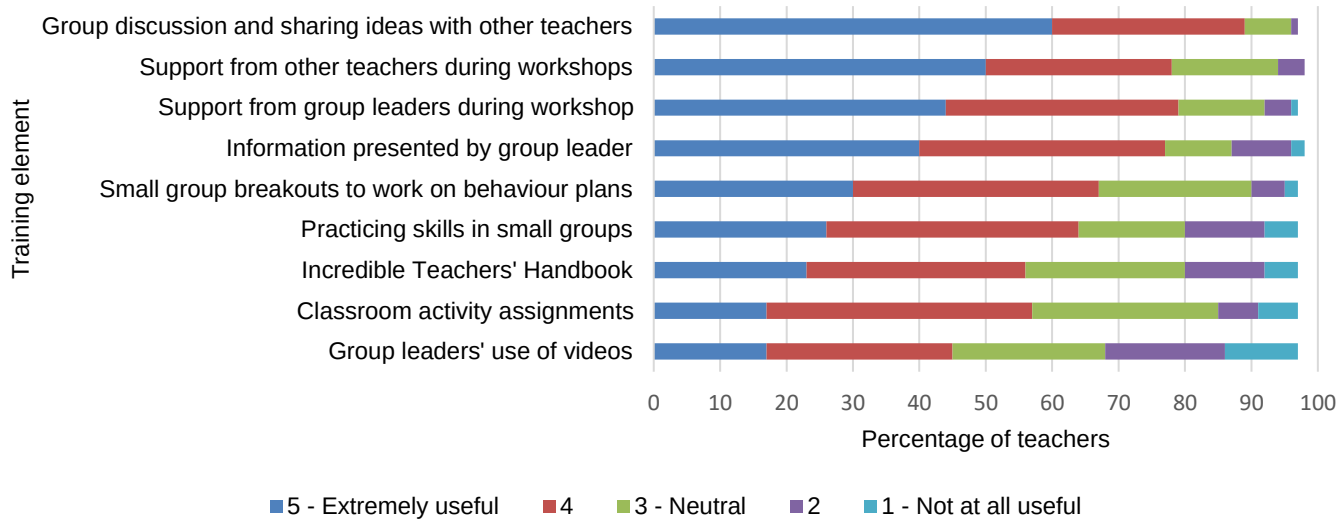
Responsiveness

How do school staff respond to the IY®-TCM programme*?*

Data from the interviews with teachers and the IY®-TCM trainers provided evidence that teachers' response to the IY®-TCM programme differed. Although most teachers involved in the interviews gave positive feedback on IY®-TCM, they, along with the trainers, highlighted several reasons why some teachers had a more negative response to the programme.

The endpoint survey asked participating teachers to indicate how useful they found nine elements of the IY®-TCM training on a scale from one (not at all useful) to five (extremely useful). Teachers found group discussions and sharing ideas with other teachers the most effective aspect of the training, with 89 per cent of the 82 responding teachers rating this as a four or five on the scale. In comparison, teachers found the use of video examples least useful. Figure 13 shows how useful teachers found each element of the training. The interviews with teachers and trainers offered further insight into these findings.

Figure 13: How useful teachers found elements of the IY-TCM training



Notes: Data from the endpoint survey: *Thinking back to the STARS: Incredible Years® Teacher Classroom Management Training, how useful did you find the…'* (N = 82). Items from the survey re-ordered to be presented in descending order. Percentages have been round to integers so may not sum to 100.

Engagement during the training

Most teachers said in the interviews that they welcomed the opportunity to discuss the behavioural challenges they faced within their classroom with other teachers during the sessions, who had a range of experiences. These conversations helped teachers to realise they were not alone in dealing with challenging behaviour, provided them with the chance to share practice and gain new ideas for strategies they could try in their own classes. Both the trainers and teachers said that a supportive group dynamic facilitated these conversations, because teachers felt they could be open and honest about the issues they faced.

However, some teachers reported a more challenging dynamic in their groups, especially at the earlier training sessions. This appeared to come about from some teachers responding negatively to their involvement in the programme. Interviewees reported that some teachers felt they had been nominated to attend by their headteachers because their behaviour management required improvement, leading to them being defensive and that some more experienced teachers were unwilling to change their existing practice. This reflects the importance of developing good relationships with participating schools from the outset and providing them with clear information about the aims and objectives of programme. In a few cases, interviewees reported that these issues led some participants to challenge the trainer or disengage from the activities. Trainers identified that the root of this problem was lack of communication between the

headteacher, who had agreed to participate in the trial, and the Year 1 and Year 2 teachers, who did not understand the reason for their involvement. One trainer commented:

'I think that there was a little bit of resistance at the beginning of the programme by some of the people that attended…We felt that some of the principles and practice, there was a little bit of scepticism. They were not sure that it would fit with what they were doing, and also, I think a little bit of fear of: "Is this going to draw up things that we don't want to change?".'

Two of the three trainers interviewed reported challenges engaging teachers in the training, particularly during the first sessions, because teachers said they were already practicing many of the techniques the training promoted, such as the importance of remaining positive and building relationships with pupils. They reported that teachers, particularly those with more experience, felt the material was 'too low level' for them, which led to the recommendation from five of the teachers interviewed that teachers new to the profession (e.g., ECTs) should be targeted for the course. However, trainers felt they were able to overcome this perception through pitching the programme in a way that made teachers realise that the first session set the foundations of IY®-TCM and that further training sessions would build upon and strengthen their knowledge of behaviour management.

Engagement with the training content

Teachers' perception of the quality of the training content and resources affected how they responded to the programme. Interviews with teachers in case-study schools provided insights into where they saw room for improvement. Some teachers commented that the Incredible Teachers' handbook chapters were long and the amount of reading expected between sessions was not manageable. They suggested that a condensed version of the handbook featuring essential reading would have been more helpful. A few teachers also questioned the usefulness of the handouts provided, because they had not felt the need to refer to them since the training.

Teachers rated the video exemplars as the least useful aspect of the training. The teachers who participated in the case-study interviews explained why they thought this. The videos featured footage of behaviour management practice filmed in kindergarten settings in America in the 1990s, so teachers considered them outdated and lacking relevance for their classroom context. Despite the fact that trainers tried to contextualise the content and make it applicable to the current UK context, some teachers still criticised the programme for being outdated and had difficultly engaging in the programme as a result:

'That part of the course definitely needs to be changed, it needs to have a British curriculum, and they need to be much more modern. The ideas were sound, but the videos were awful.' (Year 1 teacher)

To what extent are teachers implementing the strategies?

Detail on the extent to which teachers implemented the IY®-TCM techniques in their classrooms can be found with the *'Dosage'* section of this report.

Did teachers share any of the STARS training/materials with other staff in school who did not attend training?

The intention was for the intervention cohort of pupils to be taught by an IY®-TCM trained teacher for two consecutive academic years. Neither teaching assistants nor teachers of other year groups directly received training as part of the intervention, but trainers encouraged participating teachers to share information and training materials with the teaching assistants in their class. Some teachers also chose to share the training with other class teachers.

Data from the cost section of the survey showed that 35 per cent of teachers had spent time training or disseminating training to other members of staff who had not attended the IY®-TCM training. The interviews with teachers further explored how, and to what extent, teachers shared the training or materials with their colleagues.

The extent to which teachers cascaded the IY®-TCM training and shared materials differed both across and within the schools involved in the interviews. Teachers interviewed in three of the nine schools reported that together, the Year 1 and Year 2 teacher had provided training to all school staff through in-school behaviour management CPD sessions. Teachers in these schools said that their headteacher had welcomed the IY®-TCM techniques and, resulting from their participating in the intervention, would be embedding the IY®-TCM approach into the school's behaviour policy so that all teachers and support staff were using the recommended approaches. Further detail on this can be found under '*What are the impacts of the IY®-TCM programme on school behaviour policies?'* Eight teachers from five schools reported

that they had informally shared the IY®-TCM techniques with other teaching staff, including: teaching assistants; trainee teachers; partner teachers; the Year 3 teacher and the headteacher. Teachers in one school said they had not shared the training to other staff because their school already used many of the behaviour management techniques promoted by IY®-TCM.

