Department for Education

TLIF Evaluation: Tom Bennett Training Project Evaluation

Final Report

July 2022

Authors: Ben Willis, Bronwen Maxwell and Sarah Reaney-Wood (SIOE)

Julie Nelson, Suzanne Straw, Jack Worth, Dawson McLean and Ruth Staunton (NFER)

Government Social Research

Contents

List of Tables

Key Findings Summary

- The Tom Bennett Training (TBT) project aimed to increase confidence in the behaviour management practices of leaders and teachers in order to improve pupils' behaviour. Two types of courses were delivered - Running a Room (aimed at teachers) and Running a School (aimed at senior leaders). Both comprised a two-day workshop, followed by two 'Booster' workshops over six months and were supported by a course handbook and access to an online portal. Implementation of learning was supported by course activities such as baseline behaviour audits and gap-tasks but there was no direct support for implementation.
- The majority of teaching staff and senior leaders interviewed described at least moderately increased confidence and competence in employing effective behaviour management strategies at a classroom or whole-school level. Interviewees who attended the Running a Room Course, especially early career teachers (ECTs), were most likely to attribute improved confidence and effectiveness directly to the TBT project.
- There was strong evidence pupil behaviour became more prioritised and behaviour policies more embedded across the whole school workforce in participating schools. However, many schools had begun to implement behaviour management reforms prior to the TBT project.
- Some interviewees felt their school placed more emphasis on nurture-based approaches, opportunities to recognise positive pupil behaviour (for example reward systems) and increased pastoral activities. This was perceived to contribute towards a more positive school culture, improved staff-pupil relationships, heightened wellbeing (pupils and staff) and improved pupil behaviour. Certain schools also reported reductions in exclusions and detentions as well as pupils feeling more at ease in the school environment.
- Although some interviewees felt involvement in the TBT project had improved their satisfaction with teaching (sometimes significantly), there was very limited evidence of influence on wider outcomes such as attitudes towards CPD or likelihood to remain teaching.
- Data for this evaluation report is drawn from three project manager and fifteen participant telephone interviews, five school case studies and DfE management information analyses.
- DfE's management information analysis shows that:
    - recruitment targets for total participants (target 720; recruited 756) were met. However, only 59% of those recruited participants were from priority schools (Ofsted category 3 or 4) compared to a target of 70%

    - drop-out from the programme was high (109), however this is likely to mainly be due to three schools withdrawing from the programme. The proportion of participants in priority schools who withdrew was lower than for participants in non-priority schools.

- Free CPD and TBT's significant existing social media presence was considered key to recruiting target schools. TBT's pragmatic, flexible and non-judgemental approach to the provision of support were linked to high project satisfaction.
- The quality and effectiveness of TBT provision was perceived by most interviewees to have been very high. Interviewees particularly valued the provision of a set of overarching principles that they could customise to affect sustainable change within their own contexts.
- The findings indicate that the content and flexibility of the TBT project is a useful guide for future low cost, low intensity behaviour management training projects. However, it is unclear if a significant scale-up in its current form would be viable, given the importance interviewees placed on Tom Bennett's charismatic presentation style and leadership of most delivery.
- Findings from the analysis of data from the SWC on retention were mixed. At the teacher level there was some evidence to suggest that the TBT project may have had a positive impact on teacher retention in the state-funded sector two years after baseline. It is likely that the early positive impacts on teacher retention found were due to non-observed systematic differences between TBT participants and non-participants, and/or the limitations from assigning participants to a baseline year. At the school level there was one significant negative finding for retention in state-funded teaching three years after baseline. However, no other findings (positive or negative) were observed for any of the other retention measures. The negative finding for state retention appears to have been caused by a notable increase in the retention rate in comparison schools in this year (rather than the difference being caused by a change within treatment schools).
- Findings from the SWC on progression at the teacher-level provided evidence to suggest that over time, the TBT project may have had some positive impact on teacher-level progression in state funded schools at year three and in the same Local Authority District (LAD) at years two and three. Teachers that did progress likely did so at schools within the same LAD. As the initial progression effect was small and for LAD grows overtime it is plausible that this progression could be attributable to the TBT project. For school-level progression, no impacts (positive or negative) were observed.

Glossary of terms

Achieving Excellence Areas - AEA categories are DfE classifications of educational performance and capacity to improve by local authority district (LAD). They split areas into six categories from "strong" category 1 areas to "weak" category 6 areas.

Priority areas - category 5 or 6 Achieving Excellence Areas (AEAs) local authority districts, including the 12 Government Opportunity Areas - areas identified as having weakest performance and least capacity to improve.

Priority schools – schools with an Ofsted judgement of 3 or 4 (Inadequate or Requires Improvement (RI)).

Running a Room course: aimed at senior leaders and focused on whole-school behaviour management policy and practices

Running a School course: aimed at teachers and focused on their classroom practice in relation to behaviour management.

School Workforce Census (SWC): a statutory collection of individual level data on teachers and support staff from local authorities, local authority-maintained schools and academies.

Teaching and Leadership Innovation Fund (**TLIF)** — DfE programme (2017-2020) aimed at improving pupil outcomes and supporting pupil social mobility by improving teaching and leadership in priority areas and schools through outcome-focused, evidence-based and innovative professional development provision.

1. About Tom Bennett Training and the evaluation

Section 1 outlines the intentions at the beginning of the Tom Bennett Training (TBT)[1] project. Section 2 outlines how patterns of recruitment played out, and Section 3 describes how the project was adapted during delivery and implementation. TBT aimed to deliver a project that would increase confidence in behaviour management practices at leadership and classroom level in order to improve pupils' behaviour within classrooms and across schools. The TBT project was intended to be delivered in two different course formats:

1. *Running a School* - aimed at senior leaders and focused on whole-school behaviour management policy and practices

2. *Running a Room* - aimed at teachers and focused on their classroom practice in relation to behaviour management.

The project was intended to be delivered flexibly, on a whole-school basis (involving teaching staff), for groups of senior (and sometimes middle) leaders or self-referring teachers. This meant the project could be accessed in one of three ways:

- Model A): Leader(s) only – attending a Running a School course
- Model B): Teacher(s) only – attending a Running a Room course
- Model C): Whole-school projects – a combination of A and B. As detailed below there was, as intended, some variation in the model of delivery for whole-school projects.

All three course models were intended to span six months and include the following features and inputs:

- **an initial two-day course** that taught teachers and/or leaders the 'solid basics', focused on the most useful theories, strategies and principles that work in 'most schools'. Blocks of time were also protected to allow for peer-to-peer conversations and/or networking, and to produce individual and/or school-level plans. This process was then intended to inform and drive selection of contextually tailored 'projects/homework' that could be taken forward in participants' own school settings. A course-specific workbook was provided to support delivery that could be referred back to following course delivery.
- **a one-day Booster** run three months after the initial two-day course, followed by a **second one-day Booster** three months after the first. Both focused on

[1] Delivery was through Tom Bennett Training, but the contract was held by Anvil Education Limited.

the principle that training needs to be embedded into practice, revisited and consolidated.

- an **online community portal** where participants could talk to each other and discuss their issues anonymously. In addition, there were due to be online 'drop-in clinics' that afforded participants the opportunity to ask providers questions. The portal also gave participants access to other **project materials and resources.**
- **specialist analytic technology** to monitor participant engagement with the online platform, with the data used to support delivery by enabling targeted encouragement for engagement and interaction.

Every initial two-day training session and most Boosters were led by Tom Bennett, supported by a wider a pool of consultants sourced nearby to participating schools.

1.1 Theory of change

The project logic model shown in Appendix A was created by the evaluation team and was reviewed by DfE. The logic model was based on the theory of change (ToC) submitted by TBT as part of its bid; TBT's understanding of the project's underlying rationale, activities, outputs and anticipated outcomes; and subsequent conversations with the DfE project team.

The underpinning rationale for the ToC was based on an assumption that participants were committed and empowered to make changes and implement course strategies systematically and thoroughly. On that basis, it was theorised that participation in the project would lead to a range of positive intermediate outcomes at school, staff and pupil level which, in turn, would lead to longer-term impacts. Examples of intermediate outcomes for leaders included increased confidence in leading behaviour management and improved quality of leadership of behaviour management. Outcomes for teachers included deploying effective behaviour management strategies in the classroom and, for pupils, improved behaviour and discipline. The intended longer-term impacts were improved retention and progression of leaders and teachers and increased pupil attainment.

The rationale for TBT's ToC drew heavily on Tom Bennett's independent review of behaviour in schools - 'Creating a Culture' commissioned by DfE (Bennett, 2017), which recommended that school leaders (and teachers) had access to training in a range of behavioural strategies and examples of best practice in the school system. In addition, the project was underpinned by several established strategies for classroom and behavioural management (for example, Marzano *et al.*, 2003) as well as evidenced training techniques (such as Smith *et al.*, 1978; Deans for Impact, 2016). Doug Lemov's (2011) work around good practitioner practice also informed

project delivery and content, particularly around interpersonal skills, and Lemov acted as an ongoing external consultant to the project.

