Trial Evaluation Protocol
Thinking, Doing, Talking Science

Evaluator (institution): University of York
Principal investigator(s): Pam Hanley/Louise Elliott

Education Endowment Foundation

PROJECT TITLE	Thinking, Doing, Talking Science (second re-grant - a two-armed, cluster randomised trial)
DEVELOPER (INSTITUTION)	Science Oxford (The Oxford Trust)
EVALUATOR (INSTITUTION)	York Trials Unit, University of York
PRINCIPAL INVESTIGATORS	Dr Pam Hanley and Louise Elliott
PROTOCOL AUTHORS	Pam Hanley, Louise Elliott, Imogen Fountain, Jenny Roche, Laura Mandefield and Caroline Fairhurst
TRIAL DESIGN	Two-arm, cluster randomised controlled trial with random allocation at school level: Cohort 1: Year 5 (main trial) – 2022-23 Year 6 (longitudinal follow-up) 2023-24 Cohort 2 – Year 5 – 2023-24
TRIAL TYPE	Effectiveness
PUPIL AGE RANGE AND KEY STAGE	Key Stage 2 9-10 years; Year 5 (main trial) 10-11 years; Year 6 (longitudinal follow-up)
NUMBER OF SCHOOLS	180 primary schools[1]
NUMBER OF PUPILS	8100 per cohort, i.e. 16200 over the 2 years
PRIMARY OUTCOME MEASURE AND SOURCE	Year 5 Science Assessment, Centre for Industry Education Collaboration (CIEC) and York Trials Unit, University of York
SECONDARY OUTCOME MEASURE AND SOURCE	Science Attitudes Questionnaire[2], Year 6 Science Assessment, Key Stage 2 SATs attainment in Maths and Reading

[1] or middle schools if they include both Year 5 and Year 6
[2] based on Kind, Jones & Barmby, 2007

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.0 [*original*]	29/7/21	N/A

Table of contents

Study rationale and background 4

Intervention 5

Evaluation of Train-the-trainer model 10

Stage 1: Developers deliver TDTS to trainers 10

Stage 2: Developers deliver train-the-trainers sessions to trainers 11

Stage 3: Trainers train pre-trial teachers 11

Stage 4: Trainers train trial teachers 13

Impact evaluation 14

Research questions 14

Design 14

Randomisation 15

Participants 15

Outcome measures 18

Secondary outcomes 19

Compliance 19

Analysis 20

Longitudinal follow-ups 21

Implementation and process evaluation 21

Research questions 21

Research methods and analysis 22

Cost evaluation 26

Ethics and registration 26

Data protection 26

Personnel 28

Development team 28

Evaluation team 29

Risks 30

Timeline 32

References 34

Appendix 1: Changes since the previous EEF evaluation 36

Appendix 2: Brief overview of TDTS 4.5 Day training sessions 38

Study rationale and background

The primary science experience heavily influences subsequent subject attitudes but is often low priority and teachers may lack confidence teaching it (Harlen & Qualter, 2008; Slavin et al, 2014). Thinking, Doing, Talking Science (TDTS) is a continuing professional development (CPD) programme for teachers that aims to enable the teachers to adapt their pedagogy to plan and teach creative science lessons that overtly encourage their pupils' higher order thinking.

In a small-scale efficacy trial involving Year 5 pupils in 41 schools (Hanley et al, 2015), pupils of teachers trained in TDTS made three months additional progress in science, with a particularly positive effect among girls and pupils with low prior attainment. There were indications that the approach might be especially beneficial for pupils eligible for free school meals, but this required further exploration. There was an apparent positive impact on attitudes towards science. However, a subsequent effectiveness trial in 205 schools (Kitmitto et al, 2018) failed to show evidence of additional progress for most pupils - although pupils eligible for free school meals (FSM) made a small amount of additional progress and pupils' interest and self-efficacy in science showed a small improvement.

The main changes between the two trials related to teacher training. In contrast to the efficacy trial, the effectiveness trial used a "train-the-trainers" model, rather than the developers training the teachers directly, the CPD days were reduced from five to four, and the funding to cover two in-school preparation days/teacher was eliminated.

There are several other instances of success in smaller Education Endowment Foundation (EEF) trials not being replicated at scale. One commonly-shared change on scale-up is the adoption of the "train-the-trainer" model: practice shifts from training of the teachers being delivered directly by the developers in the first trial to delivery by relatively inexperienced TDTS trainers, trained by the developers, in the second trial.

Train-the-trainer is the theory that a group of individuals can be given training in a new concept and then go on to train a large group in this newly acquired skill (Ray et al, 2012). The model is increasingly implemented in business, healthcare settings and education (Gask et al, 2019) as it can be used to reach a large audience in a relatively cost-effective way (Wedell, 2005). Although there is not a wealth of literature regarding this model, evidence suggests that this multi-level process can generate a number of problems.

The main disadvantage is the dilution of the knowledge as it is passed down (Hayes, 2000). Reasons for this dilution include: knowledge transfer and the ability to train others (Turner et al, 2017); focus of knowledge at the uppermost levels and "transmissive training" (Hayes, 2000); lack of social and cultural awareness (Bax, 2002); and lack of confidence of the trained to teach their new knowledge (Dichaba et al, 2012). Other problems include rate of staff turnover (Gask et al, 2019) and the lack of "proactive technical assistance" after the initial training of the trainer (Ray et al, 2012).

After the first effectiveness trial, the Oxford team recommended strengthening the train-the-trainers model, including its length and rigour, to improve the impact of TDTS in the event of a retrial (internal report, 2019). They have introduced various amendments. Previously, training was delivered to trainers throughout the intervention year. It is now intended that, before trainers start delivering any training to teachers, they will receive the full TDTS course (as if they were teachers) as well as training in delivering to teachers. They will also train and deliver to teachers in pre-trial schools before the trial begins. The developers intend to improve quality assurance,

for instance by observing trainers delivering to the pre-trial schools, and to improve trainer resources. See Appendix 1 for details.

This study has been designed not only to re-evaluate TDTS but also to inform scale-ups more generally, with particular reference to those using a train-the-trainer model. As such, the study has a substantial initial component that evaluates the train-the-trainers model for efficacy and fidelity of delivery. The trial will then focus on the training of the main trial teachers and the experience, attitudes and performance of their Year 5 pupils. The primary outcome measure will be a general science test that covers the Year 5 curriculum, and the secondary outcome will be a pupil science attitude questionnaire. The secondary outcome measure is a slightly modified version of the one used in the previous efficacy and effectiveness trials; however, the primary outcome measure previously used is now outdated. Therefore, a new outcome measure that maps to the current science curriculum is being developed for use in this trial (see Primary outcome section). A thorough implementation and process evaluation (IPE) will take place across the year, including lesson observations, interviews and surveys of trainers, teachers and pupils. The next cohort of Year 5 pupils will also be followed and given the same science test and attitude questionnaire as the first cohort as well as the online survey elements of the IPE. This is to investigate whether the effect of TDTS appears to be modified in any way after teachers have received the entire training package (which is delivered across the academic year) and had a greater opportunity to incorporate TDTS in their science teaching. The first cohort of Year 5 pupils will be followed into the second year and will complete a science test at the end of Year 6. This test is being developed by the York Trials Unit, University of York and will reflect the current curriculum, have a mix of question types and have an emphasis on "working scientifically". We shall also follow them up based on their attainment in Mathematics and Reading in the Year 6 SATs, using data from the National Pupil Database (NPD).

The TDTS programme is led by Science Oxford and will be independently evaluated by York Trials Unit, University of York. The study is funded by the Education Endowment Foundation (EEF) and the Wellcome Trust.

Intervention

The main goal of TDTS is to develop teachers' delivery of science lessons so that they actively encourage their pupils' higher order thinking. They will enable them to think and talk about scientific concepts in every science lesson, through dedicated discussion slots (the Bright Ideas Time) linked to the topic being taught. Teachers will facilitate their pupils' thinking through practical science, providing them frequent opportunities for creative investigations and problem solving. Pupils will not record everything they do in a practical as the teacher will focus the recording on the lesson's learning objectives, so that time for thinking, doing and talking is prioritised.

It is anticipated that this will enhance Year 5 pupils' higher order thinking skills and subsequently their attainment outcomes in science. As in the previous trials, the logic model predicts that, by encouraging these higher order thinking skills, pupils will engage more deeply and actively, developing an increased interest and self-efficacy in science.