To what extent are the approaches advocated congruent with teachers/schools' existing beliefs about effective practice?

The IY®-TCM techniques mostly aligned to teachers views on effective behaviour management practice. Teachers involved in the interviews agreed with the principles of IY®-TCM, such as the emphasis upon the importance of gaining a holistic understanding the child to help to inform reasons for challenging behaviour, and the value of focussing on positive behaviour while ignoring or redirecting negative behaviour. Some teachers commented that although they already used and knew the importance of some techniques, IY®-TCM provided them with a greater understanding of the theory behind why these techniques work for managing behaviour.

What are participants' views about the suitability of the intervention?

As part of the final interviews in summer 2021, the evaluation team asked the five participating Year 2 teachers whether they would recommend IY®-TCM to other teachers or schools. All said they would recommend at least some of the IY®-TCM approach. One teacher commented that IY®-TCM was 'one of the best CPD programmes' they had attended. They said it encouraged them to work with the children to understand the reasons why they were presenting negative behaviours, rather than deal with behaviours in a punitive way. This meant they could get to the root of the problem and prevent the behaviour from recurring.

> *'It completely changes your ideas around behaviour so instead of being focussed on those negatives, it switches more to the positives and how you can prevent and stop it [poor behaviour] happening before it does become a real big problem.'* (Year 2 teacher)

Some teachers involved in interviews earlier on in the trial felt that the programme would be most effective for teachers of Early Years and Year 1 classes based on the nature of the techniques. They felt younger children would respond better to the approach compared to older pupils, for whom the programme would require adaptation to be suitable and effective. In addition, some more experienced teachers suggested that the training would be better targeted at NQTs rather than experienced teachers.

> *'It wasn't appropriate for us… a lot of what we were shown and were told we implemented anyway… If I'd done it as an NQT I'd be thinking great, this is fantastic, it's really given me some fantastic ideas.'* (Year 1 teacher)

However, this was not a universal view. Other experienced teachers said they appreciated the opportunity to attend training which reinforced they were dealing behaviour in the correct way through using positive reinforcements and gained additional techniques on how to deal with behaviour in a positive way.

What additional support is required to engage with the IY®-TCM programme and its advocated school improvement practices?

Trainers were not expected to provide teachers with extensive additional support outside of the training sessions but those who were interviewed were willing to provide support and advice, where required, to ensure teachers felt confident in delivering IY®-TCM. Interviews with the developer, trainers and teachers explored the type, and amount, of additional support provided.

The trainers and the developer provided some general implementation support to participating teachers. The trainers involved in the interviews said they emailed teachers with reminders of the techniques covered in each session and used a mobile messaging app to facilitate ongoing communication among the group in-between the sessions. Teachers reported that they used the messaging app to share how implementation of techniques was working in practice. In cases where a teacher experienced challenges, the trainer offered support and advice on how they could implement a technique differently. Teachers also reported that they shared reading or articles that they thought would be of interest to the group.

Although not intended as part of the delivery model, one of the trainers reported that if requested, they would have been willing to visit schools to observe the teacher's practice and provide tailored advice on how to implement the techniques based on the level of challenging behaviour the teacher was dealing with. Two of the teachers interviewed also commented that they would have welcomed this opportunity. IY® developers now recommend active coaching alongside the six sessions to participating schools.

Adaptations

What adaptations were made to the IY®-TCM training delivery model?

The developer planned for the IY®-TCM training to be delivered to teachers in 18 training groups. However, due to low attendance in two in different hub areas, the developer cancelled these two groups and the teachers who continued with the programme received travel expenses to attend different groups within their area.

The developer also reported a change to the timing at which they provided support to the trainers.The developer planned to meet with all trainers half way through delivery of the IY®-TCM training, yet due to the issues some trainers raised around teachers' engagement and response after the first session in a few of the hub areas, they held this after session two. The developer reported that it was the experienced trainers who raised these concerns, as they had not been confronted with the same challenges when delivering the programme previously. It appeared that these issues had only arisen in two of the 18 original groups and the developer held regular check-ins with these trainers to monitor the situation.

Did trainers adapt their delivery of IY®-TCM training workshops and if so, why?

The evaluation team asked the three trainers who participated in interviews if they had made any changes when delivering the IY®-TCM training workshops and if so, their reasons for these changes. As reported in the fidelity section of this report, there was high fidelity to the intended training approach and only few, permissible adaptions were made. Only one trainer reported that they adapted their delivery by giving more focus to the reflective elements of the sessions because this enabled teachers to develop a deeper awareness of their practice and consider how they could embed the IY®-TCM into this to have impact for themselves and their pupils. The trainer said that teachers were then more willing to implement the approach in their classrooms and from this, saw a real change to their practice over the course of the training sessions.

Another trainer reported that they did feel empowered to adapt the training delivery if necessary and reported that on reflection, they could have moved through the first session on relationship building quicker as teachers found this content too basic, which led to low engagement for some.

During the training observations, the evaluation team identified that there was evidence of trainers 'adapting with fidelity' (i.e., making adaptations allowed by the programme), focusing more in-depth on areas where teachers in the groups had specific support needs, for example in relation to their 'target child/children'. The observations did not identify any other areas of adaptation from the recommended approach.

How have schools adapted IY®-TCM and why?

During the interviews in autumn 2020 and summer 2021, the evaluation team asked teachers whether they had adapted any of the IY®-TCM techniques and if so, their reason for this. Only four of the 15 teachers involved in the interviews reported they had adapted the recommended approach and the adaptations they made were all permissible.

Two teachers reported that they had not used some of the techniques the training promoted because they did not work well with the pupils. For example, one Year 2 teacher reported that pupils found the coaching aspect of the approach patronising and they did not respond well to this, so the teacher did not persist with the technique. Two teachers reported that, in order for the rewards techniques to fit with the school policy, they had to adapt how they issued rewards. In both schools, the behaviour policy was for teachers to reward the whole class rather than individual children. This meant that teachers could only give out rewards when pupils had collectively displayed positives behaviours, rather than in response to individual pupils behaving well.

The evaluation team also asked trainers and the developer if they were aware of any changes teachers made to the approach. They reported that the only changes that they were aware of were adapting the ideas demonstrated in some

of the video exemplars to make them more appropriate for the English school context. For example, rather than giving out sweets as rewards for positive behaviour, teachers gave pupils stickers or behaviour points instead.

These findings suggest that for the most part, teachers implemented the programme with high fidelity by following the approach as taught in the training sessions and as set out in the 'Incredible Teachers' handbook.

What challenges did schools/teachers face in implementing IY®-TCM and how did they overcome these challenges?

Most of the 15 teachers involved in the interviews said they had not experienced any challenges with implementing the IY®-TCM approach, either because it aligned to their school policy or because they had support from the SLT to implement the approach. Six teachers reported that they had experienced some challenges when implementing the IY®-TCM approach. These are outlined below, along with the solutions teachers found to overcome the challenges.

Four teachers said that aspects of the IY®-TCM approach did not align to their school policy. Teachers in two schools commented on the use of rewards (see above for detail on this adaptation). The teachers in one school raised the challenge of implementing sanctions. IY®-TCM encourages teachers to be selective about when they issue sanctions, such as loss of privileges or timeout, however teachers reported that they could not use these techniques due to inconsistencies within their school's behaviour policy.

Two teachers from different schools reported the challenge of changing the mindsets of other staff in school who had not attended the training. One teacher experienced resistance from their experienced TA who believed in the use of punitive approaches, while another experienced challenges from other teachers. Teachers reported that they had tried to overcome these challenges and promote the approach through: being positive about their experience of the training; modelling how the techniques should be used within the classroom; and sharing their positive experiences of implementing the strategies.

One teacher reported challenges around the principle of all children receiving rewards. When sharing the training with colleagues, they raised questions about how to ensure all children have equal access to the rewards for positive behaviours. The teachers who attended the IY®-TCM training encouraged colleagues not to overlook the children who behave consistently well and to ensure they receive the same level of praise and rewards as their peers. Secondly, the teacher encouraged their colleagues to make a conscious effort to identify positive behaviours from children who normally misbehave.