1.2 Contextual factors

The TBT project was one of ten DfE-funded TLIF projects. The DfE wished to test out how effectively a variety of different CPD approaches could meet project-specific and fund-level outcomes; therefore, each of the ten projects were commissioned to be intentionally different in design, scale, scope and delivery method. At fund level, the evaluation sought to compare and contrast the relative effectiveness of these projects in meeting their stated aims and objectives – taking into account a range of factors related to their differences. These included:

- **impact focus and target group:** the TBT project had a behaviour management focus and targeted senior leaders and teachers from Ofsted category 3 and 4 schools
- **phase supported:** the TBT project supported both primary and secondary schools
- **per-participant cost:** relative to the other TLIF projects, the TBT project was low cost
- **intensity of the delivery model:** relative to the other TLIF projects, the TBT project had a light-touch delivery model
- **range of delivery modes:** the TBT project had a moderate range of delivery modes relative to other TLIF projects.

In the fund-level report, we take the TBT project's contextual factors into account as we compare its progress in achieving outcomes with the progress made by the other TLIF projects.

1.3 Evaluation methodology

The aim of the evaluation was to undertake a process and impact evaluation to explore indicators of effectiveness and perceptions of outcomes (including teacher and/or leadership quality, pupil behaviour/wellbeing and school culture), and to measure impacts (teacher and senior leader retention and progression) – see Section 4.2, Tables 1-4 for full details). The objective was to draw out learning and best practice, test out the project's theory of change, and identify implications for the fund-level assessment, as well as educational policy and practice more broadly. Our original evaluation design also included an impact evaluation to assess the impacts of the project on pupil attainment. However, due to partial school closures as a result of the Covid-19 pandemic, and the cancellation of Key Stage 2 assessments and

GCSE examinations for the 2020 cohort, DfE decided to remove this aspect of the evaluation. Therefore, there is no longer a pupil impact analysis aspect to the evaluation. In addition, the evaluation sought to inform learning and best practice for the ongoing development of the project, testing out the ToC and identifying implications for the fund-level assessment, as well as educational policy and practice more broadly.

This final report draws on secondary data from the School Workforce Census (SWC[2]), and rich qualitative data. It provides a measure of the project's success in achieving the TLIF programme's impacts (SWC and qualitative data), perceptions of outcomes (qualitative data) and project-specific outcomes (qualitative data).

The rich qualitative data aids understanding of the recruitment, delivery and implementation factors that influenced achievement of these impacts and outcomes. The report explores the links between inputs, outcomes and impacts, analysing the appropriateness of the project's ToC in achieving desired results. A summary of the demographic details of the fifteen case-study schools involved, along with the course they were involved in, is presented in Appendix B. The evaluation data sources underpinning this report are outlined below.

Comparison of secondary data

- a comparison of secondary data from the SWC for TBT participants, and for a matched group of non-TBT participants3. TBT participants were identified via project MI data, which was collected by DfE and shared with NFER.

Observation of project delivery

- observation of Running a School delivery by the evaluator in March 2018

Delivery perspectives

- three telephone interviews: two with the TBT project leads (Tom Bennett and the Director of Operations) in April 2017 and October 2018, and a final interview in March 2020 with Tom Bennett

School perspectives from fifteen schools in three formats

- format 1: five single school telephone interviews: four with a participating

[2] This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.
[3] Non-FPL participants were defined as any teacher who was not enrolled on the FPL project or any other TLIF intervention.

senior leader and one with a newly qualified teacher (NQT)[4]

- format 2: five school telephone case studies, involving interviews with a participating senior leader and a non-participating teacher or middle leader[5] – ten interviews in total
- format 3: five school case-study visits involving a combination of focus groups/interviews with: participating (and non-participating) teachers; pastoral and teaching support staff; special educational needs coordinators (SENCOs); senior and middle leaders, as appropriate to the case, and pupils.

The duration of data collection ranged from a 20-minute pupil focus group in format 1 through to a 90-minute interview with a senior leader in format 2. Typically, senior leader participant telephone interviews lasted around one hour and indirect participant teacher interviews approximately 50 minutes. Across all formats, as part of the interview schedules, school staff were asked a small number of Likert survey-style questions pertaining to individual characteristics (e.g., level of experience in teaching/role/at the school) and to assist with comparisons in perceived levels of pupil misbehaviour and effectiveness of behaviour management policies at a whole-school and classroom level, pre- and post-involvement in the project. This was a feature of the TLIF evaluation's qualitative-only projects, and this data was only collected because these projects did not collect survey data from participants.

Description of SWC matching and analysis methods

Appendix D describes the methods used for matching MI data to SWC data, and for constructing a comparison group. Appendix E describes the results of the impact analysis. In summary, the steps were as detailed below.

1. The MI data was matched to the SWC using Teacher Reference Numbers (TRNs), names and dates of birth. This matched 78 per cent of TBT participants as recorded in the MI data with at least one record in the SWC.

2. TBT participants were matched with non-participants using propensity score matching. Matching for the full sample used teacher and school characteristics (age, gender, years of experience, Ofsted rating, etc. – see Appendix D for the full list) observed in the baseline year, where baseline year for TBT participants was defined as the year the teacher was recruited to the project.

[4] Attendance records indicated the interviewee was a senior leader but, in reality, they were an NQT- this gave different, but very valuable insights into how the course was accessed and how course attendance influenced their classroom practice.
[5] Despite requests that the second interviewee be an individual not directly involved in training, in certain instances participants had received at least some direct training. This was particularly the case for schools involved in the Running a Room course which had a whole-school focus.

3. The retention rates in state-sector teaching of participating teachers were compared with the matched comparison group using a logistic regression model, one, two and three years after baseline and controlling for the variables used for matching. The same process was followed to estimate the impact on retention within the same school/local authority (LA)/ challenging schools[6].

4. Differences between the groups in progression rates (to middle/senior leadership) within the profession and within the same school/LA/challenging schools were estimated using a similar model as in step 3.

5. Similar analysis was then performed at the school level. Project-participating schools were matched with non-participating schools using propensity score matching. Matching for the full sample occurred on the basis of school characteristics (school phase, Ofsted rating, etc. – see Appendix D for the full list) observed in the baseline year, where baseline year was defined as the academic year that recruitment to the programme started.

6. The retention rates at whole-school level in state-sector schools among teaching staff in the treatment and matched comparison schools were compared using a logistic regression model, one, two and three years after baseline and controlling for the variables used for matching. The same process was followed to estimate the impact on retention in the same school, retention in the same LA, retention in a challenging school, progression within the profession, progression in the same school, progression in the same LA and progression in a challenging ***school.***

There were significant differences between the treatment and potential comparison group of teachers in this project. The age and experience profile of treatment teachers matched fairly closely with potential comparator teachers (at least more so than other projects), but treatment teachers were much more likely than comparison teachers to be working in more deprived schools, lower-attainment AEA category 5/6 schools. These differences are all outlined in detail in the methodology Appendix D, where a full set of characteristics for treatment and control teachers can be found, before and after matching.

Limitations of the matching are also outlined in the methodology Appendix D – since other confounding variables are unobserved, it is only possible to speculate on what might be associated with retention/promotion and selection into the project. Variables such as education, and other key psychological variables such as ambition and conscientiousness may be associated with seeking CPD training and also associated with retention and progression outcomes, but these are generally difficult

[6] Challenging' schools were defined as schools rated by Ofsted as 'requires improvement' or 'inadequate'. A teacher was defined as remaining in a challenging school if they either stayed within the school they were in at baseline, or moved to another school which was rated 'requires improvement' or 'inadequate'.

to observe and control for. Everything within the scope of the School Workforce Census that was likely to have affected the outcomes (including additional public, non-SWC school-level data that was brought in separately) was included in the matching and the regression analysis. Full details of which variables were included, how they were defined, and differences between treatment and comparison groups can be found in the methodology Appendix D.

Sampling principles

The qualitative **sampling principles** agreed with DfE (Appendix C, were designed to maximise learning from the evaluation and prioritised sampling of TLIF target schools (see Section 2.1). However, recruiting participants to the evaluation was challenging and, in order to ensure the final sample (see Appendix B) mirrored the sampling principles as closely as possible, it was necessary to contact over 50 schools. Recruitment of schools in the most challenging circumstances ('Key Performance Indicator' – KPI – target schools) was particularly difficult.

The achieved sample was broadly reflective of the sampling principles. Roughly, equivalent numbers of primary (6) and secondary schools (8)[7] participated in the telephone interview-based data collection. There was an additional focus on secondary schools (4:1) for the more intensive school case-visit work for a variety of reasons, including: greater numbers participating overall; a higher proportion with target status (see Section 2.1); more complex behavioural issues to address; and, most critically, to permit the research team to gain sufficient insights from the Running a Room courses.

Three schools involved in the evaluation participated in the Running a Room course and twelve in the Running a School course. This was reflective of the far higher numbers of Running a School courses run. Overall, only just over half of the schools involved in the evaluation were target schools. Therefore, some caution needs to be exercised in forming judgements about the findings emerging from the data. For example, it might be that schools that were in greatest need of assistance (e.g., those with the most challenging baseline behaviour issues), and that had the most to gain from TBT's project, did not feel able to be involved in the evaluation simply due to the particularly challenging nature of their schools.

Qualitative data analysis

Interview transcripts were uploaded into the qualitative data analysis software package NVivo and coded using an analysis framework based on the logic model headings (see Appendix A). Analysis was conducted looking at both fund-level and individual-level project outcomes.

[7] Please note there was one all-through school.