This trial runs across two years. The first year forms the main trial. Year 5 teachers will attend CPD sessions in the academic year 2022-23, four of which will be spread throughout the first two terms, with a further half-day in the third (Summer) term. Teachers will be given 'gap' tasks/strategies to use with their classes between the sessions and encouraged to reflect on their implementation, discuss with their in-school colleagues and then feedback at the next CPD session. Any Year 5 teachers who join the school during the year should inherit the previous

teacher's file and receive input from the other participating teacher(s) in their school as well as attending any subsequent training sessions, to reflect the real world approach.

The second year of the trial (with a second cohort of Year 5 pupils) will examine the legacy of the TDTS training and any effects of embedding of the TDTS practices. At the recruitment stage and in the MoU, schools will be encouraged, wherever possible (e.g. unless the teacher is no longer at the school or operational circumstances make it impossible) to ensure that the same teachers will be retained in Year 5 for both years of the evaluation. No training will be provided by the TDTS team to teachers new to Year 5 in the second trial year, but the final half-day of TDTS training will include a section on cascading the approach to colleagues across the school. The intervention schools in this second year of the evaluation will therefore have a mix of teachers that taught a Year 5 class at an intervention school in the first year and/or received training from the TDTS team, and teachers new to TDTS who did not teach a Year 5 class in an intervention school in the first year and have received no external training in TDTS (but may have received cascade training from an experienced teacher at their school).

The second year will also follow the first cohort of Year 5 pupils into Year 6 to assess the 'legacy' effects of exposure to the TDTS programme. Year 6 teachers may have received TDTS training if they have moved from a Year 5 class the previous year or it has been cascaded within the school but no training will be provided by TDTS to Year 6 teachers.

For the first year (2022-23), schools allocated to TDTS will not be offered a financial incentive because we anticipate attrition will be low, as was the case in the previous effectiveness trial; however, they will be given a resources grant, based on the number of teachers taking part, which can be used for equipment etc, and some low-value science equipment to take away from the training days. Intervention schools will be offered £500 for completing the requirements of the evaluation in the second year, when they receive no further input from the TDTS team. The control group schools will be offered a total financial incentive of £1500 for participating, payable in two amounts: £1000 after completion of the requirements in the first year; and £500 after completing the second year requirements at the end of the second year.

The TIDieR table below outlines the details for the procedure for the first year of the trial. Where relevant, differences in the second year are summarised in square brackets.

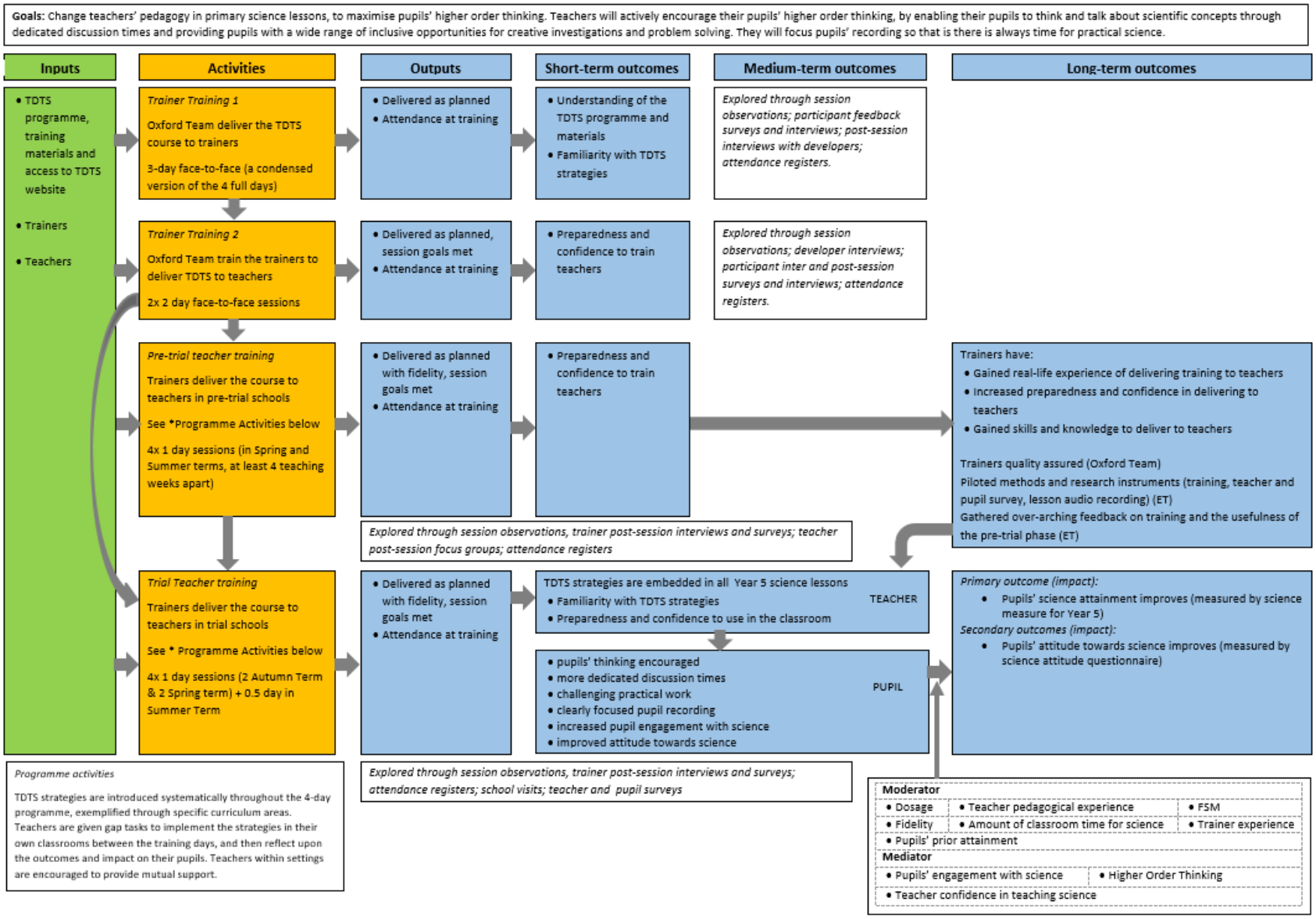
Table 1: Description of the programme using the Template for Intervention Description and Replication (TIDieR) checklist

Aspect of TIDieR	Exemplification relating to the evaluation
Brief name	Thinking, Doing, Talking Science (TDTS)
Why: Rationale, theory and/or goal of essential elements of the intervention	TDTS aims to improve Year 5 pupils' higher order thinking skills and science outcomes by improving teachers' delivery of science lessons. Government biennial sampling tests estimate that only 21.2% of pupils achieved the expected standard in science in 2018.There has been a previous efficacy (Hanley et al, 2015) and effectiveness (Kitmitto et al, 2018) trial of the intervention. This second effectiveness trial incorporates an evaluation of the train-

	the-trainer model (an amended version of the one used in the first effectiveness trial) as well as evaluating the intervention itself.
Who: Recipients of the intervention	Teachers in all Year 5 classes will be invited to attend training. Where there is only one Year 5 class another teacher, ideally the science subject lead, will also receive training. [There will be no external delivery of the intervention to new Year 5 teachers in the second year of the trial. They will be reliant on the teachers trained in the first year and accompanying physical/online materials for any learning about TDTS]
What: Physical or informational materials used in the intervention	Each teacher will receive hard copies of all TDTS course resources in a ring binder and some low-value science equipment at the point of course delivery. They will also have ongoing access to online versions of TDTS course resources via a dedicated website (https://tdts.org.uk/). [first year only, although online access will still be available in the second year to Year 5 teachers in the intervention arm]
What: Procedures, activities and/or processes used in the intervention	All Year 5 teachers in intervention schools will receive 4 one-day continuing professional development (CPD) sessions; these will be held towards the beginning and end of the first two terms of the academic year. There will be a further half-day during the third-term to share good practice and provide advice on disseminating TDTS within their schools (see Appendix 2 for further detail). [first year only] Between training sessions, teachers will be asked to try some strategies with their classes and then feedback and discuss at the next session. At least two teachers from each school will participate in the intervention and TDTS teachers will be encouraged to provide informal peer support for each other within schools.
Who: Intervention providers/implementers	Qualified TDTS trainers, certified to deliver the training course will deliver the TDTS course to Year 5 classroom teachers [first year only]
How: Mode of delivery	Teachers attend group CPD sessions delivered face-to-face. Each session will be run once per region by a pair of trainers, with expected attendance of 20-40 teachers. [first year only]

Where: Location of the intervention	CPD sessions will be run in each of the six regions [first year only]. The regions will be spread across England as far as possible, but the final choice will be a pragmatic one based on the location and reach of the final team of trainers.
When and how much: Duration and dosage of the intervention	The CPD will consist of 4 one-day sessions spread over the first two terms of the academic year and a further half-day in the Summer term. [first year only]
Tailoring: Adaptation of the intervention	No adaptations anticipated.
How well (planned): Strategies to maximise effective implementation	Six of the 24 full-day CPD sessions will be observed by the evaluation team, one in each region, and teachers will be asked for feedback on the training in the teacher surveys and interviews as part of the process evaluation. The evaluation team will also use the teacher feedback designed by the development team (in consultation with the evaluators) and completed after each training session. Short interviews will be conducted with trainers after each observed training session and they will also be asked to complete a brief survey to gather their feedback on the effectiveness of the session.