One teacher commented that the IY®-TCM did not work as well with the most challenging pupils in their class. This teacher recognised that the approach would not be suitable to fully support them with these children but said that they persevered with using the techniques with the aim of having some impact, such as the pupils being calmer and their emotions and behaviours escalating less frequently.

None of the teachers involved in the interviews reported challenges around their headteacher or SLT not supporting them to implement the IY®-TCM, however teachers and trainers picked up on teachers from other schools who were facing these difficulties. This challenge is further detailed under *'What are the impacts of the IY®-TCM programme on school behaviour policies?*

What were the main Covid-19 related adaptations?

As mentioned earlier, teachers did not report in the interviews that they had made any adaptations to the IY®-TCM approach in response to Covid-19 although they did adopt the IY®-TCM techniques, which are intended for face-to-face use with pupils in the classroom, as part of their remote delivery practices.

The developer reported that the only adaptation to delivery of the training sessions in response to Covid-19 was the cancellation of the sixth session for two training groups as these were planned for dates immediately before or after schools had to close to most of their pupils. This adaptation is detailed in the *'Compliance'* section of this report.

Cost evaluation

Time costs

Of the 80 intervention teachers who responded to the cost questions of the practitioner survey, 73 per cent[25] reported that they had spent some time on preparing to implement IY®-TCM in addition to attending the workshops. In terms of implementing the IY®-TCM strategies, 44 per cent of teachers responding reported that they spent additional time implementing the programme outside of normal planning and administration time. Table 16 shows the mean amount of time spent on different activities for the 80 responding teachers. Time spent on attending the workshops have been included in the financial costs table rather than here because more than half the responding teachers reported that their time out of the classroom was covered by supply teachers. While only one year of the intervention has taken place, the costs are shown as the costs of continuing the intervention for a further two years in schools as it is assumed that any impacts from the training would continue over the coming years.

[25] Sixty-six per cent answered 'yes' to this question, a further six per cent responded 'don't know' but then entered numbers when asked how many hours they spent on the preparation activities so have been included in the overall percentage.

Table 16: Additional time spent on training and implementing IY®-TCM

			Year 1			Year 2			Year 3		
			Number of teachers	Mean Number of hours per week (min, max) (n=80)		Number of teachers	Mean Number of hours per week (min, max)		Number of teachers	Mean Number of hours per week (min, max)	
Preparation	reading handbook	classroom teacher	3.1	3.25	(0, 25)	0	0	n/a	0	0	n/a
	trialling techniques	classroom teacher	3.1	5.71	(0, 50)	0	0	n/a	0	0	n/a
	adapting techniques	classroom teacher	3.1	2.94	(0, 20)	0	0	n/a	0	0	n/a
	other training activities outside of workshops	classroom teacher	3.1	1.56	(0, 15)	0	0	n/a	0	0	n/a
Delivery	planning specific classroom activities	classroom teacher	3.1	1.48	(0, 30)	3.1	1.48	(0, 30)	3.1	1.48	(0, 30)
	organising/creating resources	classroom teacher	3.1	1.50	(0, 20)	3.1	1.50	(0, 20)	3.1	1.50	(0, 20)
	following up behaviour	classroom teacher	3.1	2.94	(0, 50)	3.1	2.94	(0, 50)	3.1	2.94	(0, 50)
	training disseminating training to other staff	classroom teacher	3.1	1.54	(0, 20)	0	0	n/a	0	0	n/a
	re-writing the school behaviour policies	classroom teacher	3.1	0.44	(0, 5)	0	0	n/a	0	0	n/a
	other activities relating to implementing the programme	classroom teacher	3.1	0.65	(0, 10)	0	0	n/a	0	0	n/a

Financial costs

A total of 36 teachers provided data on additional costs around travel and the mean average of these costs are shown in Table 17. Of the remaining 43, 13 signalled that there were no associated travel costs with attending the training[26]. Of those responding (n=49), 73 per cent reported that there was a cost to the school for training. Four teachers reported spending their own money on travel expenses.

To enable teachers to attend the workshops, 32 teachers (64%) of responding teachers reported that there was a financial cost of supply cover. Nine teachers said that their time attending the workshops was covered by internal staff (e.g., Higher Level Teaching Assistants) at least some of the time.

Table 17: Additional financial cost spent on training IY®-TCM

		Start-up or Recurring?	***n***	**Mean Cost (min, max)***
PART A: means from those teachers who said there was an additional financial cost				
Personnel for training	cover	start-up	32	2559 (563, 5633)
Training and programme costs	travel	start-up	36	220 (19, 939)
PART B: means from all teachers who responded (i.e., including those who reported that there was no additional cost)				
Personnel for training	cover	Start-up	50	1638 (0, 5633)
Training and programme costs	travel	start-up	49	162 , 939)

* The cost per schools assuming that 3.1 teachers attended all 6 of the workshops. 3.1 is the average number of teachers attending at least 1 workshop per intervention school.

In terms of implementing the programme, 35 (out of the 77, 46%) teachers responded they didn't know whether their school had any extra financial costs associated with implementing the programme and a further 29 (38%) of teachers stated there were no additional costs. Of the remaining 13 teachers, only three reported any costs relating to implementation of the programme. There were all related to purchasing of resources for use in the classroom.

Table 18 shows that the financial costs to the school are £12 per pupil per year over a three-year period. The additional costs of delivering IY®-TCM which were not borne by schools are summarised in Table 19. These show an additional estimated cost of £14 per pupil to run IY®-TCM if the costs are spread over four cohorts[27] of pupils. The cost per pupil per year over three years is estimated as £27; IY®-TCM is therefore deemed a very low-cost intervention (see Appendix A), supporting results from the cost-effectiveness element of the original STARS trial (Ford *et al.*, 2018).

[26] The remaining 30 teachers either reported that they did not know (23) or did not provide an answer/valid answer (7).
[27] Four cohorts of pupils will benefit from the programme if it is run over three consecutive years as it was in the trial as two years are exposed to trained teachers in the same academic year.

Table 18: Additional financial cost spent on training IY®-TCM by schools

			Nominal Values			**Real Values (Deflate using Y1 as Base)**			**Present value**	**Cost in Analysis Year**
		Start up or Recurring?	£ Year 2019	£ Year 2020	£ Year 2021	£ Year 1 (in 2019 prices)	£ Year 2 (in 2019 prices)	£ Year 3 (in 2019 prices)	£PV (in 2019)	£PV (in 2021)
Personnel for training	Teacher cover	Start-up	1638			1638			1638	1730
Training and programme costs	Travel fees for training	Start-up	162			162			162	171
Total cost per school										**1901**
Number of pupils-per-school-year *										77
Cost per pupil-school-year **										**12**

* This figure is double the number of pupils in the trial cohort as the Year 2 pupils would also have been exposed to the programme in the first year of the trial.

** Over a 3-year period, 4 cohorts of pupils would have been exposed to the programme (as set up for the trial) as there are 2 cohorts in the first year so the denominator for this calculation is 4 x 39 pupils.

Table 19: Additional financial cost of IY®-TCM by the developer (market value)

			Nominal Values			**Real Values (Deflate using Y1 as Base)**			**Present value**	**Cost in Analysis Year**
		Start up or Recurring?	£ Year 2019	£ Year 2020	£ Year 2021	£ Year 1 (in 2019 prices)	£ Year 2 (in 2019 prices)	£ Year 3 (in 2019 prices)	£PV (in 2019)	£PV (in 2021)
Personnel (intervention management and administrative time)	Organising the training for the Group Leaders	start-up*	388			388			388	409
	Administration and coordination of intervention delivery	start-up	2855			2855			2855	3015
Personnel for training	Group Leader costs	start-up	89645			89645			89645	94665
	Travel and subsistence	start-up	7284			7284			7284	7692
Training and programme costs	Training the Group Leaders	start-up*	4569			4569			4569	4824
Facilities, equipment and materials	Venue hire	start-up	11391			11391			11391	12029
	Catering	start-up	6616			6616			6616	6987
	Teacher books and equipment	start-up	6720			6720			6720	7096
	IY handbook and equipment for Group Leaders	start-up*	18000			18000			18000	19008
Total cost										**155726**
Number of pupils-per-year **										5420
Cost per pupil-school-year ***										**14**
*	These start-up costs would not need to be repeated if the Group Leaders ran the training with new schools									
**	This figure is double the number of pupils in the trial cohort (2710 pupils) as the Year 2 pupils would also have been exposed to the programme in the first year of the trial.									
***	Over a 3-year period, 4 cohorts of pupils would have been exposed to the programme (as set up for the trial) as there are 2 cohorts in the first year so the denominator for this calculation is 4 x 2710 pupils.									