1.4 Focus of this report

This final report focuses specifically on:

- **Section 2 – Recruitment and retention:** whether TBT met its targets for school and participant recruitment, and the factors that supported or impeded this
- **Section 3 – Delivery and implementation:** whether this progressed according to plan; what worked well and not so well, and what lessons could be learned for future continuing professional development (CPD) offers
- **Section 4 – Perceived outcomes and impacts of the provision:** the extent to which TBT met, or had the potential to meet, the TLIF programme's expected outcomes and impacts, and its own bespoke project outcomes
- **Section 5 – Sustainability:** discussion of the potential for sustainability of new ways of working, new learning and outcomes in schools, which had come about through involvement with the TBT project
- **Section 6 – Evaluation of the TBT theory of change**
- **Section 7 – Summary and indicative implications for policy and CPD development.**

2. Recruitment and retention

2.1 Recruitment targets

The initial target was for a minimum of 720 participants to be recruited from at least 100 primary and secondary schools during the programme, with a minimum of 70 per cent of participants sourced from target schools (i.e., Ofsted category 3 or 4). Management Information (MI) data shared by the DfE revealed that, over the course of the project, TBT recruited more participants (756) from fewer schools (75), something that was agreed with DfE (See Appendix F). Fifty-nine percent of all participants recruited were from target schools so the target of a minimum of 70% was not met.

DfE MI data, revealed 109 'teachers dropped out' meaning 647 participants completed the project. The MI data was not broken down by course attendance so there was no way of knowing the extent to which participants dropped out from either the Running a Room or Running a School course. However, the MI data stated that three schools withdrew during the course of the project. If any of these were participating in the Running a Room course (where large numbers of participants were frequent), it might account for the relatively high number of drop-outs. It is important to note that retention of participants from target schools was higher than from schools that did not meet the target criteria. This suggests that schools outside of target categories may have been less committed to the programme and arguably less suited to it. Analysis of the following additional MI data can be found in Appendix F:

- school phase
- schools by region
- schools by AEA category
- schools by Index of Multiple Deprivation Decile
- participant role.

As outlined in Section 1.1, TBT Project Leads attempted to operate as flexibly as possible in order to fulfil school preferences for delivery. This meant the number of participants and schools represented varied across the courses delivered. The composition ranged from all attendees being from a single school, to collections of unconnected schools within a particular region, through to more formal linkages such as school representatives from across a specific multi-academy trust (MAT).

The Running a School course (Model A) did, as intended, recruit senior leaders (and some middle leaders) in regional groups or leaders from across an individual MAT. Recruitment to the Running a Room course (Model B) predominantly occurred when

school leaders enrolled their teaching staff after they had completed the Running a School course. The demand for this course from individual teachers was extremely low, to the extent that TBT made the judgement that courses aimed at this group were not financially viable. In the small number of instances where demand came from individual teachers or NQTs, arrangements were made for them to be incorporated into scheduled Running a Room days at a different, nearby school setting.

Recruitment to Model C, 'whole-school projects', was intended to involve a school committing their senior leaders to the Running a School course, followed by signing their teaching staff up to the Running a Room course. However, in reality, recruitment and the consequent delivery model were sometimes less clear-cut. Despite TBT Project Leads claiming to not '*consolidate the two groups together* [as part of] *the programme*', there was a pragmatism that was intent on arriving at a model that worked best for individual schools. For example, as project delivery evolved, it was recognised by TBT that approximately half of the content from Running a Room and Running a School courses was identical in terms of *'basic principles*'. This led TBT to invite senior leaders to Running a Room classes in cases where they had not previously attended a course - for clarity, these additional senior leader numbers did not count towards the KPIs.

2.2 Recruitment methods

Social media, using Tom Bennett's personal profile, was perceived to be the most important marketing collateral. System leaders were also contacted directly through Tom Bennett's networks, courses advertised on the TBT website and a few schools were targeted by regional DfE officers. Word of mouth became more important over the duration of the project.

The delivery team confirmed that the predominant means of recruiting schools and participants was via Tom Bennett's significant social media profile:

> Having good name recognition was very useful, and having social media reach was extremely useful as well, because it enabled low-cost comms access…that was a huge asset, because it enabled us to…short-circuit some of the challenges which you can experience getting [into] some schools.

Tom Bennett conjectured that projects set up to benefit the schools '*that need it most*' often have the '*greatest difficulty…reaching out*', due to school's '*administrative*' and '*managerial*' structures being overloaded responding to '*crisis*' after '*crisis*'. On balance, TBT argued that their social media presence was '*quite useful in getting around the door*' [of target schools], requiring only a member of a

school [or an associate of] to become aware of it through a 'Twitter' or 'Facebook' post. However, a downside was that the '*vast majority*' of initial enquiries in year 1 were from contexts that did not meet the eligibility criteria - so it was necessary for TBT to do quite a lot of filtering. Interviewee data reinforced that social media was the primary method of recruitment, although smaller numbers also identified website searches for behaviour management training and being informed by senior colleagues. Tom Bennett also directly approached system leaders through the networks he had developed over several years in various high-profile roles including Director and founder of ResearchED[8] and leading the Initial Teacher Training - Behaviour Review Group on behalf of the DfE (2015-17):

> Focussing on the big-ticket players, the heads, the multi-academy trust heads and so on, get them to send their school leaders and get those leaders onto courses, get them so happy about it they want to then set up the schools [for whole-school courses involving teachers]. - *TBT Project Lead*

Courses for individual teachers were advertised through the TBT website and Tom Bennett's social media platform (as with the senior leader courses). There was no evidence of anything more targeted for individual teachers. By the time of the second Project Lead interview (October 2018), word of mouth from satisfied participants was perceived to be an additional recruitment method. This was particularly the case with regard to schools having completed the Running a School course signing up their whole school staff to the Running a Room course, where capacity permitted this.

Finally, a minority of target group school leaders reported being approached directly by DfE about the 'opportunity' to become involved in the TBT project. Despite conveying a genuine enthusiasm to become involved, one interviewee suggested that involvement was non-negotiable owing to the circumstances their school was in.

2.3 What influences effective recruitment?

A number of themes emerged through the analysis, which had acted as either enablers or barriers to recruitment. They included:

- project set up, commissioning and contracting
- using key strategic contacts
- supporting engagement in the project
- geographic location.

[8]ResearchED is a teacher lead project that aims to make 'teachers more research literate' - https://researched.org.uk/about/

Further details are provided below.

Project set-up, commissioning and contracting

Project set-up, commissioning and contracting issues appeared to be influential factors in driving a recruitment strategy that, while successful in achieving the intended number of participants, did not initially focus sufficiently on the TLIF target schools (See Appendix C). **A perceived barrier** in TBT progress towards realising KPI targets, especially in the early stages, was the perception of TBT Project Leads that there had been some 'initial contractual ambiguity' surrounding what constituted a target school. TBT concluded contract negotiations later than other cohort 1 TLIF providers but were very aware that their year one target for school recruitment remained non-negotiable. TBT were particularly motivated to meet this deadline and avoid possible financial consequences. This set of commissioning circumstances seemed to drive a TBT recruitment strategy that was focused more on securing participant and school numbers than necessarily ensuring all eligibility criteria were met.

However, following further discussions with DfE in 2018/19, TBT were given explicit confirmation that strictly only schools with a category 3 or 4 Ofsted would count towards KPI targets. TBT responded by altering its recruitment approach to '*focus* [solely] *on target schools in target* areas' thereafter. This tightening of focus was perceived as a crucial enabling factor by TBT project leads for increasing the proportion of participants trained from target schools by the close of the project However, as stated previously, according to DfE management information data, ultimately the number of participants recruited to the project from target schools (59% of those recruited) remained under the KPI target (70%).

As noted in Section 2.2, despite significant interest from schools, initially the '*vast majority*' were not eligible. Substantial time was taken to sift through applications and check eligibility, which often meant that by the time TBT could make a formal offer, the window of opportunity had been missed, which acted as **an additional barrier** to recruitment: *'It took us weeks to get back to some schools, by which point they'd say, 'oh, we can't come there for January' or whatever'* (TBT Project Lead). However, updating their *'comms'* facilitated greater precision in targeting eligible schools and acted as an enabler to recruiting KPI schools.

Utilising key strategic contacts

The use of Tom Bennett's school system leader network contacts acted as both an enabler and a barrier to recruitment. Initially, it acted as an asset towards fast recruitment but, in some instances, this led to participants not being fully informed about the commitment required. For instance, a number of MAT leads were contacted and relatively straightforwardly were able to identify significant numbers of senior leaders to attend the Running a School course. However, particularly in the

early phase of project delivery, this led to a more devolved relationship with participants and a greater reliance on 'intermediaries' (e.g., MAT leads) to set up the event and communicate with participants. TBT recognised these early teething problems and resolved to '*get a lot more muscular about who's signing up*' and ensure that every participant was fully aware of the commitments involved over a six-month period. Later interviews with the TBT Project Leads indicated that implementing this revised approach reduced the number of enquiries from ineligible schools.

Supporting engagement in the project

An **enabler to recruitment was TBT's pragmatic and flexible approach** which enabled schools in challenging circumstances to participate. TBT Project Leads described a number of ways in which they attempted to be flexible and responsive to schools', or MATs', circumstances to support and enable engagement in the project. TBT Project Leads inferred that their preferred model was to recommend that schools send some committed leaders to the Running a School course first, and then to subsequently sign up all (or most) teachers to the Running a Room course. As Tom Bennett explained:

> …when you are doing CPD in schools which have existing [trouble], when the plane is already flying, you will always be involved to some extent in triage and battlefield medicine, which means that you work with the people you have in the circumstances you've got, to achieve as many of the outcomes as you're looking for… if the only way to get some of these ideas into a school is by having selected members of staff sent, with some level of seniority, then to my mind that's better than a school which feels like it can't participate because it cannot devote the resources, as many can't. Schools with high level of challenge, with high level of demand on their material and personnel resources, and particularly smaller schools find it very difficult just to shut everything. Special schools find it very difficult just to shut everything down. So, there are practical considerations which mean that the non-ideal model is still ideal for that circumstance.