Figure 1: TDTS Logic Model



Evaluation of Train-the-trainer model

Background

This section of the protocol focuses on the evaluation of the "train-the-trainers" element of TDTS. This divides into three consecutive stages which will take place from July 2021 through the academic year 2021-22. During each stage, trainers fulfil a different role:

Stage 1: trainers, as if they are participant teachers, receive training in TDTS from the developers;

Stage 2: trainers are trained by the developers to deliver TDTS training to teachers;

Stage 3: trainers will train pre-trial teachers (i.e. who will not be involved in the main trial). Trainers will work in six pairs comprising one 'experienced' trainer (either one of the developers or someone who has trained teachers to use TDTS in the past) and one 'new' trainer.

All trainers, whether categorised as experienced or new, will participate in Stages 1, 2 and 3.

Four main research questions will be addressed in this part of the evaluation:

- To what extent is the training model implemented as planned?
- How effective is each element within the model at achieving its aim(s)?
- How necessary is each element to the training model overall?
- What improvements could be made to the model to benefit the TDTS intervention and training design more widely?

This stage of the evaluation will also provide an important opportunity to trial new/amended research instruments before the impact evaluation.

On completion of the three stages of the train-the-trainers model, a final team of trainers will be selected to train the teachers participating in the intervention arm of the trial. Details of this are covered in the "Implementation and process evaluation" section.

Stage 1: Developers deliver TDTS to trainers

The Evaluation Team (ET) will observe the course in which trainers will receive training in TDTS. The intention is that trainers will experience the course as if they were participant school teachers, although the content will be condensed from four separate days into a three-day residential. The ET will complete an observation schedule and fieldnotes for each session. They will also ascertain developers' satisfaction with the event, and (in addition to on-the-day participant evaluations) obtain trainer feedback via a survey. For further depth, towards the end of the course, two of the new trainers will take part in a paired interview. This will help determine whether the course objectives have been met.

Focus	Method	Why?	Number
Developer-run TDTS course	Observation schedule and fieldnotes	To gain more insight into TDTS and how the developers run the training	All 3 days
Developer	Post-session interview	To establish whether sessions ran as planned	At the end of each day

Trainers	Survey [devised by Oxford team in consultation with ET]	To obtain feedback from sessions	At the end of each day
Trainers	Paired interview with "new" trainers	To gain deeper insight into training experience	1 (2 participants)

Stage 2: Developers deliver train-the-trainers sessions to trainers

Stage 2 consists of two 2-day sessions (2 months apart) designed to prepare the trainers to deliver TDTS training to participating teachers. We will determine from developers their training goals for each session. Each day will be observed by the ET and developer feedback collected after each 2-day session. Trainer surveys between the sessions will check learning, recall and issues from the previous session and expectations of the next. After the final session, there will be a survey reflecting on the full course. For more depth, two trainers will be interviewed for their feedback after the full training course, and another two just before delivery to pre-trial teachers to explore their preparations and confidence. Assuming trainers have been assigned to their regional pairings by this stage, the two trainers will comprise a working pair where possible.

Focus	Method	Why?	Number
Developer	Pre-session communication	To determine training goals for each session	1 (2-3 participants)
Developer-run train-the-trainer course	Observation schedule and fieldnotes	To understand the expected model of TDTS training delivery	All (4 days)
Developer	Post-session interview	To establish whether sessions ran to plan and goals were met	2 (after second and fourth days)
Trainers	Inter-event survey	To check learning, recall, issues from the first session and expectations of the final session	1 (all trainers)
Trainers	Post-course survey	To check learning, recall, issues from the final session and reflect on the course overall	1 (all trainers)
Trainers	Paired interview with "new" trainers	To gain deeper insight into training experience (first pair); to explore preparations and confidence (second pair)	2x2=4 (one post-training, one pre-delivery)

Stage 3: Trainers train pre-trial teachers

In the third stage, trainers will provide TDTS training to pre-trial teachers, delivering the four sessions across a 6-month period.

Approximately 60 schools, ten in each of the six geographical areas will be recruited by the trainers, with oversight from Science Oxford during the Autumn Term 2021-2022. Schools will be eligible All state primary schools[3], including academies, in England can take part in the pre-trial as long as the following eligibility criteria are met:

- Nominated teachers have not been TDTS-trained.
- School will not take part in the main trial.

[3] or middle schools if they include both Year 5 and Year 6

- School has not been part of EEF Stop & Think trial.
- School has not been part of EEF Focus4TAPS trial.
- If the school is part of a Multi-Academy Trust (MAT), then no school in the MAT will take part in the main trial.
- School is aware the training is Year 5-focused.

It is recommended that two Year 5 teachers from each school take part, however this is not an eligibility requirement.

Each school will receive an Information Sheet explaining the pre-trial and a MoU, which will outline the schools' commitment/obligation.. In order to take part in the pre-trial schools must complete the MoU and return it to the developers who will forward a copy to the ET.

This pre-trial phase is part of the training model to give the trainers additional experience (thus bringing them to the level they might be expected to be at in real life). It will also allow the Oxford team to observe the trainers and select the final team for the trial. It offers the ET some benefits as well. Firstly, the importance of this stage to the development of the TDTS trainers can be assessed through observation, interview and survey. Secondly, it allows evaluation of new or amended research instruments: teacher surveys, lesson observation forms and pupil measures. Thirdly, an alternative approach to obtain examples of more "typical lessons" for evaluation can be explored.

Exploring an additional approach to in-person lesson observations is recommended because TDTS is not a prescriptive programme or pedagogy but uses various techniques and strategies. Lessons that are observed in-person might be particularly susceptible to the "measurement effect" where teachers prepare lessons with more TDTS characteristics than usual. Although this can be triangulated with pupil feedback, it might also be possible to reduce the problem by using remote recording of lessons. Teachers in two schools will be asked to audio-record their science lessons over several weeks. After discarding the very first as potentially atypical, the evaluators would randomly sample two from each teacher for analysis. The intention is that this would minimise the "measurement effect" as teachers/children become used to the recorder. We will assess whether the recordings work as planned and, if judged successful, they would supplement, rather than replace, in-school observations in the IPE of the main trial. Success will be judged on several criteria including the lessons being recorded as planned, recordings returned to YTU, usable audio quality, and the content providing information that would enhance lessons captured by a physically present observer.

Focus	Method	Why?	Number
Trainer-run TDTS course	Observation	To assess fidelity of delivery	6 sessions (one per region and covering each of the four days)
Trainers	Post-session interview	To establish whether sessions ran to plan and goals were met	6 (at above sessions)
Teachers	Focus group	To explore feedback from the sessions	2 groups (different regions) with 4-6 teachers
Trainers	Survey	To gather over-arching feedback on training and how useful they find the pre-trial phase	All (4 sessions)
Instrument (teacher)	New survey	Develop and pilot teacher survey for main trial	15 schools (30+ teachers)
Instrument (lesson)	New observation schedule	Develop and pilot observation schedule and fidelity measures for main trial	4 lessons (2 schools)
Instrument (pupil)	New survey	Develop and pilot pupil survey (including attitude statements) for main trial	2 classes x 4 schools
Technique (lesson)	Audio recording	Triangulate face-to-face observation with data from audio-recorded lessons (two sampled from several). Test usability of recording in main trial. To overcome possible “measurement effect” when observing in-school lessons that may have been more carefully prepared than usual.	2 classes x 2 schools

Stage 4: Trainers train trial teachers

This is covered in the “Implementation and process evaluation” section.