Conclusion

Table 20: Key conclusions

Key conclusions
6. Due to Covid-19, the primary outcome for this evaluation was not collected, so no measure of impact on KS1 maths attainment is reported and no padlocks are assigned. Key conclusions are based on secondary outcome measures and qualitative data from the implementation and process evaluation.
7. The secondary outcome measures found no evidence that the IY®-TCM programme influences behaviour, social and emotional wellbeing, concentration, prosocial behaviour or student-teacher relationships for pupils receiving the intervention compared to pupils in the control group. These results are based on surveys completed by Year 2 teachers at the end of the project, comparing the intervention and control group. However, due to low response rates, they are based on 55 per cent of pupils in the trial. Furthermore, 21 per cent of the pupils in the second year of the trial were taught by teachers who were not on the training register for the IY®-TCM training, so these pupils did not have any exposure to the trained teachers when in Year 2[28].
8. Only 60 per cent of the intervention teachers attended four or more of the training sessions. Pupils of those teachers demonstrated better behaviour, compared to the control group, using surveys completed by classroom teachers at the end of the trial. This impact was small and there may be other differences between teachers that attended sessions and those that did not which could contribute to the difference in impact.
9. Teachers who attended the training showed larger improvement in perceived self-efficacy compared to those in the control schools, when measured through surveys at the start and end of the trial. However, due to low response rates, this is based on fewer than half of teachers in the trial.
10. Findings from interviews with teachers, the developer team, and trainers suggest that the combination of the costs to schools in releasing all Year 1 and Year 2 teachers and covering for their absence from the classroom was a factor in the low attendance of training sessions, when considered alongside the perceived value of the programme in addressing schools' needs.

Impact evaluation and IPE integration

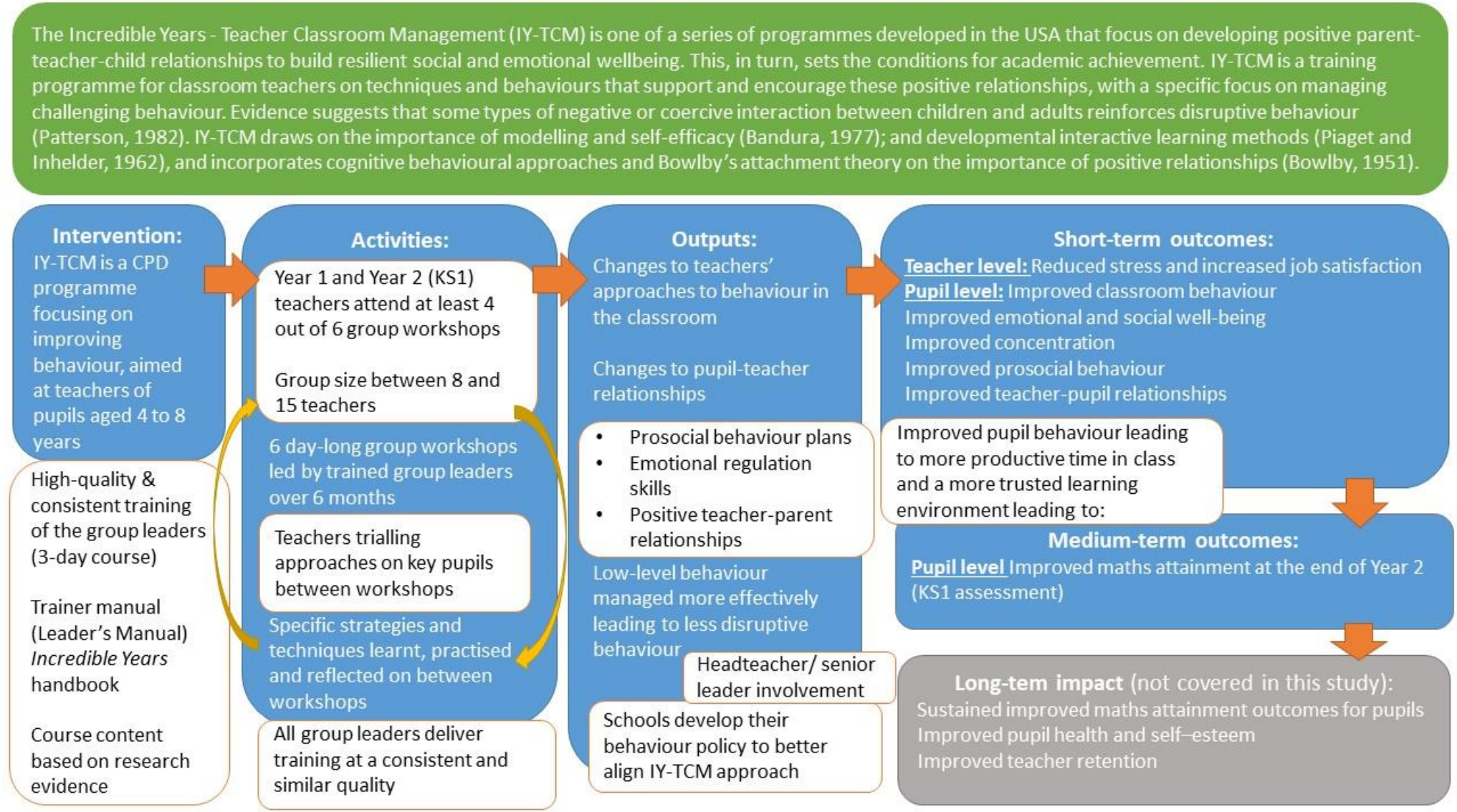
Evidence to support the Theory of Change

The evaluation was not able to provide any evidence as to whether there is an impact on pupils' maths attainment as the end of key stage assessments for Key Stage 1 were cancelled for summer 2021. The secondary analysis examined pupil behaviour, social and emotional well-being, concentration and prosocial behaviour as well as student-teacher relationships and did not find any evidence of an impact of the IY®-TCM programme on pupils in the intervention schools as per of the intention-to-treat analysis. In line with the Theory of Change, however, there was evidence that supports the attendance requirements as teachers attending at least two thirds of the training sessions had pupils who experienced a positive impact on behaviour.

While it was not possible to provide evidence from the impact measures, there were some findings from the IPE which support the initial stages of the causal chain of hypothesised impact in the Theory of Change. For example, there is some evidence from teachers who responded to both the baseline and endpoint surveys that those who attended training had seen larger improvements in their perceived self-efficacy compared to the control group. Similarly, there was also some evidence (although again self-reported by those teachers who responded to both surveys) that behaviour management had improved more for the intervention teachers than the control over the trial period. There is a lack of evidence from the IPE to suggest the intervention leads to 'more productive time in class' and a 'more trusting learning environment' but teachers were unable to deliver in their classrooms in the usual way for a significant duration during the trial. An updated Theory of Change is provided in Figure 14.

[28] The proportion of pupils in Year 2 taught by a teacher who had not been trained in IY®-TCM is likely to be slightly higher than 21 per cent as some teachers were on the training register but did not attend any workshops.