Cost, geographic location and availability

A key enabler for school leaders was that there was no charge for the CPD. This made participation appealing, particularly when it was being led by Tom Bennett who was widely recognised and respected. In addition, the relatively light-touch nature of the project acted as a further enabler, as the staff release commitment was perceived as manageable. However, a barrier in certain instances was the

geographical location of courses, occasionally significant distances from interviewees' own school settings, something that was said by some to be problematic and that might have limited the number of leaders able to attend.

A theme that arose in a small number of instances was disappointment that school leaders had not been aware of the possibility of whole-school training, or that they were informed that there was no remaining availability:

> I think I probably would have signed up for the whole-school rather than just the leaders' [course], because what happened was, we did the leaders' and then, in our inset day in September, myself and the SENCO did an inset based on what we'd done, with the rest of the school staff. But clearly, it was always going to be better coming directly from Tom Bennett himself.
> - *Headteacher, Primary school*

Finally, an issue for a limited number of individual teachers was that Running a Room courses aimed at self-selecting teachers from across different schools had to be cancelled owing to a lack of demand.[9]

[9] In most instances, individual teachers were offered the opportunity to join a specific whole-school, school-based Running a Room course nearby

3. Delivery, and implementation of learning

Section 3 outlines the extent to which the TBT project was delivered and implemented in accordance with intentions described in Section 1, as well as identifying any adaptations for improvement.

3.1 Progress in delivery

Perceived overall effectiveness of TBT provision

The overwhelming majority of participants found the overall quality and effectiveness of the TBT provision to be very high, as is further evidenced in Sections 3.2 and 4. This aligned with the positive feedback TBT Project Leads reported receiving from their end-of-course feedback questions, that stated that 'well over' 90% of participants gave a score of over 4 on a 5-point scale[10] for course quality, meaning the satisfaction target was met.

The most consistently referenced beneficial dimension across all of TBT delivery was the sharing of a set of overarching principles and a language that could be utilised to affect sustainable change in whichever direction schools felt appropriate to their own context. Across all components delivered, there was a strong emphasis placed on the need to purposefully model expectations (whole-school and in the classroom) in an explicit and transparent manner, thus greatly reducing the risk of ambiguity and misunderstanding:

> He was really ethos-driven… these are the ways that you can do it. His presentation style is modelling all the time. So, it was just extending that toolkit, building confidence. - *Senior leader, Secondary school*

Participants frequently referenced engaging with and applying key project concepts/terminology such as '*relentless routines*', being a '*conscious architect*' and '*what you permit you promote*' into their own practices. This focus on transparency and consistency was referred to on one level as not being '*rocket science*', but at the same time being an appropriate and extremely effective approach:

> …it's an increased awareness and use of the language as a result of us saying this is the way we do things here…Now we're deliberate architects of it and, therefore, need all staff buy-in. - *Senior leader, Secondary school*

[10] Question and scale provided by DfE.

However, it needs to be acknowledged that, in a small minority of case-study schools (two secondary target schools), programme delivery was not well received and was perceived to be overly reliant on personal anecdotes and lacking in the necessary substance and specificity to help address either entrenched and/or severe behaviour management difficulties. These senior leaders expected some more concrete strategies to help deal with more extreme instances of poor behaviour they were faced with:

> They were general [strategies]. Do you know what I mean? Actually, OK, when they are really [hard-hitting], what do I do with them? Because I'm not meant to permanently, exclude them. So what else can you do?... it gave you a great oversight, but then there wasn't the sort of nitty gritty answers that I was hoping I would get. - *Senior leader, Secondary school*

Effectiveness of delivery staff

The programme was delivered mainly in accordance with the expectations outlined in Section 1. **Tom Bennett led** delivery on **all** of the two-day workshops and the majority of the Boosters.

Interviewee feedback was particularly positive about the manner in which Tom Bennett delivered the course, with comments frequently praising his 'engaging' style and ability to interweave relevant personal anecdotes and refer to effective school practices in relation to behaviour management:

> …his skill as an orator, as a presenter, his charisma, his jokes, everybody was sort of a little bit in love with him. And that's actually really important, because he's ultimately telling you you're not doing it quite right, but he's doing it in a way that makes you feel like we're all in this together. - *Senior leader, Secondary school*

TBT delivery was always supported by at least one additional consultant, typically sourced locally to the school the course was taking place at. Consultants tended to either be in a current school leadership role (e.g., Executive headteacher) or had classroom-specific expertise. This aspect was perceived to add to the '*credibility to the course delivery*' and provided '*concrete examples*' from practice. However, a number of respondents commented that other supporting CPD deliverers whilst perfectly competent, did not have the equivalent delivery presence that Tom Bennett himself was able to bring.

Given the reliance the programme placed on Tom Bennett delivering training, there is a question about whether course delivery would have remained viable in the event

that he became unavailable. This poses some wider questions about the extent to which the TBT programme could realistically be scaled up in its current form.

Delivery and effectiveness of specific course components

Initial two-day course

The vast majority of interviewees stated that this was the single most useful component of the project. Interviewee responses, along with observation data of the Running a School training, confirm that the initial two-day course was delivered through a '*lecture-style/workshop experience*' where the core behaviour management principles and overarching change management techniques became established.

This largely followed the intended structure outlined in Section 1, with the two-day course delivered sequentially. However, it was reported that TBT Project Leads were amenable to accommodating school requests to deliver this component less intensively, for example over three days or four half days, where it was necessary to do so. All interviewees referenced having attended these initial two-day courses in one structure or another.

Interviewees universally appreciated opportunities to connect with and learn from colleagues from other schools or their own, depending on which course they accessed. These opportunities continued across the delivery, but were particularly prevalent in the two-day course because it lent itself to the exchange of ideas to inform plans and there was more time specifically reserved for this:

> …establishing relationships with those members of other schools was a lot more beneficial to me in cementing ideas for my school than I perceived it would be before I went. - *Senior leader, Secondary school*

> TBT's training was not so much [deliverer-receiver], it was an environment created for debating, discussing, interacting together, because it was over a longer period. And that hugely benefited me. - *Senior leader, Secondary school*

In one example, a senior leader from a secondary school was connected with schools from a neighbouring authority and they continued to network outside of the TBT project.

Booster days

Interview data shows Booster days were valued by the vast majority of interviewees as they focused on consolidating and embedding the key

learning from the initial two-day workshop (as opposed to introducing new content). Even those that were initially a little sceptical about their purpose praised their timing and the opportunity they afforded to refocus. **Interviewees particularly welcomed having enough time to have trialled implementing their plans (or aspects of them) before returning to the Booster sessions to discuss their experiences with TBT Project Leads and peers** that could help troubleshoot or advise on ways through any obstacles encountered:

> I think the Booster days were the most useful bit…until you go back and try it, and then you hit the barriers that you hadn't envisaged, and then go back to Tom and being able to say, 'Well this is really good, but this bombed completely', then other colleagues being able to say 'Well actually we found that was really hard as well'. That was the bit that kind of gave me the momentum to carry on I think, because you can get discouraged quite quickly when you're met with resistance. So, the Booster days were the bits that I found most useful. - *Senior leader, Secondary school*

Although Booster days were completely focused on behaviour, what was covered and the format in which it was delivered varied. The extent to which TBT Project Leads had capacity to tailor the Booster days to each school's specific preferences was contingent on the particular model of delivery. Where the Running a School course was delivered with a '*mixed group from different schools, then pretty much the only way you could make that work on a Booster day would be to reconvene the exact same group in a similar circumstance and make it a reflection coaching model*'. Contrastingly, for Running a Room courses and occasionally Running a School courses where the whole Senior Leadership Team (SLT) was derived from just one school (or MAT), there was greater scope for the Booster days to become '*more tailored and boutique*' to suit individual school's preferences and to be delivered *'in situ*'. However, it was acknowledged by TBT Project Leads that accommodating Booster days for Running a Room delivery was 'more complex' because of schools' understandable reluctance to shut a school for four days.

Below, we outline some of the different delivery approaches that were applied in relation to Booster days for whole-school models:

- further workshops
- working with school data managers in relation to tracking pupil behaviour data and facilitating a wider discussion about how the whole school staff had been engaged

- observing lessons, assemblies, and other areas of school life in situ [through the lens of particular policies or features that schools had committed to implementing] to inform subsequent activities and discussions
- focus groups with a selection of teachers (different criteria were applied, but which included: new staff members; teachers perceived to be struggling; staff who had 'particularly run with the programme'; or typical teachers) who TBT delivery staff would request feedback from to inform understanding of how the project was being implemented (or not)
- lunchtime drop-in clinics for anyone who wanted to discuss any individual behaviour-related issue they were encountering
- afternoon or lengthy twilight sessions with the whole school staff.