Impact evaluation

Research questions

Main trial: Cohort 1 – Year 5

RQ 1. What is the impact of the TDTS programme, in comparison to usual Year 5 provision, on the science attainment of Year 5 pupils? [primary outcome]

RQ 2. What is the impact of the TDTS programme, in comparison to usual Year 5 provision, on pupils' attitudes towards science? [secondary outcome]

RQ 3. What is the impact of the TDTS programme, in comparison to usual Year 5 provision, on the science attainment of Year 5 pupils who are eligible for Free School Meals?

RQ 4. What is the long-term impact of the TDTS programme, in comparison to usual Year 5 provision, on pupils' science attainment at the end of Year 6 and on Key Stage 2 outcomes (Year 6 SATs attainment in Reading and Maths) ? [secondary outcomes]

Second year: Cohort 2 – Year 5

RQ 5. What is the impact of the TDTS programme, in comparison to usual Year 5 provision, on the science attainment of Year 5 pupils given the mix of experienced and inexperienced teachers in the intervention group?

RQ 6. What is the impact of the TDTS programme on pupils' attitudes towards science, in comparison to usual Year 5 provision, given the mix of experienced and inexperienced teachers in the intervention group?

RQ 7. What is the impact of the TDTS programme, in comparison to usual Year 5 provision, on the science attainment of Year 5 pupils who are eligible for Free School Meals given the mix of experienced and inexperienced teachers in the intervention group?

Design

Table 2: Trial design

Trial design, including number of arms		Two-arm, cluster randomised
Unit of randomisation		School
Stratification variables (if applicable)		Region and % FSM (school level, taken at the time of recruitment from the latest census data)
Primary outcome	variable	Science attainment
	measure (instrument, scale, source)	Year 5 Science Assessment, 15-item measure scored 0-45, Centre for Industry Education Collaboration (CIEC) and York Trials Unit, University of York (currently under development)

Secondary outcome(s)	variable(s)	Attitudes towards Science
	measure(s) (instrument, scale, source)	Science Attitudes Questionnaire, 25-item measure, 5-point Likert scale, based on Kind, Jones & Barmby, 2007
Secondary outcome(s)	variable(s)	Science attainment
	measure(s) (instrument, scale, source)	Year 6 Science Assessment, York Trials Unit, University of York (currently under development, scale tbc)
Secondary outcome(s)	variable(s)	Attainment in Mathematics and Reading
	measure(s) (instrument, scale, source)	Key Stage 2 (Year 6 SATs attainment in Reading and Maths) from the National Pupil Database English Reading (KS2_READSCORE_noSpeccon, range 0-120) Maths (KS2_MATSCORE_noSpeccon, range 0-120)
Measure of prior attainment	variable	Early Years Foundation Stage Profile (EYFSP).
	measure (instrument, scale, source)	Average EYFSP point score obtained by combining all 17 Early Learning Goals (ELG)

Randomisation

Randomisation will be conducted at the school level using a 1:1 ratio to intervention or control. Minimisation will be used in order to ensure groups are balanced across the following important school characteristics: region (n=6) and percentage of pupils in the school who have ever been eligible for free school meals (ever-FSM). Ever-FSM will be dichotomised in the minimisation process at the median values. An independent trial statistician at YTU will be responsible for conducting the minimisation using minimPY software (Saghaei & Saghaei, 2011). Once schools have completed the necessary baseline tasks, they will be ready to be randomised. When a group of schools is ready to be randomised, they will be entered into the minimisation program in one go, in a random order within each batch. Therefore, even if the minimisation factors for the schools are known it will not be possible to predict the allocation sequence in advance and so allocation concealment is assured. It will therefore not be necessary to introduce a random element to the minimisation, which can be used to minimise predictability when schools are randomised one-by-one on a rolling basis.

Participants

Schools

Recruitment of schools will be led by the developers with support from the evaluation team. Schools will be recruited during the academic year 2021-22. The Oxford team will lead the

recruitment with most trainers recruiting schools through their own contacts. Methods of recruitment will include using existing contacts, conferences, publicity through third parties and social media. It may be necessary for the Oxford team to fund additional partners to assist with the recruitment. Schools will be recruited from six geographical areas across England. The schools will be representative of their area whilst targeting those that are higher than average in the percentage of pupils receiving free school meals (FSM).

All state primary schools, including academies, in England can take part in the trial as long as the following eligibility criteria are met:

- The school must have a minimum of one full class of Year 5 pupils (mixed year group classes will not be eligible to take part).
- The school does not operate a two-year science curriculum that involves Year 5 pupils (i.e. either Year 4/Year 5 or Year 5/Year 6).
- The school will allow all Year 5 teachers to be available for the 4.5 days of training. If a school only has one Year 5 teacher, another teacher (ideally the science co-ordinator) would also need to attend the training.
- The school or individuals involved have not been involved in the previous trials of TDTS, been trained in TDTS or taken part in the pre-trial. If the school is part of a MAT then none of the schools within the MAT have taken part in the pre-trial.
- The school is not involved in the EEF Stop & Think trial.
- The school has not been involved in the EEF Focus for Teacher Assessment of Primary Science (Focus4TAPS)
- The school agrees to all requirements outlined in the Information for Schools and Memorandum of Understanding (MoU) documents (including commitment to keep same Year 5 teachers across the two years wherever possible).

Schools within a multi-academy trust (MAT) will be eligible to participate on the understanding that schools within the same MAT must agree that they either do not usually, or will not during the period of the trial, collaborate on science teaching. This is essential to minimise the risk of contamination between schools in the intervention and control groups. Also the MATs must accept that their schools will be randomised individually and so may be allocated to different groups. Alternatively, a MAT can nominate just one school to take part.

Each school will receive an Information Sheet explaining the trial and a MoU, which will outline the schools' commitment/obligation to the trial. In order to take part in the trial schools must complete the MoU and return it to the developers who will forward a copy to the ET.

Pupils

As the TDTS programme is designed to be delivered at a whole-class level, all the Year 5 pupils within the school will be able to participate in the trial. At the beginning of the academic year, parent/carers will be informed about the research through an information sheet sent on behalf of the evaluation team by schools to parents/carers. Parents/carers will be asked to return a signed 'withdrawal from research' form if they are unwilling to share their child's data with the ET and/or they do not wish their child to take part in any assessments, surveys or focus groups. This will apply for both cohorts of Year 5 pupils. This will be repeated for the second cohort of Year 5 pupils.

Incentives

Schools randomly allocated to the intervention arm of the trial will receive the TDTS course for free and will also receive a resources grant as acknowledgement of the evaluation work that is required of them as part of the trial, along with some low-value science equipment. After completing the second year, they will be eligible for a financial incentive of £500.

The schools allocated to the control arm will be eligible for a financial incentive of £1000 in the first year and £500 in the second year.

Sample size calculations

Table 3: Sample size calculations

		OVERALL	FSM
Minimum Detectable Effect Size (MDES)		0.15[a]	0.19[a]
Pre-test/ post-test correlations	level 1 (pupil)	0.5	0.5
	level 2 (class)	-	-
	level 3 (school)	-	-
Intracluster correlations (ICCs)	level 2 (class)	-	-
	level 3 (school)	0.15	0.15
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two	Two
Average cluster size (at randomisation)		45	~8
Number of schools	Intervention	90	90
	Control	90	90
	Total	180	180
Number of pupils	Intervention	4050	700
	Control	4050	700
	Total	8100	1400

[a]Accounting for 15% attrition

A summary of the assumptions used in the calculation of the sample size are given in Table 2. The primary outcome will compare science attainment scores between the intervention and control groups for cohort 1. All pupils in the intervention and control groups will be tested. The following calculation is for a single year group. Based on the previous TDTS trials we have assumed an ICC of 0.15, and an average year group (cluster) size of 45 at randomisation. In the efficacy TDTS trial, the observed correlation between the pre-test (Science Knowledge Questionnaire administered in Year 4) and outcome (Science Knowledge Questionnaire administered at the end of Year 5) was 0.51. In the first effectiveness trial, the analysis model for the outcome (Science Knowledge Questionnaire) included achievement at KS1 in

reading/writing and mathematics as a covariate (as a measure of prior attainment). The proportion of variance explained by level 1 covariates (R^2) was 0.4, suggesting a pre-post test correlation of around 0.6. In this trial, we shall use the average score from the 17 ELGs of the EYFSP (obtained via the National Pupil Database [NPD]) as the measure of prior attainment. This was similarly used in the EEF Stop and Think trial (Roy et al, 2019), for which the post-test was GL Assessment's Progress Test in Science 10 measured at the end of Year 5, and the correlation between pre-test and post-test was 0.53. Based on these estimates, but acknowledging the differences in outcome measures used as pre- and post-tests, we conservatively assume a pre- and post-test correlation of 0.5 for this calculation. Hence, to detect an effect size of 0.15 with 80% power and two-sided alpha of 0.05, assuming pupil-level attrition of 15%, a total of 180 schools would be required (8100 pupils per year group).