Figure 14: Updated Theory of Change for IY®-TCM (mediators and moderators in white boxes)



Further evidence from the IPE found that there are implicit assumptions about the intervention that are not made explicit in the Theory of Change and that may not hold across all schools. Firstly, is the idea that behaviour management issues in schools are largely to do with low-level disruptive behaviours. If individual schools do not agree with this assumption, the intervention is not considered useful. Responses from the practitioner survey suggested that teachers feel that their schools generally have effective strategies to address low-level disruption in their classrooms but often lack support for pupils with higher-level behavioural needs. The intervention does not attempt to address the more severe behavioural needs that some teachers were reporting. In the original trial there were mixed reports by teachers as to how effective the TCM training was in the management of more challenging pupils (Allen *et al*., 2019), although two groups were successfully delivered with pupils with SEND previously; one group of staff were working in alternative provision, and the other were teacher / TA dyads working with pupils in mainstream (Ford, 2016).

Findings from case-study interviews suggest that in some instances, headteachers and participating teachers did not understand what to expect from the intervention until training had started. It would have perhaps been beneficial for them to have a deeper understanding of the aims of the intervention and methods used at the recruitment stage so that an informed judgement can be made about the appropriateness of the intervention for their schools' needs and how a schools' involvement could be best communicated to participating teachers by the headteacher.

One outcome in the Theory of Change was that working with both the Year 1 and Year 2 class teachers could lead to a positive change in the school's behaviour policy. The mechanism for how involvement in the intervention may change school behaviour policies is not explicit and this is likely to involve more direct involvement from senior leaders. Only a quarter of intervention teachers, who responded to the survey, reported that changes had been made to their school behaviour policy due involvement with IY®-TCM. The lack of involvement of headteachers is likely to have impacted on the extent of whole school change, particularly in relation to changes to school behaviour policies. In other work the developer team has worked directly with teaching assistants and would recommend a whole school approach, with all staff, including lunch time and playground facilitators experiencing the training (Ford, 2016). Direct involvement of head teachers or other senior leaders in the intervention, at some level, is an important mechanism for achieving change and should have been a factor included in the ToC. Direct SLT support might amplify the effects of individual teacher's practice by supporting a whole school approach. While not emphasised within the ToC in this report, the initial STARS trial did emphasise the importance of school context and influences from local and national educational policy (Ford *et al.,* 2019; see Appendix E). In earlier work with a whole team from alternative provision, the delivery team and participants found the presence of SLT was clearly challenging for all concerned in terms of group dynamics (Ford, 2016). Separate training groups for SLT would provide an alternative, which would have the additional benefit of sharing best practice and experience across schools, which was highly valued by teachers from the earlier STARS trial (Allen *et al.*, 2019)

Interpretation

The trial cohort spent a significant proportion of the time from the end of the IY®-TCM training to end of the trial either being schooled remotely or adapting to returning to school. Schools closed for the majority of pupils from 18 March 2020 until 1 June 2020. A minority of pupils were in school during this time – children of key workers for example – although their experience of school over this period would likely have been considerably different to normal. From June to the end of the academic year schools were generally open to Year 1 pupils (the trial cohort at the time) although the offer schools could make (full time or part time for example) varied a lot between schools. Restrictions in place at the time, in order to restrict unnecessary mixing, meant it likely that schools would have operated very differently to normal over this period. Pupils not being taught by their normal teacher, being taught in smaller groups and only attending school on a part-time basis were common occurrences (Timmins, 2021). Of the teachers who responded to the survey, over 80 per cent used praise/encouragement, incentives and good commands/clear limit setting at least once a week during remote learning. Interviews with teachers also highlighted examples of where teachers had used emotional coaching and relationship building online. While the survey responses and interviews demonstrated that some IY® techniques could be used during remote learning, it is clear that school closures did impact on the amount of exposure pupils had to the intervention strategies as intended. Some teachers reported that they were able to use some of the IY®-TCM approaches to support remote learning.

It is worth noting that in the previous RCT run in England on IY®-TCM by the developer team (Ford *et al.*, 2019) there was a small positive impact of the programme on emotional and social wellbeing at nine months from the start of intervention delivery but that at later follow ups (at 18 and 30 months), this effect was not found. In this trial exposure to

a trained teacher was intended to be significantly higher as Year 2 teachers were training in IY®-TCM alongside Year 1 teachers. Given the school closures in 2020, the poor match between Year 2 teachers attending training and those teaching the eligible pupils, and the loss to follow up, this may not actually have been the case. Under this evaluation, there were midpoint measures (PBQ and SDQ) planned which would have mirrored the timings of the initial follow-up in the Ford *et al*., trial but it was not possible to collect these measures due to Covid-19. As such the finding based on the SDQ measure do not contradict the previous RCT. In terms of the main secondary outcome – the PBQ - the findings presented in this report do contradict those in the initial Ford *et al.* (2019) RCT who found at a small positive impact of the intervention on pupil behaviour at each of the follow ups. The previous trial reported considerably higher levels of attendance with 90 per cent of intervention teachers attending at least four sessions compared to 60 per cent in this trial and had much lower levels of attrition in terms of pupils and schools (*ibid*). The high level of attrition in this trial increases the probability of failing to identify an effect of the intervention if it exists (a type 2 error) for all secondary analysis.

The main analysis takes an intention to treat approach where all those schools and pupils allocated to the intervention are included in the analysis whether they attended the training or not. The CACE analysis differs in that it considers the impact of the intervention for only those pupils whose teachers complied with the intervention. Compliance at teacher level was attendance for at least four out of the six training workshops. The CACE analysis on the main secondary outcome – the PBQ - found a small statistically significant effect of the IY®-TCM. Where a pupil's Year 1 and Year 2 teacher had both attended a minimum of four out of six of the training sessions their PBQ score was lower compared to those in the comparison group indicating a better relative level of behaviour. This finding should be treated with some caution as there is an increased risk of obtaining false positives (type 1 error) the more hypotheses that are tested. The size of the effect is also relatively small (standardised effect size= -0.13 [-0.24, -0.01]). It is also worth noting that while CACE analysis provides a causal estimate for those teachers who complied, there are likely to be differences between these teachers (and their schools) and those who technically had the opportunity to attend (i.e., they were in the intervention group) but did not. It is therefore not possible to conclude from this finding that if all teachers were compelled to attend at least four sessions there would be a similar effect.

Along with school closures, the other key factor impacting on pupils' reduced exposure to the intervention strategies from a trained teacher was the low levels of teacher attendance at training sessions. Findings from the IPE suggest that targeting the intervention at teachers who are perhaps less confident or less experienced (e.g., trainee teachers or ECTs) about routine behaviour management could increase the attendance at training. A key difference between this trial and the previous STARS trial run by the developer team was that schools in the previous trial chose one teacher (from Reception to Year 4) to attend the training rather than sending all teachers teaching in Year 1 and 2. This teacher is unlikely to have been chosen at random and perhaps more likely to be a teacher who needed more support in the area of behaviour management. More consideration could also be given to making the intervention more flexible; enabling it to move away from the highly manualised approach found in this trial. Adaptations to content and materials to make them more obviously applicable to English schools could have been useful as teachers did not rate some aspects of the materials and training delivery highly.

Teachers involved in the IPE who did attend the training typically reported that the IY®-TCM impacted on themselves and their pupils in a positive way. What teachers found most effective about the intervention was support from other teachers. IY®-TCM did appear to affect teachers' self-efficacy positively though the impact of IY®-TCM on teacher stress was inconclusive. The previous study of IY®-TCM in English schools lead by the University of Exeter found no quantitative evidence of an impact of the programme on teachers' mental health, self-efficacy or work-related stress (Hayes *et al.*, 2020), although the process evaluation did indicate that teachers perceived themselves to be calmer, more confident and in control (Allen *et al*., 2019).

Limitations and lessons learned

The main limitation to this trial design was the relatively high level of attrition. Only 55 per cent of pupils randomised were included in the main analyses. Under the original design the levels of missing data were likely to have been considerably lower for the planned primary outcome as the baseline measurement was to be obtained via the NPD and the endpoint outcome was part of a statutory assessment (although the data collection was planned to be direct from schools). Although this trial was powered to detect a difference in attainment, the attrition will have affected our ability to detect effects in the secondary outcomes that we could measure.