Online portal and community

Use of the online portal, including the community, was consistently described by interviewees as the aspect of the project they engaged with least, if at all. A significant proportion of participants, particularly those recruited early, referred to a number of technical issues with the portal including difficulty gaining access to the online resources and community and, in one instance, failure to receive certification following completion of online units:

> No, because my login never worked. Every time I contacted them they said 'Yes, we fixed it' and then every time I got to the next meeting, I'd say 'It doesn't work' and someone would write down my email address again and it never worked again. So, I've never even seen the beautiful portal. - *Senior leader, Primary school*

> Terrible. The actual ones [resources] that worked – brilliant. But the website is absolutely shocking. There are so many errors on it. There are resources that don't load, and yet there are questions on there that you have to answer. - *Senior leader, Secondary school*

Where interviewees did access the portal, they tended to state that the website and materials were of high quality, but often cited a lack of time to fully do justice to what appeared to be available:

> The online stuff looks great, but I didn't really do a huge amount with it, but like I said that's not reflective of the quality of it. - *Senior leader, Primary school*

There was very limited evidence that the online communities, 'drop-in clinics', or the analytic technology (where TBT leads monitored the online portal usage and interacted with participants in a more informed manner) referenced in Section 1, had been meaningfully engaged with at any stage throughout the evaluation. As a TBT Project Lead recognised *'… to be honest I don't think the online platform took off a great deal'.*

Tour of the school hosting delivery

Providing a tour of the host school was a later addition to the initial two-day programme for a number of Running a School courses. TBT Project Leads referenced that host schools tended to be strategically selected to include a senior leader able to provide current practice insight into effective behaviour management systems. For the majority, the tour offered a valuable additional dimension. One interviewee praised how the host school did not attempt to 'sugar-coat'. However, the tour was not a guaranteed component and, whilst the majority of interviewees were positive about the practice they observed, a minority felt that there was a disconnection between what was being taught during the formal sessions and what they observed informally at the school (i.e., behaviour not being dealt with effectively).

3.2 Progress in the implementation of learning

TBT provided training and resources specific to each model of delivery (see Sections 1 and 3.1) but viewed the implementation of learning to be the responsibility of participating schools and individuals. As noted in Section 3.1, TBT was not about the implementation of a 'one size fits all' behaviour management policy or set of practices. Instead, **participants were trusted to self-reflect on the behaviour and efficacy of behaviour management systems within their own settings** (school or class level) in order to make **informed decisions on which learning to implement:**

> …it was constant discussions and tweaks all the way along. So, there wasn't a 'This is how we're doing it and you have to…'
> *- SENCO, Primary school*

The decision-making of the participants involved was influenced by an array of different contextual and individual considerations including:

- baseline levels of whole-school or class-level behaviour issues
- effectiveness of existing school and class-level behaviour management systems
- how long the SLT had been in post

- extent of other competing whole-school priorities
- receptivity to core TBT CPD messages and style of delivery.

These factors influenced both the focus of in-school or in-class implementation and the extent of that implementation. Caution should therefore be exercised about making simplistic associations between participant-level implementation and course delivery effectiveness.

How was learning implemented?

The course delivery model that participating schools (or individuals) signed up to influenced the mechanisms through which learning was translated and ultimately implemented.

The **Running a School course** typically required senior leaders involved to make judgements as to what learning and strategies they wanted to cascade or translate to wider staff groupings and ultimately pupils and parents. In instances where the headteacher elected to not directly attend the training, some interviewees described occasions where a more inexperienced senior leader (without the authority to implement change) would schedule a de-briefing or short consultation session very shortly after attending TBT inputs in order to arrive at next steps. However, in most instances, the headteacher or a senior leader with autonomy to implement change directly attended TBT inputs.

The **Running a Room course** required less direct translating from senior leaders because teaching staff (and in most cases all school staff) received direct inputs from TBT Project Leads. TBT Project Leads and senior leaders referred to communication in advance of the whole-school pitched course to ensure that the key messages were as relevant as possible to the school's circumstances.

Below we highlight how senior leader interviewees reported implementing learning derived from attending the TBT course, with some attempt to share typical chronological ordering (please note schools/individuals did not necessarily take part in all stages).

Reflection and consultation stage

- Formal staff consultation periods were run often involving the whole staff (including, for example, Teaching Assistants (TA), midday supervisors and the office team). Opinions were canvassed through a variety of mechanisms such as surveys, baseline audits and more informal opportunities for conversations as part of existing CPD structures.
- Schools sought out the 'student voice'. Examples of strategies included focus groups, whole-school surveys and consultation with student councils.

Launch of revised strategies and policies

- Changes to behaviour policies and expectations were communicated, for example, as part of dedicated twilights or inset days, through discrete training days and departmental meetings (particularly in large secondary contexts) to the wider whole-school staff (for example TAs, pastoral support workers and lunchtime supervisors were also regularly invited to attend).
- Key behaviour policies (and associated reward systems) were updated on websites and/or hard copies reprinted.
- Awareness-raising of key changes to behaviour policies was undertaken and expectations were communicated to pupils through whole-school and house assemblies, posters around school corridors and in classrooms.
- Greater emphasis was placed on explicit modelling of expectations. For example, students were directly shown how to enter the classroom or were physically walked through where/how to stand in queues for assemblies or when returning from break time.
- Modifications were made to existing job descriptions and adverts for new staff, taking account of changes in expectations.

Maintaining awareness and implementation

- Behaviour-focused learning walks and greater peer observation was undertaken, along with any revised behaviour management expectations being incorporated into staff appraisals and other performance management structures.
- The message was maintained through continued inputs into existing CPD structures and occasional one-off dedicated training sessions.
- Compliance with revised behaviour policies was incentivised through implementing positive praise and seeing through promised reward activities, for example being entered into a draw for cinema tickets.

What learning was implemented within schools?

In the following section, we thematically outline the key learning and practices that became implemented at schools.

A clearer understanding of behavioural expectations: large numbers of interviewees suggested this began with senior leaders simplifying whole-school behavioural policies that made the fundamental expectations visible and clear. This improved leadership clarity was perceived to have fed down to teachers' classroom practices, with many school leaders and classroom practitioners reporting that there was more explicit modelling of behavioural expectations provided to pupils within the

classroom. This took a variety of forms including showing PowerPoint slides at the start of a lesson, having posters around the classroom and the teacher verbally reinforcing them. It is interesting to note that even teachers who did not directly participate in the TBT project used phrases often used by TBT Project Leads such as being 'deliberate architects', 'identifying the non-negotiables', 'having consistently high expectations' and 'establishing relentless routines'.

Senior leaders were also reported by non-leadership staff as being far more 'hands-on' and visible around the school. For example, being more likely to take on a break duty, walking around the corridors more and increasing the number of classroom drop-ins. This could be interpreted as sending a tacit message about the greater importance they were placing upon adherence to behavioural expectations.

Greater whole-school adherence to behaviour policies: Many interviewees told us that a culture was beginning to take hold in their schools whereby decisions about whether to enforce behavioural policies (or an aspect of them) were no longer perceived as discretionary. One senior leader spoke about there now being a 'united front' across staff. There was significant evidence from both teaching and non-teaching staff that there was a more systematic enforcing of behavioural expectations. Examples included no longer tolerating low-level disruption within classes, confiscating mobile telephones, and ensuring the school uniform was worn correctly:

> … the classic one is the shirts tucked in…the onus is on us to challenge…If you look around the school, I'd say 97% of them [pupils] will probably have their shirts tucked in… No one goes by unchallenged, and you can hear it down the corridors – 'Get your shirts in'. Then another member of staff – 'Get your shirts in!'…Sometimes you don't even have to say anything, you just go [gestures] and they know what you're going to say – because that's how much it's embedded. - *Pastoral Worker and TA, Secondary school*

More analytic and systematic approaches: Senior leader interviewees regularly conveyed a need to constantly review and avoid complacency: '*even when it's* [behaviour's] *good it's still not good enough*'. One way they put this philosophy into practice was through improving the tracking and monitoring systems for logging behavioural incidents. This made it easier for school staff to analyse the data and discern any trends or patterns that could assist with implementing an appropriate response:

> We log incidents of disruptive behaviour, so we can then track which children these are and deal with those children... We can

> also add an action to that and tag in other staff… they know that we've seen it and also that there's an action for them to do, so it's very proactive. - *Senior leader, Secondary school*

Senior leader interviewees reported this made them more focused on **pre-empting issues**, being proactive and **identifying trouble spots** in order to avoid needing to be overly 'reactive.' Several interviewees also claimed to have closed off 'loopholes' pupils had previously been exploiting, by making adaptations to their consequence systems to make them more of a deterrent. In one example, this meant extending the school day and amending school bus timetables, so they arrived later to allow for same-day detentions to be given. Senior leaders at the school in question argued that moving towards same-day consequences was more meaningful (see vignette of School C below).