As of January 2020, 17.3% of pupils were eligible for free schools meals. Assuming we recruit 180 schools and an anticipated total of 8100 pupils per each year of the trial, there will be approximately 1400 pupils eligible for FSM each year (approximately 8 per school). Under the same assumptions as above, an MDES of 0.19 will be detectable. All calculations were conducted in Stata (Version 15).

Outcome measures

Baseline measures

In order to minimise the burden on schools, as well as cost, the baseline measure used in both the primary and secondary analyses will be the average point score from the 17 Early Learning Goals (ELGs) that make up the Early Years Foundation Stage Profile (EYFSP). This average GLD baseline measure has been chosen as an alternative to the KS1 English (Reading) and Mathematics scores used in the previous effectiveness trial as KS1 results are not available for the first cohort of pupils (2022-23). This cohort was in Year 2 during the academic year 2019-20 when schools were closed from March 20th due to Covid-19 and national assessments were cancelled. The EEF Stop and Think trial (Roy et al, 2019) used a GLD average as the pre-test for their co-primary outcomes of Maths and Science and observed a correlation of 0.53 for the science outcome (GL Assessment's Progress Test in Science 10) in the Year 5 cohort.

Primary outcome

The measure used for both the efficacy and effectiveness trials of TDTS is no longer fit for purpose. Its creation (Abrahams et al, 2014) preceded the new science curriculum (DfE, 2013) with its changed content and emphases (e.g. more focus on "working scientifically"/science enquiry). The main alternative (GL Progress Test in Science) is not considered to be a varied enough test (for instance, it is predominantly multiple choice and its coverage of "working scientifically" has been criticised) to be an adequate replacement. Instead we are intending to use a new measure, the Year 5 Science Assessment, developed by the Centre for Industry Education Collaboration (CIEC) and York Trials Unit (YTU), University of York. It was originally developed for use in two EEF-funded RCTs in 2020. However, both these trials were delayed because of school closures due to Covid-19. This new measure has been developed to better reflect the current curriculum, have a mix of question types and have greater emphasis on "working scientifically" than the alternatives.

The same assessment will be used for Cohort 1 Year 5 and Cohort 2 Year 5, with the first year (Cohort 1 Year 5) being used as the primary endpoint.

Invigilators, recruited and trained by the ET, will administer the tests within schools.

Tests will be marked according to a detailed mark scheme by a team recruited and trained by the ET. Ten percent of the tests will be second-marked to ensure consistency and all will be double-entered to confirm input accuracy.

Both invigilators and markers will be blind to condition as they will not have access to any information about allocation.

Secondary outcomes

The existing science attitudes instrument, used in both the efficacy trial (Hanley et al, 2015) and the previous effectiveness trial (Kitmitto et al, 2018) will be utilised but with a strengthened self-efficacy scale. The instrument from the efficacy trial is shown in Appendix 3. The instrument allows pupils' interest in science, including practical work, as well as self-efficacy to be gauged (relating to medium and long term benefits in the logic model).
The science attitudes questionnaire will be completed in-class supervised by class teachers. It will be inputted by the team recruited to mark and input the primary science test.

Follow-up analysis will be conducted for the first cohort of pupils, who will participate in the intervention in the academic year 2022-23. At the end of Year 6, we intend to collect the secondary outcome of science attainment, assessed via a new measure, the Year 6 Science Assessment, developed by the York Trials Unit (YTU), University of York. This new measure will be developed to reflect the current curriculum, have a mix of question types and have an emphasis on "working scientifically". Invigilation, administration and marking details will be the same as outlined for the Year 5 assessment.

In addition, we will assess for any impact on Maths and English by considering attainment based on pupils' Key Stage 2 (KS2) results (English Reading and Maths), which will be obtained from the National Pupil Database (NPD) in Autumn 2024.

Compliance

Compliance will be measured at a class level rather than school level. Definitions of compliance:

Cohort 1 - first year of trial: *Has a (pupil in a) class been taught by a teacher who attended at least 3 out of the 4 full days of training?* This would include a class that (because of long-term sick leave, resignations etc.) has been taught by two teachers who together have attended 3+ training days. For example, the Year 5 teacher attends two days training in the Autumn term then leaves the school; the new teacher TDTS attends at least one further full-day training session.

Cohort 2 - second year of trial: *Did the predominant teacher of a class attend at least 3 full days of training in the first year of the trial?* The predominant teacher will be defined as the teacher who taught the class for the majority of the academic year based on termly updates from each school.

Analysis

The analysis outlined in brief below follows EEF statistical guidance (2018). A detailed statistical analysis plan (SAP) will be produced within three months of randomisation and will be peer-reviewed.

Analyses will be conducted using the principles of intention to treat including all schools and pupils in the groups that they were randomised to, irrespective of whether or not they went on to receive the intervention. Baseline data will be summarised by trial arm and presented descriptively both for schools and pupils as randomised, and as included in the primary analysis. No formal comparison of baseline data will be undertaken, except to report the difference between the groups in pre-test scores as Hedges' g effect sizes and 95% confidence interval (CI).

Statistical significance will be determined at the 5% level and tests will be two-sided. Estimates of effect will be presented as Hedges' g alongside corresponding 95% CIs and p-values. ICC's for pre- and post-tests at the level of the school and class will be presented alongside 95% CIs.

The correlation between average EYFSP score and all outcomes (separately) will be presented, as will the correlation between science attitude and attainment outcomes for each year.

The cohort 1 data analysis will be carried out from September-December 2023, and the cohort 2 data analysis from September-December 2024.

Primary Analysis

The primary analysis will investigate any difference in science test scores between the two arms. Unadjusted scores will be summarised by trial arm. A linear mixed effects model will be used to estimate the adjusted mean difference in scores. School will be included as a random effect and group allocation, average EYFSP score and the minimisation factors (region, ever-FSM) will be included as fixed effects. Ever-FSM will be used as a dichotomous variable at the pupil level (using the indicator EVERFSM_6_P from the NPD) in the analysis rather than dichotomised at the school level as is planned for the randomisation.

This analysis will be repeated for both cohorts separately. The difference between the intervention and control groups in cohort 1 will be the primary comparison.

Sensitivity Analyses

A Complier Average Causal Effect (CACE) analysis will be conducted for the primary outcome to account for non-compliance with the intervention as defined above. An instrumental variable (IV) approach will be taken using randomised group as the IV. This will be repeated in both cohorts separately.

A mixed effect logistic regression model will be run to predict the presence of missing primary outcome data including group allocation, pre-test score and other school- and pupil-level baseline data. Where more than 5% of cases are excluded from the primary analysis due to missing data, the impact of missing data on the primary analysis will be assessed by repeating the analysis on a data set where missing data has been completed using multiple imputation. This analysis will be repeated for both cohorts separately.

Subgroup Analyses

The following subgroup analyses will be conducted for the primary outcome:

- everFSM-eligible pupils (yes/no) – in both cohorts separately

This will be conducted both by including an interaction between FSM and group allocation in the primary analysis model, and by repeating the primary analysis only within the FSM subgroup.

Secondary Analysis

Scores from the science attitudes questionnaire will be compared between the two trial arms. Unadjusted scores will be summarised by trial arm. As for the primary analysis, a linear mixed effects model will be used to estimate the adjusted mean difference in scores. School will be included as a random effect and group allocation, average EYFSP score and minimisation factors (as in the primary analysis) will be included as fixed effects. This will be conducted in both cohorts separately.

Longitudinal follow-ups

The secondary outcomes of Science, Maths and Reading attainment assessed at the end of Year 6 will be analysed similarly to the primary outcome.

Implementation and process evaluation

Research questions

In line with EEF guidance (EEF, 2019; Humphrey et al., 2016) the IPE aims to explore the relationship between delivery and programme outcomes, in particular to provide greater context and understanding of the results of the impact evaluation. The IPE will use a mixed methods approach and address the following questions:

RQ1: To what extent was TDTS implemented as planned?

a. Training
b. Classroom practice

RQ2: What processes are involved for teachers and schools implementing TDTS – what are the main facilitators and barriers?