Related to the attrition were the low levels of attendance at the training sessions. The IPE shed some light on this and highlighted that teachers were not fully aware of the aims and content of the programme before they started the training, which indicates the need to support the SLT in communication of the aims of the intervention as well as the research

process. Relatively high numbers of intervention teachers did not attend more than one or two sessions. The intervention was free to attend for the schools in the trial and the cost to schools came later through needing to release teachers to attend the training. While this was explained at recruitment, the SLT had clearly not appreciated that all teachers needed to attend all sessions, and the disruption and cost to the school that would result. The initial trial had provided supply cover and achieved much higher engagement and attendance (Ford *et al.*, 2019), which might be worth considering if large numbers of teachers were invited from the same school in future studies. An alternative model would have courses on different days so that not all teachers were out of the school on the same day.

In terms of evaluation design, it is worth noting that Year 1 and Year 2 teachers were trained together at the beginning of the trial period for the intervention schools. The normal churn of teachers between schools and year groups within schools means that some Year 2 teachers who were trained in the intervention would not actually teacher the trial cohort pupils and thus reducing compliance. This is shown in the drop in pupils who were taught by a trained teacher (who had attended at least four workshops) between Year 1 and Year 2.

Another limitation to the revised study is the reliance on teacher-reported outcome measures who were clearly not blinded to their group allocation. This was not intended at the outset as the primary outcome measure was planned to be the maths scales of the end of key stage assessment for Key Stage 1 which, although it is marked locally by teachers, is a statutory assessment and therefore completed by all eligible pupils.

The IPE is limited by the small number of observations and interviews conducted and the relatively low response rate to the teacher endpoint survey. Furthermore, elements of the data collection were only partially completed due to Covid-19, so less data was available on dosage, quality, responsiveness, programme differentiation and adaptation (from the CLASS observations and follow-up interviews) than originally anticipated. This means that insights from the IPE are limited and must be viewed as tentative – this affects the generalisability of the IPE results more broadly.

Future research and publications

There are several elements of the trial, as originally planned, that were not fulfilled due to challenges related to Covid-19 which could be considered in future research. One element would be the inclusion of the CLASS tool throughout the trial. This tool, if used at midpoint and endpoint, could enable a deeper understanding of how the strategies are adopted and used in the classroom across teachers of different levels of experience.

Given the importance of parents and home factors, the focus of the role of parents and the home-school relationship in the programme could also be an element that would be worth exploring in future studies on IY®-TCM. Evidence provided through the IPE and previous studies led by the developer team highlight that it may be worth investing in updating the training material and making it more relevant to the English context as the current materials did appear to have an impact on teacher engagement in the training.

A short report detailing the external validation of the STRS is planned (by the developer team). The data collected as part of the trial will be archived by the EEF's data archiver (FFT) and held in the EEF's data archive.

References

Allen, K., Hansford, L., Hayes, R., Allwood, M., Byford, S., Longdon, B., Price, A., & Ford, T., 2019. Teachers' perceptions of the impact of the Incredible Years® Teacher Classroom Management programme on their practice and on the social and emotional development of their pupils. *British Journal of Educational Psychology*. **https://doi.org/10.1111/bjep.12306**

Allwood, M., Allen, K., Price, A., Hayes, R., Edwards, V., Ball, S., Ukoumunne, O. C., & Ford, T., 2018. The reliability and validity of the pupil behaviour questionnaire: a child classroom behaviour assessment tool, *Emotional and Behavioural Difficulties.* DOI: 10.1080/13632752.2018.1478945.

Burge, P., Lu, H., & Phillips, W., 2019. *Understanding Teacher Retention: Using a discrete choice experiment to measure teacher retention in England*. Cambridge: RAND. [pdf] Available at: **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/958634/Understanding_Teacher_Retention_Report_by_RAND-February_2021.pdf** [Accessed 20 July 2022].

Ford T., 2016. *Supporting Teaching staff to Enable children's Emotional and behavioural Resilience: a feasibility study of the application of the Incredible Years® Teacher Classroom Management Course to the support of children with emotional and behavioural difficulties within mainstream and special education. Final report*. Available on request.

Ford, T., Hayes, R., Byford, S., Edwards, V., Fletcher, M., Logan, S., Norwich, B., Pritchard, W., Allen, K., Allwood, M., Ganguli, P., Grimes, K., Hansford, L., Longdon, B., Norman, S., Price, A., & Ukoumunne, O. C., 2019. 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, 49*(5), pp.828–842. **https://doi.org/10.1017/S0033291718001484**.

Goodman, R., 1997. The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38(5), pp.581-586. **https://doi.org/10.1111/j.1469-7610.1997.tb01545.x**.

Goodman, R., 2001. Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), pp.1337-1345. **https://doi.org/10.1097/00004583-200111000-00015**.

Goodman, A. & Gregg, P., 2010. *Poorer children's educational attainment: How important are attitudes and behaviour?* York, UK: Joseph Rowntree Foundation. [pdf] Available at: **http://www.jrf.org.uk/publications/educational-attainment-poor-children** [Accessed 21 December 2021].

Gutman, L. M., & Feinstein, L., 2008. *Children's Well-Being in Primary School: Pupil and School Effects. Wider Benefits of Learning Research Report No. 25*. Centre for Research on the Wider Benefits of Learning. London: Institute of Education, University of London. [pdf] Available at: **https://discovery.ucl.ac.uk/id/eprint/10002050/** [Accessed 20 July 2022].

Hay McBer, 2000. *Research into teacher effectiveness: A model of teacher effectiveness report by Hay McBer to the Department for Education and Employment—June 2000.* London: DfE. [pdf] Available at: **https://dera.ioe.ac.uk/4566/1/RR216.pdf** [Accessed 21 December 2021].

Haydn, T., 2014, To what extent is behaviour a problem in English schools? Exploring the scale and prevalence of deficits in classroom climate. *Review of Education,* 2(1), pp.31–64. **https://doi.org/10.1002/rev3.3025**.

Hayes, R., Titheradge, D., Allen, K., Allwood, M., Byford, S., Edwards, V., Hansford, L., Longdon, B., Norman, S., Norwich, B., Russell, A.E., Price, A., Ukoumunne, O.C., Ford, T., 2020. The Incredible Years® Teacher Classroom Management programme and its impact on teachers' professional self-efficacy, work-related stress, and general well-being: Results from the STARS randomized controlled trial. *British Journal of Educational Psychology*, 90(2), pp.330-348. **https://doi.org/10.1111/bjep.12284**.

Hickey, G, McGilloway, S, Hyland, L, Leckey, Y., Kelly, P., Bywater, T., Comiskey, C., Lodge, A., Donnelly, M., & O'Neill, D., 2017. 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(2), pp.174-194.

Lord, P., Styles, B., Morrison, J., White, R., Andrade, J., Bamford, S., Lushey, C., Lucas, M., Smith, R., 2018. *Families and Schools Together (FAST).* London: EEF. [pdf] Available at: **https://d2tic4wvo1iusb.cloudfront.net/documents/projects/FAST.pdf** [Accessed 21 December 2021].

Marlow, R., Hansford, L., Edwards, V., Ukoumunne, O. C., Norman, S., Ingarfield, S., Sharkey, S., Logan, S., Ford, T., 2015. Teaching classroom management – a potential public health intervention?, *Health Education*, 115(3/4), pp. 230-248. **http://dx.doi.org/10.1108/HE-03-2014-0030**.

Ofsted, 2019. *Teacher well-being at work in schools and further education providers*. Manchester: Ofsted. [pdf] Available at: **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/936253/Teacher_well-being_report_110719F.pdf** [Accessed 20 July 2022].

Pianta, R. C., 2001. *STRS Student–Teacher Relationship Scale. Professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A., 2015. Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, 56(3), pp. 345-365. **https://doi.org/10.1111/jcpp.12381**.

Reinke, W. M., Herman, K. C., & Dong, N., 2018. The Incredible Years Teacher Classroom Management Program: Outcomes from a group randomized trial. *Prevention Science*, 19(8), pp.1043-1054. **https://doi.org/10.1007/s11121-018-0932-3**.