Establishment of more positive learning environments: Alongside a drive for the consistent enforcement of school consequence systems, there was a conscious effort by school leaders to promote more nurturing, positive and reward-led learning environments. One senior leader sought to implement *'dignity and learning'*. There were numerous examples of senior leaders and wider staff creating or building upon existing reward systems and/or house competitions/club activities run by staff for pupils, and end of term prizes (for example cinema tickets) for pupils who had displayed consistently good behaviour. This operated alongside a wider focus on respectful communication. The following quotation illustrates how one school set about implementing a programme of lunchtime and break time activities[11] run by school staff:

> The students get to see the staff taking part. So even in the bottle flipping, on one round we had the headteacher competing against other senior members of staff, and obviously with things like Deal or No Deal, you're not getting friendship groups, you are getting different individuals, different ages, different ethnic groups. So, it's broken down that divide between different ethnic groups. With that, when you have limbo, obviously you'll get 200 kids participating and you'll get another 200 actually sitting and watching it, so obviously it's giving them something to do at lunchtime and it's going to improve their behaviour in social time. They're not gathering outside in groups; they're not running riot on a field… they've been really popular. - *Middle leader, Secondary school*

[11] The TBT project did not recommend specific activities

A further practice implemented across various schools was a conscious effort to create more positive communication channels between teachers and parents. For example, teachers making positive phone calls to parents or sending home positive postcards in instances where pupils (particularly those with a history of poor behaviour) had behaved well. Collectively, such initiatives were often credited with improving teacher-pupil relationships. Attempts to improve relationships and to dovetail support, reward and behaviour systems were recognised and appreciated by pupils:

> However, if you're constantly misbehaving then you do miss out on opportunities. We had a rewards thing in Year 8 that was like an Alice in Wonderland themed tea party at the house... Christmas discos…which if you're misbehaving, you'd miss out on the things like that. - *Pupil, Secondary school*

> People get upset in the school and they might get angry and all that. That's the reason why we have things like gateway and [self-study rooms] and blue room. We have many counsellors who come in. And we have a school counsellor. And that's the reason why our school is so structured. It's not like out of control or anything. Because we have a lot of support – like tutors, teachers, and counsellors. - *Pupil, Secondary school*

Personnel and infrastructural changes: A small number of interviewees stated that their school had renovated and/or built new areas connected to behavioural management systems. For example, one secondary school launched a brand-new restorative room which provided a larger, more learning-friendly environment with a full set of computers, whilst one primary school launched a dedicated nurture/reflection area with purposely calming colours. This was designed to facilitate restorative conversations either between pupils who had fallen out, or between a pupil and a member of staff. In addition, there were some isolated examples of schools investing in additional staff to more directly assist with the restorative dimensions of the behaviour policy, for example a trust-wide counsellor and additional pastoral staff.

A greater toolkit of techniques for teachers to apply within the classroom: While the focus of the TBT project (especially the Running a Room course) was on wider strategies and reflection on current practices, the course did offer some specific techniques to illustrate their points. For example, interviewees attending the Running a Room course reported that they had developed particular pedagogical skills, de-escalation techniques and practical resources to assist with managing their classroom more effectively. This dimension appeared to be particularly important for

NQTs and early career teachers. However, in most instances, it was at the discretion of individual teachers as to whether they used them in their own practice or not:

> …invisible ladder stuff, so making sure that we are putting those rewards [in place] allowing a successful atmosphere basically. So, we're setting the children up to be successful rather than this is really strict…because it's not creating an environment where you want to come in and learn. So, I think that was quite good, highlighting those kinds of positive and subtle things that you can do in and around the classroom, like tapping tables or gently prompting students to kind of get on with the work or whatever it might be… much more preventative rather than curative or looking into the school system…. - *NQT, Secondary school*

Other examples included:

- use of **flow diagrams** to make consequence systems clearer:

  > We had new staff joining as well; we did the flow diagrams, so they were clear about the steps that they need to take if things weren't going… [referring to papers] … so they knew… OK, so this is happening, this is what I need to do, this is making it really explicit to them. - *Senior leader, Primary school*

- **Seating plans** to try to manage the classroom environment in a manner that optimised learning opportunities and minimised misbehaviour:

  > Yesterday I had a Year 10 group and they'd been absolutely fantastic up until when I've had two students come in from managed moves. Because of those two students…the dynamics have changed completely. Yesterday it was absolutely horrendous… [led to] a complete seating plan change, and the behaviour today was just spot on. It had gone from probably the worst lesson I've ever had to the best lesson I've ever had. - *Teacher, Secondary school*

- **Use of scripts**[12] (sometimes differentiated for certain groups, for example pupils with special educational needs and disability (SEND) to assist with implementing the behaviour management system and disseminating a shared comprehension of expectations to facilitate restorative actions:

[12] Certain interviewees referenced using scripts with key phrases and a set of language to assist with agreed behaviour management processes and to improve the consistency of response. Scripts were bespoke to different schools and sometimes particular pupils.

> Scripts…we've introduced this year... there are different children with different scripts for different needs, the wording is very similar and age-appropriate, so that it uses the same four key words, simple language, easy to understand, easy to hear. I can see you're upset, tell me and I'll listen. So, the idea of the use of these scripts is certainly something that the [headteacher] has filtered through. - *SENCO, Primary school*

Categorisations of implementation

Below, we have categorised case-study schools according to the extent to which learning from TBT became implemented into their practices. There are three categorisations, and, for each, we provide an illustrative vignette based on data from one of the school case-study visits.

- **Learning implemented to a certain extent:** This category comfortably accounted for the largest proportion of schools and participants. Learning implemented was mainly aligned with a set of practices and existing trajectory relating to whole-school behaviour management. Implementation in this category tended to be limited to 'tweaks' to existing policies, but with a deeper focus on achieving greater, whole-school levels of consistency which often enabled accelerated progress. Such schools were usually in more 'secure' circumstances from a school performance perspective, with an established SLT and pupils with fewer perceived behavioural issues (especially more serious examples).

School A - Non-target secondary school (Running a School)

This school was rated 'Good' in terms of Ofsted, with a mixed profile of pupils, a well-established SLT and baseline behaviour perceived to be largely good. It had pockets of low-level disruption and a small core of pupils (often with underlying SEND issues) involved with more serious behavioural issues. These characteristics were broadly typical of the schools that took part in the evaluation. School A's motivation for involvement in TBT, like many other schools, was to take their school to the 'next level' which, in their particular case, meant more targeted work on pupils who persistently infringed the behaviour policy and to make application of the behaviour policy even more consistent across non-teaching staff:

> I think behaviour has always been good at the school… [But] how we're dealing with that top 10%, those frequent offenders, those [non-compliers]. How do we identify students, what methods are we putting in place to ensure that there aren't repeat offences there, how are we helping rebuild and reform those students to help them? - *Senior leader*

School A had a dynamic Senior Leader, especially committed to the implementation of an effective whole-school behaviour system, which had also been the focus of their National Professional Qualification for Senior Leadership (NPQSL). Steady incremental gains had been made over time and the behaviour policy was purposely tied up with an existing school philosophy:

> The [School A] Way – we have one school rule, which is respect. Everything else is built around that one school rule. We have certain aspects that staff are expected or encouraged to follow... it's the thing that everything else hangs on. - *Middle leader*

During the course of the TBT project, the school's SLT attempted to implement the following things:

- dissemination of key learning from the TBT training filtered through: in-school CPD training; significant inset training; and CPD PowerPoints to teachers and wider staff including support staff, site managers and attendance officers
- use of restorative languages/practices, particularly across the behavioural and pastoral teams.

Increased ways of recognising good behaviour were also implemented, for example through rewards of hot beverages, postcards and phone calls home to parents.

- **Learning implemented to a significant extent:** Although sharing a number of similar traits to the previous categorisation, these schools tended to be on a less secure footing from a school performance perspective, with less established SLTs and greater perceived pupil misbehaviour. In addition, these schools tended to have less established and/or effective whole-school behaviour policies. Therefore, there was often greater need to implement CPD learning to a more significant extent. For example, a more radical overhaul of behaviour policies (e.g., explicitly addressing a contentious or recurrent behaviour issue, such as the banning of mobile telephones, through to consciously streamlining overly complex policies). Given the stage of the journey many schools were on, and their often difficult recent history, there was perceived to be a greater need to consider ways of nurturing pupils and rebuilding staff confidence.

School B - Target school (Running a Room)

A converter academy, with a history of underperformance and recently overhauled SLT. On arrival, the new headteacher described the school as being 'in crisis', with '*180 kids in detention every day'.* Teaching staff identified behaviour management as their most pressing concern and as something that consistently impeded teaching quality:

> Teachers couldn’t teach – I was trying to teach about pedagogy and curriculum change and assessment, but they were just trying to survive…they couldn't even think about teaching and learning. - *Senior leader*

The new SLT pursued an unorthodox approach to behaviour management, signalling a conscious move away from 'saturation of top-down strategies’, perceived to have created a culture of 'learned helplessness' and staff feeling 'powerless'. The SLT instead aimed to move towards *'subject teams'* working together to arrive at their own approaches. '*Holding their nerve*', the SLT purposely refused to create a whole-school behaviour policy. This was the stage School B was at when it signed up to TBT.

Certain senior leaders held reservations that TBT Project Leads might not 'grasp' what they were attempting to implement, something that risked 'undermining' progress made. Reassuringly, TBT Project Leads emphasised they were not there to 'judge' or be 'critical' of school approaches, emphasising any policy *'has got to be right for your school'.* The overarching strategy School B resolved to implement, was to '*set their stall out*', '*revisit expectations*', '*be consistent*' and '*to keep going*'.

The fact that **all staff experienced the message directly from TBT Project Leads was seen as an enabler for universal staff implementation**. It meant there was no departmental variation in how the message was cascaded throughout the school. However, the onus remained on 'empowering' individuals to use their own professional judgement in determining exactly how they implemented a set of whole-school principles (not rules) within their own classrooms and across the wider school environment:

> The whole school [project] was absolutely right, absolutely right for everyone to get that same message…it’s a great thing to come back on… No one can say that they haven’t had training. No one can say they haven’t had the support. No one can say they don’t know what we’re doing, because we’ve all sat in that room and we were all in there. - *Senior leader*

This approach was said to be further aided by deliberately heightened senior leader presence throughout the school. '*It’s SLT being on the corridors backing staff up... being present at break and lunch time... actually getting hands dirty'.* Finally, there was pragmatism from the SLT that what had been implemented was not yet the ideal, but what befitted the stage pupils and staff were currently at, and fast tracked the school's progress:

> We had the policy and processes, but they’re only as good as the bit of paper they’re written on. They’ve got to be implemented by the staff. And that was the real message we wanted... we’re three, four months ahead of where I thought we’d be. - *Senior leader*

Section 4.3.1 includes a brief vignette on School B outcomes.