RQ3: What are the perceptions of teachers as regards TDTS?

a. What are their opinions about training and support, including cascading from colleagues where relevant?
b. What are their views of TDTS strategies and techniques?
c. What impacts has TDTS had on their classroom practice?
d. How has it affected their engagement with and confidence in teaching science?
e. How do they think it has impacted on pupils?

RQ4: How do pupils respond to TDTS?

a. What is their experience of, and reaction to, the different TDTS strategies?
b. What is their experience of practical work in the science classroom?
c. What is their engagement with science lessons?

RQ5: How does TDTS compare with practice in business-as-usual science lessons?

a. What strategies and techniques are used in science lessons?
b. How interested and engaged are teachers and pupils in science teaching and learning?
c. What is the frequency and length of science lessons?

d. What practical science takes place?
e. How much training have Year 5 teachers received in science?

Research methods and analysis

We will use a mixed methods approach incorporating the following elements, which will have been developed and refined during the pre-trial phase. Interview and focus group data will be transcribed and analysed thematically using NVivo software and triangulated with observation and survey data:

Teacher surveys

All teachers involved in the evaluation will be asked to complete an online teacher survey pre-intervention to establish a baseline of school and teacher contextual factors, current science provision (both the amount of science teaching and the strategies used) and teacher attitudes towards and confidence in teaching science. Follow-up surveys will be administered towards the end of the first (Year 5 Teachers) and second year (Year 5 and 6 teachers) of the intervention with additional questions to explore feedback about training sessions, use of the TDTS approach in the classroom and the effect on their confidence and practice of teaching science. Teachers will also be asked about the perceived effects on pupils, including engagement and confidence in their understanding of science.

Pupil questionnaires

At the end of the first year Year 5 pupils would complete a science attitude measure (similar to the instrument used in the previous TDTS trials) along with a questionnaire about their science lessons (to compare TDTS with business-as-usual, and triangulate against teacher feedback and observations). These measures would be repeated towards the end of the second year with the second cohort of Year 5 pupils.

Case studies

We will select two intervention schools in each region to visit. One will be visited twice, in autumn and summer terms, to allow direct comparison of the experience and perception of the TDTS programme towards the beginning and end of the intervention period. The other will be visited once, in the spring term, to pick up on any aspects of the intervention that might be particular to the midway point. The alternative would be to visit each of six schools three times, but this would (a) be burdensome for the school, and (b) restrict the evaluation to a more limited spread of settings. Lesson observations, teacher interviews and pupil focus groups would assess implementation fidelity and the attitudes/engagement of teachers/pupils. A sample of pupils' work would be examined to assess the move to more focused recording of investigations. If earlier recording of lessons was successful in the pre-trial we would record a sample of lessons in another 6 schools (2 teachers/school) to capture shifts across time and minimise "hothouse" effects. Additionally, we will select three control schools where we will ask a teacher to send us samples of pupils' work and (if applicable) recordings of lessons.

Training observations

We will observe one training session per region, to ensure each pair of trainers is observed at least once. The observations will be spread across the first four days of the four-and-a-half-day programme.

Trainer surveys and interviews

After each training day that is observed, we will interview the pair of trainers who delivered the session to get their feedback on how it went. There will be a short online survey for all the remaining trainers after each day to obtain feedback.

Developer interviews

Towards the end of the first year, the developers will be interviewed to obtain their views of how the intervention has been implemented.

Table 4: IPE Methods Overview

Research focus	Data collection methods	Participants/ data sources	Data analysis methods	Research questions addressed	Implementation/ logic model relevance
Trainer feedback (after observed sessions)	Paired semi-structured interview	6 interviews (12 trainers)	Combination of inductive and deductive analysis	1a	Implementation activity (4-day professional development programme)
Trainer feedback (after each delivery phase)	Survey	4 time points (12 trainers per occasion)	Descriptive analysis	1a	Implementation activity
Teacher feedback on training (Collected by developers)	Survey	All teachers (c180) attending TDTS training (collected after each training event)	Frequency counts; Descriptive/thematic analysis	1a 3a 3c	Implementation activity
Pre-randomisation baseline	Survey	All participating teachers (c360) in intervention and control schools		3c 5a-d	Pre-implementation practice
Follow-up at the end of the first and second years	Survey	All participating teachers (c360) in intervention and control schools; in intervention schools only, a member of senior leadership team as appropriate; Year 6 teachers (end of second year only; c360)		1b 2 3 a-e 5 a-d	Comparison of pre- and post-intervention; TDTS and control

Case study school lessons	Lesson observations	Case study schools/teachers (18)	Descriptive analysis (of schedule and fieldnotes)	1b 2 4a-c	Whether teachers are implementing strategies in lessons; confidence; pupil engagement
Case study school teachers	Semi-structured interviews	Teachers (36)	Combination of inductive and deductive analysis	2 3 a-e 5 a-d	Whether teachers are implementing strategies in lessons; confidence
Case study pupil feedback	Focus groups	Pupils in case study schools (18 groups of 4-5)		1b 4a-c	Pupil response to TDTS strategies; triangulation of practices etc.
Case study pupil written recording	Examination of samples of pupils' work (5 pupils x 5 pages/visit)	Pupils in case study schools (c100)	Descriptive analysis (pro forma and fieldnotes)	1b 4c 5b	Evidence of TDTS affecting written work (e.g. focused recording)

Pupil feedback	Survey	All Year 5 pupils in all schools (c8100)	Frequency counts; Descriptive/thematic analysis	1b 4 a-c	Pupil experience of and engagement with science lessons
Developer feedback	Semi-structured interviews	Development team	Descriptive analysis	1 a,b 2	Developer reflections on TDTS programme implementation

Cost evaluation

Data on intervention costs (including training and materials) will be collected from discussions with the development team and from participating schools using cost-specific questions during teacher interviews (case study schools) and follow-up surveys (all participating teachers). Following EEF guidance (EEF 2019), the evaluation team will provide the total cost per school for the intervention as implemented over three consecutive years, and the cost per-pupil-per-school-year.

Ethics and registration

Ethical approval for this study was granted by the University of York Health Sciences Research Governance Committee (HSRGC) in May 2020. All outputs will be anonymised so that no setting or student will be identifiable in the report or dissemination of results. The statistical database will hold non-identifiable data. Confidentiality will be maintained and no one outside of the evaluation team will have access to the database which will be held securely on the department servers.

The evaluators will register the trial with ISRCTN on agreement of the protocol.

Data protection

Data will be handled in accordance with the General Data Protection Regulations (GDPR). Personal data will be processed under Article 6 Section (e) of the GDPR ('Tasks carried out in the public interest') as the research is being conducted to support education provision in the UK (and, if applicable, Special Category data under Article 9(2)(j)). A Data Protection Impact Assessment (DPIA) will be conducted and Data Sharing Agreements will be put in place with schools.

The University of York will be the Data Controller and will also process data. Data subjects are the participants in the evaluation, which includes pupils and teachers in participating schools and the trainers.

Personal data will be processed under Article 6 (1) (e) (*Processing necessary for the performance of a task carried out in the public interest*) and Special Category data under Article 9 (2) (j) (*Processing necessary for ... scientific ... research purposes)* of the General Data Protection Regulation (GDPR; 2018).

All participant data will be treated with the strictest confidence and will be stored in accordance with the GDPR. Identifiable information about participants will be shared by the evaluation team, with the Department for Education, the EEF's archive manager and, in a pseudonymised form, with the Office for National Statistics and potentially other research teams. Matching to the National Pupil Database and other administrative data may take place during this and subsequent research. There will be no international data transfers outside of the EU.

Parent/carers will be informed about the research though an information sheet sent on behalf of the evaluation team by schools to parents/carers. Parents/carers will be asked to return a signed 'withdrawal from research' form if they are unwilling to share their child's data with the

ET and/or they do not wish their child to take part in any assessments, surveys or focus groups. This will apply for both cohorts of Year 5 pupils.

For the purposes of the research, details of participating pupils (e.g. name, date of birth, gender and UPN) will be collected from schools and further details from the National Pupil Database (FSM, EYFSP and KS2 results). The details will be fully specified in the Data Sharing Agreement which will be put in place with participating schools before data transfer.

Schools will transfer data directly to YTU on an encrypted spreadsheet via the University of York's secure file transfer service (DropOff).

A unique trial identification number (Trial ID) will be generated for each participant when their details are entered into the trial management system.

The trial management system and all electronic data will be held on secure University of York servers with access limited to specified members of YTU staff. Paper documents and assessment papers will held securely in a controlled access area in locked cabinets.

The dataset for statistical analysis will hold pseudonymised data and no schools, teachers or children will be identifiable in the report or dissemination of any results.