Sharp, C., Nelson, J., Lucas, M., Julius, J., McCrone, T. & Sims, D., 2020. The challenges facing schools and pupils in September 2020. [pdf] Available at: https://www.nfer.ac.uk/schools-responses-to-covid-19-the-challenges-facing-schools-and-pupils-in-september-2020/ [Accessed 21 December 2021].
Skripkauskaite, S., Shum, A., Pearcey, S., McCall, A., Waite, P., & Creswell, C., 2021. *Changes in children's mental health symptoms from March 2020 to June 2021 (Report 11)*. Co-SPACE study. [pdf] Available at: **http://cospaceoxford.org/findings/** [Accessed 21 December 2021].

Steer, A., 2009. *Learning Behaviour: LESSONS LEARNED A review of behaviour standards and practices in our schools.* [pdf] Available at: **http://www.educationengland.org.uk/documents/pdfs/2009-steer-report-lessons-learned.pdf** [Accessed 9 December 2021].

Timmins, N., 2021. *Schools and coronavirus: The government's handling of education during the pandemic*. London: Institute for Government. [pdf] Available at: **https://www.instituteforgovernment.org.uk/sites/default/files/publications/schools-and-coronavirus.pdf** [Accessed 9 December 2021].

Wang, T. & Merkle, E. C., 2018. merDeriv: Derivative Computations for Linear Mixed Effects Models with Application to Robust Standard Errors. *Journal of Statistical Software*, Code Snippets, 87(1), pp. 1–16. **https://doi.org/10.18637/jss.v087.c01**.

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), pp.383-420.

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (Eds.), 2004. *Building academic success on social and emotional learning: What does the research say?* New York: Teachers College Press.

Appendix A: EEF cost rating

Figure 2: Cost Rating

Cost rating	Description
£ £ £ £ £	*Very low:* less than £80 per pupil per year.
£ £ £ £ £	*Low:* up to about £200 per pupil per year.
£ £ £ £ £	*Moderate:* up to about £700 per pupil per year.
£ £ £ £ £	*High:* up to £1,200 per pupil per year.
£ £ £ £ £	*Very high:* over £1,200 per pupil per year.

Appendix B: Trial documents

Information sheet - RCT Evaluation of "Incredible Years: Teachers Classroom Management"

The Incredible Years® Teacher Classroom Management (TCM) course is part of a suite of programmes for children, parents and teachers. The parent and child programmes have been researched extensively and have been shown to be extremely effective in reducing child problem behaviours and improving the child's emotional wellbeing. The TCM course comprises a one-day session for teachers every month for six months. It is thought to help teachers support the behaviour of children and thus reduce disruption in the class, leading to a general improvement in academic attainment and social and emotional development. The focus is on collaborative learning, discussions of staff's own experiences and group work to find solutions to problems encountered in the classroom. The following principles underpin the TCM approach:

- The importance of attention, encouragement and praise
- Motivating children through incentives
- Preventing behavioural problems – promoting proactive teaching
- Decreasing children's inappropriate behaviours
- Building positive relationships with children and encouraging them to problem solve.

Research suggests that although staff already use a lot of these techniques, the TCM course can increase their skills and confidence in using them and produce significant improvements in the behaviour of their pupils.

Who is conducting the trial?

The Education Trials Unit at the National Foundation for Educational Research (NFER) is conducting the trial. The University of Exeter is providing the TCM training. The Education Endowment Foundation (EEF) is funding the trial.

What are the trial aims?

We want to know if training teachers with TCM improves children's mathematics attainment, social and emotional development and their behaviour in school. In order to do this and understand the reasons for any change for children whose teachers have received TCM training, we are conducting a randomised controlled trial (RCT). As part of the trial we will also investigate if there are any changes in teachers' classroom practice and sense of wellbeing as well as their views of the training and programme overall.

Which schools, staff and pupils will be involved?

The trial will involve around 140 primary schools across Bristol, Cornwall, Dorset, Liverpool and Southampton. Schools that are under Ofsted 'special measures', that are privately funded or that primarily cater for children with special educational needs will not be eligible.

The children in the study are all children currently in Reception. Schools have a 50% chance of being randomised to receive TCM and in the schools that do receive TCM, we want the study children to be taught by a TCM trained teacher for the next two years. Therefore, schools will need to identify which teachers are most likely to teach the study children when they are in Year 1 (19/20) and Year 2 (20/21). To be eligible to receive the training, teachers must have at least four days of teaching per week (unless both parts of a job share attend).

Each school will need to provide a named school contact for the duration of the trial. This person will act as the single point of contact between the school and the University of Exeter and NFER and could be a class teacher, teaching assistant, SENCO or a member of the leadership team. It does not need to be one of the trained teachers. The named contact will ensure the return of requested data and questionnaires, and ensure teachers are released and covered for the training days.

What will the trial involve for schools?

Date	Activity
2018/19 Year All schools	Schools sign up to take part in the trial, returning the signed Memorandum of Understanding to Exeter in the pre-paid envelope. All schools commit to honouring the condition they are placed in: intervention or control. Schools will provide pupil lists/IDs for all current Reception children and the names of the teachers that will teach these children when they are in Year 1 (2019/20) and Year 2 (2020/21). Schools will be provided with a parent letter, enabling parents to withdraw their child from data sharing. All Reception teachers will complete questionnaires about each child's social and emotional wellbeing and classroom behaviour. These take less than 5 minutes per child.
Jul 2019 All schools	**Schools Randomised** Schools are randomly allocated to the intervention or control group. Schools will receive an email before the end of term confirming which group they have been allocated to and the next steps.
Sep 2019 All schools	All Year 1 and Year 2 teachers will complete questionnaires about their own classroom practice and sense of wellbeing. Intervention schools will begin the training.
2019/20 Year Intervention Schools	Year 1 and Year 2 teachers are released to attend six one-day workshops. Not all teachers have to be released on the same day if this makes timetabling easier. All Year 1 and Year 2 teachers will incorporate TCM techniques into their teaching.
2019/20 & 2020/21 Years Control Schools	All Year 1 and 2 teachers will teach in their normal way.
Jun 2020 All schools	All Year 1 teachers will complete questionnaires about each child's social and emotional wellbeing and classroom behaviour, these take less than 5 minutes per child.

2020/21 Year Intervention Schools	All Year 2 teachers will incorporate TCM techniques into their teaching. Year 1 teachers may continue to use TCM techniques if they choose to.
Jun 2021 All schools	All Year 2 teachers will complete questionnaires about each child's social and emotional wellbeing, classroom behaviour and relationship with their teacher. These take less than 10 minutes per child. All Year 1 and Year 2 teachers will complete questionnaires about their own classroom practice and sense of wellbeing.

In addition, we will complete some interviews and observations in around ten schools in order to further understand teachers' experiences of the training and delivering the intervention.

Due to the continued disruption in schools from Covid-19 we have had to make changes to this trial and the data that will be collected as part of the evaluation. Key Stage 1 mathematics results will no longer be collected as part of the trial.

How will NFER and partners use and protect the data collected?

For information on how the parties involved in this research will gather, use and protect data, please refer to the Privacy Notice for the STARS: Incredible Years® Teacher Classroom Management Evaluation, available for schools here and for parents and carers here.

How will schools benefit from taking part?

All schools in the trial will contribute to the evidence-base on what works in supporting children's mathematics attainment and social and emotional wellbeing in primary schools. Schools allocated to the intervention group will have the opportunity to receive free TCM training.

Do schools have to take part?

Schools only have to take part if they wish to do so. However, all data is important to the trial, and the NFER and the University of Exeter really appreciate schools' support in participating in the trial. How will the findings be used? The findings from the project will be freely available on NFER's and EEF's website. They will be used to inform the education sector about improving mathematics attainment and social and emotional wellbeing in primary schools.

Who can I contact for more information?

Who to contact		**Telephone**	**E-mail**
Kinnery Koria, NFER	For queries about providing data	01753 637205	incredibleyears@nfer.ac.uk
Rachel Hayes, University of Exeter Project Team	For queries about signing up, training and intervention	01392 722978	r.a.hayes@exeter.ac.uk

Evaluation of "Incredible Years: Teacher Classroom Management"

Privacy notice for parents and carers

Why are we collecting this data?