- **Learning not implemented**: For a very small minority, project delivery was not well received, to the extent that it was claimed that no learning was implemented. Given the very small numbers involved in this category, caution needs to be applied to making generalisations. However, in the limited number of cases there were, there was a fundamental misalignment between the style and content of the CPD delivered by TBT and what participants were a perception that TBT was overly reliant on personal anecdotes, lacking the necessary specificity to help address entrenched and/or severe behaviour management difficulties. These schools were exclusively from target schools.

School C - Target secondary School (Running a School)

This school had had a significant turnover of senior leaders and was categorised as requiring improvement and had a long history of 'underachievement'. The community served was characterised as socio-economically disadvantaged and there were examples of serious criminal activity that had involved pupils and/or their relatives.

However, this was a school being reinvented, driven by a new and dynamic SLT. By their own admission, both the headteacher and deputy tended to lead by their 'gut'. Despite not being fundamentally opposed to the strategic messages of TBT (in many ways the changes implemented over the previous few years appeared largely aligned to the core principles), there was frustration about the lack of practical advice to accompany it:

> I'll go back to it, that nitty gritty, when your back is against the wall, how do I make it better? - *Senior leader*

The SLT also expressed disappointment that TBT Project Leads did not 'commit' to supporting their proposed reforms. Nevertheless, the deputy headteacher backed by a 'strong headteacher' was given permission to implement the following two key changes:

- Removing a second warning step in the behaviour consequence system, meaning four steps became three:

  > Students were saying they were getting too many warnings. That's when we dropped it down to C1 your first warning, C2 detention, C3 lesson removal. - *Senior leader*

- Combining the 'remove room' (previously where pupils were sent if removed from a lesson until the next lesson began) with the 'internal exclusion room' - IER (where a pupil was sent for something more serious either from the same day or from the day before). The expectation being that pupils enter the IER ready to learn and reflect and, if by the end of the day they were deemed to have not done so, they would repeat a second day in IER (this included break times and staying after school).

The TBT Project Leads' chief concern was that, despite sounding good in 'principle', staff managing the IER room would become overwhelmed and that it would be impossible to maintain a calm environment conducive to effective working. The deputy headteacher disputed this and set about implementing their vision through making the pastoral team available (on a rota). When a student was brought to the IER but refused to go in, the pastoral team would speak with the student; if that did not work, their parents would be called, and if that still did not work, they would go to SLT. See section 4.3.1 for a brief outline of School C's outcomes.

3.3 Challenges and enablers in effective delivery and implementation of learning

Below we outline key enablers and barriers under a series of themes that are relevant to the TBT project.

3.3.1 Factors related to the provider/provision

Scheduling and timing of training: One consequence of the tightened window for delivery in the early phase of the project was that scheduling of course sessions appeared to lack a broader awareness of significant events in the educational calendar. This was described as frustrating and acted **as a barrier** in certain instances:

> ...the dates [of the project] were rubbish because they clashed with things like SATs [Standard Assessment Tests] work, phonics weeks, stuff like that. So, we ended up having to swap a group, from memory, and I just think there needs to be a little bit more foresight. - *Senior leader, Primary school*

The consensus among interviewees was that the optimal time to take part in the TBT project was well in advance of the summer term, to allow sufficient time to incorporate revisions to school behavioural systems and planning for the forthcoming academic year. Where this was not possible, one response was to delay whole-school roll out of revised policies and expectations and to pilot with certain classes or year groups first.

Pragmatic, flexible and non-judgemental approach to provision: This was said to be **a significant enabling factor** in allowing schools (particularly target schools) to implement learning tailored to their current circumstances. In the example below, Tom Bennett outlines instances where middle leaders were struggling to implement key principles derived from the training, due to senior leaders' unwillingness to fully commit to the necessary reforms needed to truly improve behaviour systems:

> I mean some of them had been sent by school leaders to say, you know, ‘fix’ this person and, as I say frequently, it’s the school systems themselves which are problematic. And frequently people say things like ‘I’m doing everything that you said Tom, I’m doing everything that we’ve been trained to do on this course, and I’m referring things to people they should be referred to’, but frequently they’re not very good at it, so we would have to do a lot more work with them, and workshopping with them, on how to manage upwards, how to manage a system [around you], how to liaise with effective colleagues and so on which was interesting, because this is more of [a case of] coping mechanisms in imperfect circumstances. - *TBT Project Lead*

Strategic-level pitch: The majority of leader participants appreciated the strategic-level pitch, which focused on delivering a contextually appropriate school vision for behaviour management, as opposed to recommending adherence to a pre-specified blueprint. However, it should be acknowledged that **most schools taking part in the evaluation felt their pupils' behaviour was not a serious concern, with largely effective behaviour management strategies already in place, which was an additional enabler**. Typically, participants described wanting to '*enhance*', '*tweak*' or '*make more consistent*' their behaviour management systems as opposed to feeling the need to overhaul them:

> I think it just reinforced what we were already thinking, but in a way in which we could clarify where we wanted to go, with that consistency and the culture side of things. - *Senior leader, Primary school*

Nevertheless, the interviewees from a minority of schools were dismayed about the lack of project specificity, finding its generality underwhelming and a **barrier to implementation, particularly in contexts with severe behavioural difficulties**.

Course delivery model: **Participants welcomed being able to exercise judgement in deciding which course their school signed up to**. There were dimensions associated with both the whole school (Running a Room) model of delivery, and delivery to leaders only (Running a School), that were perceived to have enabled or impeded delivery and implementation effectiveness.

A number of senior leader interviewees welcomed the option of CPD focused on whole-school leadership strategies. They described deliberately not wanting their school to take up the Running a Room course, in order to have time to strategically consider their approach and to manage the messages being delivered to their teaching staff. **In contrast, many argued whole-school delivery**

approaches were more effective because they engaged more staff across the school and were focused on one school, something that inevitably leant itself more towards greater contextual specificity. As noted in the School B vignette (Section 3.2), whole-school delivery was favoured because it meant that all staff received the same unadulterated messages. This was something interviewees felt unified staff, encouraged joint accountability, and ultimately acted as an enabler to the implementation of learning:

> The whole school [project] was absolutely right, absolutely right for everyone to get that same message…It's a great thing to come back on. You've got that and it's: 'Come on, we've done this, you were in the training with Tom Bennett, we've all agreed this'. No one can say that they haven't had training. No one can say they haven't had the support. No one can say they don't know what we're doing, because we've all sat in that room, and we were all in there. It wasn't like, 'Oh we don't need to be in there, we know what we're talking about as SLT, we're good at behaviour'. Everyone was in there. - *Senior leader, Secondary school*

However, a **barrier** for whole-school approaches (involving all or most teaching staff) was that, in certain instances, staff attending were either not fully invested in, or lacked awareness of, course expectations. In contrast, the Running a School course tended to be populated with fully committed senior leaders who had actively sought to be involved. This nuance in different delivery models is outlined in the quotation below:

> So, the people who have signed up to our individual courses, big, big, buy in, really, really big buy in. One of the things we're going to have to really nail down is to make sure that, attitudinally, people are signing up to whole-school courses... Obviously, we can't reach into the schools so thoroughly, so the school have to make it happen itself. - *TBT Project Lead*

In a separate example, TBT Project Leads discussed how certain schools requested that support staff such as TAs and learning mentors be allowed to attend the Running a Room course. Subject to DfE approval, TBT were happy to include support staff and there were arguments for doing so, because of the whole-school nature of pupil behaviour. However, project delivery and pitch needed to be adjusted to accommodate both groups, and TBT Project Leads acknowledged that this did have a slight adverse effect on the pace and flow of training designed to be delivered to just teachers.

Finally, owing to the lack of numbers to make course delivery viable, individual teachers who applied to the Running a Room course were sometimes incorporated into another school's whole-school project delivery. This limited the extent to which they were exposed to colleagues from different contexts (as unlike the original intention, the whole-school project delivery they were incorporated into was focused on participants from just one school), potentially inhibiting the effectiveness of delivery. In addition, in one example, an NQT working in the primary sector was connected into a secondary school project. Although this was perceived by the NQT to be useful in terms of understanding the overarching principles, it did act as a barrier to the delivery of content with primary-level specificity.

Participants on the leadership course valued the dedicated blocks of time built into the project. During these blocks of time, participants were able to reflect and discuss (with TBT Project Leads, colleagues from their own schools and others) both intended and actual approaches to behaviour management in their own schools. This was felt to have **enabled** effective embedding delivery of the learning when they returned to school. Having the time and space across the four days of training to undertake this kind of strategic thinking was reportedly something very distinctive, when compared to other forms of CPD participants had experienced in the past:

> Tom said, 'You go back, you've got [101 things waiting for you the next day] and you won't do it'. Exactly right. So, one definite thing I needed was the time on the day to get my thoughts onto paper, otherwise it would have been lost. - *Senior leader, Secondary school*

However, a barrier to effective delivery was that participants did not have enough time to engage with the website or online communities. This barrier, as has been previously referenced, was compounded by technical issues.