Electronic data and paper documents including identifiable personal child data will be securely archived and disposed of by YTU 5 years after the end of the study. Pseudonymised electronic data and paper documents will be kept indefinitely.

The University of York’s data protection policy is publicly available at:
https://www.york.ac.uk/records-management/dp/

Personnel

Development team

The Delivery Team is responsible for recruiting and training the trainers, coordinating the training of teachers, recruiting participants in cooperation with trainers/other local partners, and liaising with the Evaluation Team in order to ensure the smooth-running of the evaluation and associated data collection activities.

The Delivery Team comprises:

Bridget Holligan is the Director of Education and Engagement for Science Oxford and has spent her career in the informal science learning sector, with a particular focus on working with primary teachers and pupils in science. She jointly developed and leads the Thinking, Doing, Talking Science projects (2013-23) with Helen Wilson, funded by the Education Endowment Foundation and others. She led the creation of the Science Oxford Centre for primary schools and families, which is founded on the TDTS ethos, and which opened to the public in 2019.

Helen Wilson is an Affiliate Lecturer at Oxford Brookes University, having been a Principal Lecturer in Science Education there. She began her career as a secondary physics teacher and then moved into primary teaching. She then went into Initial Teacher Education, eventually leading the primary teacher training programmes at Oxford Brookes University. As a primary science consultant, she continues her research into the links between creative, challenging primary science lessons and pupils' attitudes and attainment. She jointly developed and leads the Thinking, Doing, Talking Science projects (2013-2023), funded by the Education Endowment Foundation.

Andy Kensley is the Head of Education Outreach for Science Oxford, having formerly been an engineer and project manager for National Grid (and STEM Ambassador) and then a primary school teacher. He leads on the development and delivery of Science Oxford's local CPD for teachers, including courses for STEM Learning and the Primary Science Quality Mark as well as TDTS-based twilight sessions. He is a TDTS-trained trainer and project manager (from 2021) for the TDTS effectiveness trial 2020-23.

This core team will be joined by a number of trainers:

Bryony Turford – Primary Science Geeks

Wendy Precious – Precious Learning Ltd

Rachael Webb – Lancashire County Council

Sarah Earle – Bath Spa University

Alison Trew – Primary Science Teaching Trust

Allie Beaumont – Independent Consultant

Mandy Hodgkinson – East Riding of Yorkshire Council

Nicky Waller, Jane Winter, Joy Parvin – Centre for Industry Education Collaboration, University of York

Evaluation team

York Trials Unit, University of York

Pam Hanley (Co-PI) has an extensive background in education research, including many RCTs at the University of York. Her EEF experience includes other science-related interventions in addition to the TDTS efficacy trial (Let's Think Secondary Science, Sci-napse). She previously worked for CIEC (primary science specialists) and was course evaluator for Science Learning Centre South-East (CPD providers). Pam will be jointly responsible for the day-to-day management and coordination of the trial along with leading on the qualitative aspects of the project.

Louise Elliott (Co-PI) will be jointly responsible for the day-to-day management and coordination of the trial and lead on the impact evaluation. She has been involved in a large number of trials, including several for the EEF and is currently joint Principal Investigator of the 5Rs and Lexia evaluations. She has broad experience of education research and has worked on a wide range of trials covering science, including the efficacy trial of TDTS, literacy and mathematics.

Imogen Fountain has supported many education trials, being responsible for data collection from schools as well as IPE visits and surveys. Her previous trials include TDTS efficacy, Let's Think Secondary Science, Wellcome Trust Primary Science Specialist CPD, ReflectED and Lexia.

Caroline Fairhurst, a senior statistician who has worked on many education trials for the EEF, will oversee all statistical aspects of the trial. Her previous EEF-funded trials include Sci-napse, ABRA, LEXIA and ReflectED and she is currently involved in Maths Champions II.

Professor David Torgerson is Director of the York Trials Unit. He will provide consultancy on methodology and design.

The team will also include a qualitative researcher to support the IPE.

Risks

Risk	Preventative measures/mitigation	Likelihood
Insufficient schools recruited	• TDTS is a promising and low-burden intervention • Inclusion of 4.5 days staff development might be attractive for non-specialists • Development team have previous experience of recruitment	Medium
Attrition of schools	• At the recruitment stage the expectations and commitment of the project will be made clear to schools and they will be required to sign an MoU • Ensure buy-in at head and teaching staff level • TDTS is a promising and low-burden intervention • Regular communication with key contacts throughout the project • Offer control schools financial incentive (first instalment payable after first year) to reduce dropout • Multiple schools from the same MAT will be eligible to participate (subject to certain conditions, see Contamination section below)	Low
Attrition of teachers	• All Year 5 teachers will be invited to be involved in TDTS to allow for increased attrition over two-year project • Main staff loss will be between first and second year as staff leave school or change Year Group. Leaving teachers will be expected to hand over all TDTS materials to their replacement and, where possible, cascade their knowledge. New teachers will be expected to attend any outstanding training sessions. • TDTS techniques should not significantly increase workload • Check staff changes regularly with key contact:new staff in first year to receive training as soonas possible • MoU commitment to have initial Year 5 TDTS-trained teachers teaching Year 5 in second year of trial wherever practicable	High
Attrition of pupils	• Allowed for 15% pupil-level attrition in the sample size calculation over two-year project • Keep number of outcome measures to a minimum and as engaging as possible	Medium

Risk	Preventative measures/mitigation	Likelihood
High drop-out from intervention or poor implementation	• Regular CPD sessions and supporting resources should assist strong implementation and mitigate against withdrawal • Poor implementation should be picked up by the process evaluation and will inform the evaluation	Low
Project staff turnover	• York Trials Unit has a range of experienced staff to substitute if necessary • All procedures will be documented to assist any replacement personnel	Low
Delays in schools providing necessary documentation	• Provide some details as a prerequisite of randomisation • The evaluation team has extensive experience of chasing up data/documentation from schools • The team includes dedicated project support	Low
Contamination	• Schools whose teachers have received TDTS training at any time will be ineligible • Schools from the same MAT will be eligible to participate on the agreement and understanding that they do not share practices from intervention schools to control schools during the trial.	Medium
School closures	• In-school fieldwork does not start until Spring 2022, when hopefully the Covid-19 pandemic will be more under control and/or schools will have better-developed alternative strategies	Medium

Timeline

Table 5: Timeline

Dates	Activity	Staff responsible/ leading
December 2019 - February 2020	Set Up Meetings 1 & 2 and IDEAS meeting	ET/Developer
May 2020	Ethical approval granted	ET
Early-mid 2020	Recruitment of Trainers	Developer
April 2021	ISRCTN registration	ET
July 2021	Trainers experience the TDTS course as participants	Developer
September &November 2021	Trainers attend train-the-trainer course	Developer
July, September &/ November 2021	Evaluation of train-the-trainer course – Observation / Interviews	ET
September – December 2021	Recruitment of pre-trial schools	Developer
January– July 2022	Trainers deliver TDTS to teachers in pre-trial schools	Trainers
January– July2022	Evaluation of Trainers delivery of TDTS to teachers in pre-trial schools	ET
January– June 2022	Recruitment of trial schools	Developer
January – June 2022	Randomisation of trial schools	ET
September – October 2022	Cohort 1 collect pupil details	ET
September 2022	Cohort 1 Teacher baseline survey data collection	ET
September 2022 – June 2023	Cohort 1 Intervention period (delivery of TDTS to teachers in trial schools)	Trainers
September 2022 – April 2023	Evaluation of trainers delivery to trial schools	ET
October 2022 – May 2023	Cohort 1 Case study visits to schools	ET
June/July 2023	Cohort 1 science testing	ET
June/July 2023	Cohort 1 pupil survey and attitude to science questionnaire collection	ET
June/July 2023	Cohort 1Teacher survey follow-up data collection	ET
January 2023 – September 2023	Cohort 1 Code/analyse IPE data	ET

July - September 2023	Cohort 1 data analysis	ET
September – October 2023	Cohort 2 collect pupil details	ET
30th November 2023	Cohort 1 draft report submitted to EEF	ET
31st May 2024	Cohort 1 submission to EEF of final report	ET
June/July 2024	Cohort 1 longitudinal science testing (Year 6) Cohort 2 science testing	ET
June/July 2024	Cohort 1 (Y6) and Cohort 2 (Y5) Teacher survey follow-up data collection	ET
September– December 2024	Cohort 1 longitudinal analysis - Year 6 testing and KS2 results	ET
July– December 2024	Cohort 2 data analysis	ET
15th December 2024	Draft addendum report (Cohort 2 and Cohort 1 longitudinal including KS2 results) submitted to EEF	ET
29th March 2025	Submission of final addendum report to EEF	ET

References

Abrahams, I., Bennett, J., Cheung, A., Elliott, L., Hanley, P., Oberio, Z., ... Turkenburg, M. (2014). *Evaluation of the impact of a Continuing Professional Development (CPD) course for Primary Science Specialists: Final report*. London: The Wellcome Trust.