The Incredible Years Teacher Classroom Management (IY-TCM) is a teacher training programme that aims to strengthen teachers' confidence in classroom management and to improve the classroom climate. The programme ultimately intends to enhance children's emotional and social wellbeing.

The National Foundation for Educational Research (NFER) has been commissioned to evaluate the IY-TCM programme. The evaluation will be conducted as a randomised controlled trial; this means that of all schools that apply to take part in the trial, half will be randomly selected to receive the programme, and half will continue with their normal practice. NFER will be collecting data about pupil's wellbeing and behaviour, from all schools that take part in the trial.

The Education Endowment Foundation (EEF) has commissioned and funded the IY-TCM evaluation. NFER and the University of Exeter (joint data controllers) will jointly decide what happens to childrens' data as part of the evaluation. When the trial is completed and the data is submitted to the archive, NFER and the University of Exeter will delete all children's data, and the Education Endowment Foundation (EEF) will become the data controller.

Due to the continued disruption in schools from Covid-19 we have had to make changes to this trial and the data that will be collected as part of the evaluation. Key Stage 1 results will no longer be collected as part of the trial.

What personal data is being collected by this project?

We will collect the following personal data about each child from their school:

- The child's name, date of birth, Unique Pupil Number (UPN), gender and eligibility for Free School Meals. This data will be kept up to date during the trial.
- Teacher's questionnaire responses about each child's behaviour. These include the Pupil Behaviour Questionnaire, and the Student Teacher Relationship Scale.

Originally as part of this evaluation we had hoped to collect the pupils' Year 2 (Key Stage 1) SATs results, however due to the continued disruption in schools from Covid-19, there will be no SATs for pupils in KS1 this year and therefore we will no longer be asking schools to submit this data as part of evaluation.

We will collect the following special data about your child, directly from the school

- Teachers' responses about your child's mental health and wellbeing, gathered through the Strengths and Difficulties Questionnaire.

What is the legal basis for processing activities?

The legal basis for processing your child's personal and special data is covered by:

- 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.

Our legitimate interest for processing your child's personal data is to administer the randomised controlled trial.

- Special personal 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 subject.

Special data can be lawfully processed under 'research exemption' 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.

How will personal data be obtained?

NFER will collect data about each child directly from schools, which will be uploaded to a unique secure school portal.

We will also collect teachers' questionnaire responses about your child's wellbeing and classroom behaviour.

Who will personal data be shared with?

Only NFER will see your children's data during the evaluation.

At the end of the trial, NFER will share the data with EEF's data archive processor through secure data portals, where it will be encrypted and saved to secure servers. At that point EEF will become the data controller and NFER will no longer have any responsibility for it.

Is personal data being transferred outside of the European Economic Areas (EEA)?

No personal data is stored or transferred outside of the EEA.

How long will personal data be retained?

NFER will share all the data with EEF's data archive processor within three months of the trial completion. NFER and the University of Exeter will delete any personal and special data after one year from the publication of the final report.

Data is not kept longer than is necessary and is deleted in accordance with NFER's internal policy.

Can I stop my personal data being used?

NFER handles personal data in accordance with the rights given to individuals under data protection legislation. If at any time you wish us to withdraw your child's data or correct errors in it, please contact **incredibleyears@nfer.ac.uk**

In certain circumstances, data subjects have the right to restrict or object to processing. They also have the right to make a subject access request to see all the information held about them. NFER will cooperate fully when a subject access request (SAR) is made. To exercise these rights, please contact our **Compliance Officers (compliance@nfer.ac.uk,** dataprotection@exeter.ac.uk

Who can I contact about this project?

For any queries, please contact **incredibleyears@nfer.ac.uk**

If you have a concern about the way this project processes personal data, we request that you raise your concern with us or the University of Exeter in the first instance. Alternatively, you can contact the Information Commissioner's Office, the body responsible for enforcing data protection legislation in the UK, at **https://ico.org.uk/concerns/**.

Last updated

We may need to update this privacy notice periodically, so we recommend that you revisit this information from time to time.

This privacy notice was updated on 6 May 2021. The update covered changes to the personal data being collected as a result of changes to Key Stage 1 standardised testing this year.

Evaluation of "Incredible Years: Teacher Classroom Management"

Privacy notice for teachers and schools

Why are we collecting this data?

The National Foundation for Educational Research (NFER) is collecting personal data to enable the evaluation of the 'Incredible Years: Teacher Classroom Management' (IY-TCM) programme, using a randomised controlled trial. The main aim of the IY-TCM teacher training programme is to strengthen teachers' confidence in classroom management and to improve the classroom climate which ultimately intends to enhance children's mental health and promote their prosocial behaviour. The trial aims to ascertain the impact of the intervention on pupil wellbeing and behaviour.

NFER and the University of Exeter will be the joint data controllers for the trial. When the trial is completed and the data is submitted to the archive, the Education Endowment Foundation (EEF) will become the data controller.

Due to the continued disruption in schools from Covid-19 we have had to make changes to this trial and the data that will be collected as part of the evaluation. Key Stage 1 mathematics results will no longer be collected as part of the trial.

What is the legal basis for processing activities?

Due to the nature of the intervention NFER requires to collect and process teachers' 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.

Our legitimate interest for processing teachers' personal data is to administer the randomised controlled trial.

How will personal data be obtained?

For recruitment purposes, The University of Exeter will collect school names, contact details and nominated teacher details, which will be transferred to NFER via a secure portal to carry out the randomisation and proceed with the trial and evaluation.

During the trial, the University of Exeter will take registers of teacher attendance at training, which will be provided to NFER also via the secure portal. NFER will perform training session observations, classroom observations and phone interviews with selected teachers. At two points of the trial, teachers will also complete program-implementation surveys.

What personal data is being collected by this project?

NFER will collect personal data about nominated teachers directly from the recruited schools. This includes teacher name, job title and contact details for Reception, Year 1 and Year 2 classroom teachers. In addition, all Year 1 teachers will complete program-implementation surveys at the beginning of the trial and Year 2 teachers will complete the survey at the end of the trial. The University of Exeter will also take attendance registers at teacher training events and provide this information to NFER.

NFER will also conduct a number of case studies which comprise classroom observations (using the CLASS tool) and phone interviews with teachers. These will be matched with personal data by NFER (though teachers remain anonymous within all reporting).

Who will personal data be shared with?

Following recruitment, the University of Exeter will share teachers' personal details with NFER via secure portal.

Is personal data being transferred outside of the European Economic Areas (EEA)?

No personal data is stored or transferred outside of the EEA.

How long will personal data be retained?

NFER and the University of Exeter will delete any personal and special data after one year from the publication of the final report.

Data is not kept longer than is necessary and is deleted in accordance with NFER's internal policy.

Can I stop my personal data being used?

NFER handles personal data in accordance with the rights given to individuals under data protection legislation. If at any time you wish us to withdraw your data or correct errors in it, please contact **incredibleyears@nfer.ac.uk**

In certain circumstances, data subjects have the right to restrict or object to processing. They also have the right to make a subject access request to see all the information held about them. NFER will cooperate fully when a subject access request (SAR) is made of a data controller. To exercise these rights, please contact our **Compliance Officer**.

Who can I contact about this project?

If you have any queries, please contact **incredibleyears@nfer.ac.uk**

If you have a concern about the way this project processes personal data, we request that you raise your concern with us or the University of Exeter in the first instance. Alternatively, you can contact the Information Commissioner's Office, the body responsible for enforcing data protection legislation in the UK, at **https://ico.org.uk/concerns/**.

Last updated

We may need to update this privacy notice periodically, so we recommend that you revisit this information from time to time.

This privacy notice was last updated on 6th May 2021.

Further appendices

Appendices C-N are available as a separate document (Technical Notes).

This document is available for download at https://educationendowmentfoundation.org.uk

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk
@EducEndowFoundn
Facebook.com/EducEndowFoundn