Bax, S. (2002). The social and cultural dimensions of trainer training. *Journal of Education for Teaching 28*(2), 165-178.

DfE (2013). *The National Curriculum in England: Key Stages 1 and 2 framework document*. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study/national-curriculum-in-england-science-programmes-of-study

Dichaba, M. M., & Mokhele, M. L. (2012). Does the Cascade Model Work for Teacher Training? Analysis of Teachers' Experiences. *International Journal of Educational Sciences, 4*(3), 249-254. doi:10.1080/09751122.2012.11890049

EEF (2019). *Implementation and process evaluation guidance for EEF evaluations*. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_guidance.pdf

EEF (2019). *Cost evaluation guidance for EEF evaluations*. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/Cost_Evaluation_Guidance_2019.12.11.pdf

Gask, L., Coupe, N., & Green, G. (2019). An evaluation of the implementation of cascade training for suicide prevention during the 'Choose Life' initiative in Scotland - utilizing Normalization Process Theory. *BMC Health Services Research, 19*(1), 588. doi:10.1186/s12913-019-4398-1

Hanley, P., Slavin, R., & Elliott, L. (2015). *Thinking, Doing, Talking Science: Evaluation report and executive summary*. London: Education Endowment Foundation.

Harlen, W., & Qualter, A. (2008). *The teaching of science in primary schools*. London: Fulton.

Hayes, D. (2000). Cascade training and teachers' professional development. *ELT Journal 54*(2), 135-145.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). *Implementation and process evaluation (IPE) for interventions in education settings: An introductory handbook*. Available at https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Guidance_Final.pdf

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. International Journal of Science Education, 29(7), 871-893.

Kitmitto, S., González, R., Mezzanote, J., & Chen, Y. (2018). *Thinking, Doing, Talking Science: Evaluation report and executive summary*. Available at https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/TDTS.pdf

Ray, M. L., Wilson, M. M., Wandersman, A., Meyers, D. C., & Katz, J. (2012). Using a training-of-trainers approach and proactive technical assistance to bring evidence based programs to scale: An operationalization of the interactive systems framework's support system. *American Journal of Community Psychology 50*(3-4), 415-427.

Palak Roy, P., Rutt, S., Easton, C., Sims, D., Bradshaw, S. & McNamara, S. (2019). *Stop and Think: Learning Counterintuitive Concepts: Evaluation Report.* Available at: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Stop_and_Think.pdf)

Saghaei, M., & Saghaei, S. (2011). Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials. *Journal of Biomedical Science and Engineering, 4*(11), 734.

Slavin, R., Lake, C., Hanley, P., & Thurston, A. (2014). Experimental evaluations of elementary science programs: a best-evidence synthesis. *Journal of Research in Science Teaching, 51*(7), 870–901.

Turner, F., Brownhill, S., & Wilson, E. (2017). The transfer of content knowledge in a cascade model of professional development. *Teacher Development, 21*(2), 175-191. doi:10.1080/13664530.2016.1205508

Wedell, M. (2005). Cascading training down into the classroom: The need for parallel planning. *International Journal of Educational Development, 25*(6), 637-651.

Appendix 1: Changes since the previous EEF evaluation

Appendix table 1: Changes since the previous evaluation

	Feature	Efficacy	First effectiveness stage	Second effectiveness stage
Intervention	Intervention content		No change	The course content has been updated to more closely align with the Year 5 curriculum
	Delivery model	Developer led training of teachers	Developers Train-the-trainers to train teachers with sessions spread throughout the intervention year.	The train-the-trainers model adds: TDTS programme delivered to trainers; all training on how to train teachers delivered before the intervention year; and pre-trial experience before trainers are selected to participate in the main trial. Improved quality assurance and trainer resources.
	Intervention duration	5 days of teacher CPD spread across one academic year. Delivery to pupils over the whole academic year.	CPD days reduced to 4 but teacher delivery to pupils still over one academic year.	CPD of 4.5 days - additional final half-day consolidation CPD day. Sessions shifted slightly earlier in academic year. Delivery to pupils still over one academic year.
Evaluati	Eligibility criteria	Schools in Oxfordshire	Schools were located across 7 regions of England.	Schools located across 6 regions of England. (Not yet confirmed but not intended to be the same as the previous effectiveness trial).

o n	Two teachers per school were required to attend the training (unless otherwise arranged)	All Year 5 teachers were required to attend the training (minimum of two); where there was only one Year 5 class another teacher was required to attend (ideally the subject lead). There was no set minimum number of pupils.	All Year 5 teachers are required to attend the training (minimum of two); where there is only one Year 5 class another teacher is required to attend (ideally the subject lead). Schools must have a minimum of one full class of Year 5 pupils (mixed year group classes cannot take part).
Level of randomisation	School level	No change	No change
Outcomes and baseline	Baseline: Science Knowledge Questionnaire Year 4 Post-test: Science Knowledge Questionnaire Year 5 Science Attitude Questionnaire	Baseline: – KS1 Maths and KS1 Reading/writing Post-test: As efficacy The questionnaire content was unchanged but the measure was split into two indices.	Baseline: Early Years Foundation Stage Profile (EYFSP). Proposed post-test: new Year 5 Science Assessment (currently under development) Possible adaptations to attitude measure depending on findings from pre-trial phase.
Control condition	Business as usual	No change	No change

Appendix 2: Brief overview of TDTS 4.5 Day training sessions

Training Day	Sessions include	Brief content notes
Day 1 Materials	Introduction to the TDTS project	Overview of Teacher Folder and website; Background and evidence; Ethos of CPD and mapping to curriculum content
	Various practicals	e.g. Paper Flowers & Protect an Egg practicals
	Challenge and Higher Order Thinking (HOT) in science	Definitions and evidence
	Introduction to the Bright Ideas Time (BIT): Odd One Out (OOO)	Examples of pupil responses and teacher feedback
	The Science of Materials	States of matter: role play
	Practical Prompts for Thinking	e.g. Use PPT to go from 'wow' to 'wonder' and HOT
GAP TASK – trying an Odd One Out with pupils in science lessons		
Day 2 Forces	Introduction and Practical Prompts for Thinking (PPT)	TDTS strategies reminder
	Various practicals	e.g. Which shoes have the best grip & The Marble Maze
	Sharing of good practice	Discussion: OOO gap task teacher feedback
	The Science of Forces	Pushes/pulls and getting a 'feel' for Newtons
	Higher Order Questioning (HOQ) BIT: Big Question (BQ)	HOQ and inclusive challenge BQ examples and pupil feedback
	More PPTs for HOT	
GAP TASK – trying a Big Question discussion with pupils		
Day 3 Earth & Space	Introduction BIT: The Big Question (BQ) cont.	TDTS strategies reminder Constructivist view of learning
	Further thinking about Higher Order Questioning (HOQ)	Types and examples of questions teachers ask Planning (& celebrating) HOQ and HOT: OOO, BQ
	Sharing of good practice	Discussion: BQ gap task teacher feedback
	Various practicals	e.g. Strongest legs & Glider Challenge
	PPTs for HOT in Earth and Space Galaxies – seeing history	Scale and use of models
GAP TASK – practicals for Higher Order Thinking with Focussed Recording		
Day 4 Living Things	Introduction The Bright Ideas Time (BIT)	TDTS strategies reminder E.g.s of OOO, PMI, BQ and making your own Questioning – value of open and closed questions
	Life Cycles	Researching secondary sources – HOTS Observation over time - HOT and FR
	Sharing of good practice	Discussion: Practical and FR gap task teacher feedback
	Various Practicals	e.g. Create an Animal & Seed dispersal
GAP TASK – crafting a lesson		

Day 5 Leading TDTS in your school	Sharing of TDTS practice	Based on 'crafting a lesson' gap task
	Leading TDTS in Your School Part 1: your classroom practice Part 2: working with others	Link to OFSTED 2019 primary science research Discussion: how TDTS practice addresses issues Small Changes Big Impact: value of TDTS evidence Discussion: effective staff meetings, dissemination