3.3.2 Factors related to the school climate/context

Autonomy to implement change: In the vast majority of instances, senior leaders attending the Running a School course were self-selecting and incredibly committed to the project and implementing learning in practice. This was regarded as a key enabler:

> I didn't dream this up over a night... I spent six weeks – I spent from the January to the February half term, really looking at - where are we going wrong?… And this evolved probably three or four times until I had the final model. - *Senior leader, Secondary school*

However, in instances where a senior leader representative did not directly attend TBT sessions this was regarded as a significant barrier to implementing learning, although the worst effects of this could be mitigated if delegates were granted authority to implement changes and/or there were structures in place where plans could be discussed and actioned promptly, in conjunction with senior leaders with the authority to implement changes:

> …one very, very clear differential... was the seniority of the member of staff attending. When we had the mixed groups, there would typically be two extremes. There would be the CEO or the head teacher or principal who would be attending with perhaps their deputy. And that was one of the things that had the biggest impact, because they would then leave the programme bristling with good intentions and buying in, and then devolve it at the highest level throughout the school. That was effective…At the opposite end of effectiveness, we found less impact when, for instance, a junior member of the leadership team, perhaps even a pastoral leader, like a head of year or a head of learning had been sent along by a headteacher, who… had low levels of structural authority, and so would report back to the headteacher. And the headteacher would accept or not accept aspects of the programme depending on how much they bought into it, or indeed how well the returning member of staff could sell it to them. - *TBT Project Lead*

Perceived school stability and cultures of trust: In certain instances, KPI target schools had undergone a significant overhaul of senior leaders, teaching staff and in some cases school governance structures. Where there remained ongoing uncertainties about the leadership or status of a school, this acted as a barrier to the extent to which learning became implemented and/or would be sustained:

> There was a strong sense of nobody knew if the things from this programme would be implemented, because a new academy trust was taking over. So, there was a high level of uncertainty there…we've had as much impact as we could have…because of the uncertainty. - *TBT Project Lead*

TBT Project Leads claimed staff from target schools became especially committed when a headteacher signed up to attend the course, which acted as an additional enabler to school-level implementation:

> In schools which were suffering extreme behavioural challenges, we got high buy-in from the staff if they felt like the leadership

> was taking it seriously, which tended to be when leadership also attended the course.

In one instance, senior leaders reported feeling unable to sign up to the Running a School course because they had reservations about how effectively the individual responsible for behaviour would be able to filter down the message. They, therefore, opted to sign up to the Running a Room course instead, which although well-rated was not their preference.

Baseline pupil behaviour and alignment of existing behaviour management strategies: As has been explained elsewhere, baseline pupil behaviour was typically perceived to be largely good. Where behaviour issues were generally restricted to low-level disruption it was usually regarded as easier for schools to implement learning from the TBT training.

Also, where schools already had a well-embedded set of school-level principles, for example a one-school rule based around 'respect' or allowing everyone to 'shine', interviewees spoke about it being easier for the learning connected to TBT to 'hook' on to. However, senior leaders from a newly-built primary school with only three year groups currently on roll argued that their lack of ingrained systems enabled greater learning, because staff were not so set in their ways and were more receptive to revising their culture.

Competing priorities and time to dedicate towards behaviour: Four full days of staff's time (particularly for whole-school models) to attend behaviour specific professional development, even without the additional time to undertake gap tasks, amend policies and implement learning, was considered a significant investment of staff resource by some. This sometimes manifested itself as a barrier in two ways. Firstly, TBT Project Leads reported that not all staff attended all sessions, highlighting that Booster sessions, in particular, were often less consistently attended:

> Some people who had less affinity for the programme decided that something else they had to do was more important, and then trooped along for the last half an hour to show face.

Secondly, behaviour or project leads within schools sometimes outlined a reluctance to push too hard for additional time in the school calendar to dedicate to further whole-school, behaviour-focused CPD. This acted as a barrier to being able to get the learning as fully embedded into the school as leads would have liked:

> So, I've almost had my [limelight] and have to back off and let other things come in, as opposed to keeping this as the top priority of the school for a period of time. And that is just hugely

> frustrating for trying to get an initiative embedded. And the way I'm doing it, we have a staff briefing every Monday morning. And I do that drip-drip-drip. - *Senior leader, Secondary school*

3.3.3 Wider, external factors

Insufficient wider support services: The most extreme instances of bad behaviour were said to be instigated by relatively small numbers of pupils on a recurrent basis. A small number of respondents referenced wider systematic issues, such as a lack of access to alternative provision and/or specialist support services including Educational Psychologists and Counselling services, which hampered their ability to support such pupils as effectively as desired:

> Our children have significant challenges as any children that grow up in disadvantaged areas have. As well, there's a paucity of alternative provision so, if children have extreme needs, there's no alternative provision for them to go to, or there are very few places. If a child does get a place, they might have to wait up to seven or eight months for that place. So, it's quite a significant challenge. Our challenge was eliminating low-level disruption and then having to manage the behaviour of a real minority of children, under 10 I would say, but with extreme behaviour needs, where really they needed some kind of specialist provision not mainstream provision. - *Senior leader, Secondary school*

Difficulty with SENCO recruitment: In one instance, it was reported that a school consistently struggled to recruit an appropriately qualified SENCO. This lack of specialist expertise was believed to have, on occasion, impaired the school's capacity to implement their vision for behaviour management as effectively as they would have liked across all groups:

> We get on with it on a day-to-day [operating without a fully qualified SENCO] … you try and compensate and you deal with it and you still succeed. But you're aware… and if there's something we could do about it instantly, we would, but you just have to work with what you've got... Any key leadership position, you're going to be hamstrung by if that's [not being] done by someone who's doing it at a top level… - *Senior leader, Secondary school*

4. Perceived outcomes and impacts of the provision

This section considers the extent to which the TBT project achieved its intended project outcomes and impacts as well as the perceived contribution it made to the TLIF programme's intended outcomes and impacts (See Appendix A and Tables 1-4). It draws on qualitative data, exploring different stakeholders' perceptions of the outcomes of the project, and providing context for interpretation of these, and secondary analysis of SWC data to report changes in teacher retention and progression.

The analysis of impact on teacher retention and progression utilises a comparison group design. This enables us to estimate counterfactual outcomes for teachers, and infer whether or not changes in teacher retention and progression might have happened in the absence of the TBT project.

Please note that, **as the evaluation design does not include surveys, we are unable to provide a measure of the relationship between the project and any reported outcomes. The outcomes reported here are based on perceptions data and, therefore, should be regarded as illustrative rather than conclusive.**

4.1 Context for interpretation of outcomes and impacts

Although we have attempted to collect comparable fund-level outcome data for all TLIF projects, in practice the projects' intentions, with regard to achieving these outcomes, differed. The TBT project attempted to achieve all of the fund-level outcomes **within the context of behaviour management**. This should be borne in mind when interpreting the outcomes reported in Section 4.2 below.

A contextual issue to consider in interpreting findings on outcomes specific to the TBT project is that each school started the project at a different baseline stage of pupil behaviour and with a set of unique contextual circumstances (see Section 3.2). We have also previously cautioned that the proportion of participants from target schools involved in the evaluation was lower than the overall profile of target schools recruited (see Section 1.3). Furthermore, certain schools referenced drawing on other support to improve pupil behaviour alongside participating in the TBT project, most frequently consultant specialists in pupil behaviour management. Taking all these aspects together means that caution needs to be exercised in forming judgements about the outcomes and impacts of the project.

4.2 TLIF and bespoke project outcomes and impacts

Tables 1-4 set out the intended TBT project outcomes and impacts as agreed with TBT at the beginning of the project.

Table 1: TBT Project outcomes and impacts for senior leaders

Outcomes and Impacts	Outcome or Impact
Improved quality of senior leadership	
• Improved confidence in creating the conditions for effective behaviour management across their school	Outcome
• Changes in leadership practice related to effective behaviour management that lead to the implementation of a culture conducive to good behaviour and effective behaviour strategies implemented at school level, which are consistently deployed by all staff.	Outcome
Satisfaction and retention	
• Higher level of job satisfaction	Outcome
• Improved senior leader perceptions of school and their academic and civic identity	Outcome
• Improved leader retention	Impact
• Improved leader progression	Impact

Table 2: TBT Project outcomes and impacts for teachers

Outcomes and Impacts	Outcome or Impact
Improved quality of teaching	
• Improved confidence in ability to employ effective behaviour management strategies	Outcome
• Changes in classroom practice related to effective behaviour management	Outcome
• Effective teacher-pupil relationships	Outcome
• Improved quality of teaching (underpinned by improved behaviour strategies)	Outcome

Outcomes and Impacts	Outcome or Impact
Satisfaction and retention	
• Higher level of job satisfaction	Outcome
• Improved teacher perceptions of school and their academic and civic identity	Outcome
• Improved teacher retention	Impact
• Improved teacher progression	Impact

Table 3: TBT Project outcomes and impacts for schools

Outcomes and Impacts	Outcome or Impact
• Effective strategies used across the whole school to improve behaviour	Outcome
• Improved school culture (underpinned by improved behaviour strategies)	Outcome
• Valuing of CPD	Outcome
• Increased engagement in /demand for CPD	Outcome
• Improved teacher retention	Impact
• Improved teacher progression	Impact

Table 4: TBT Project outcomes and impacts for pupils

Outcomes and Impacts	Outcome or Impact
• Improved pupil wellbeing, behaviour and discipline (e.g. via perceptions of pupil attendance, exclusions data)	Outcome
• Improved pupil perceptions of school and their academic and civic identity	Outcome
• Increased pupil attainment	Impact

Below, we discuss and outline the extent to which involvement in the TBT project was perceived to influence participant and school-level realisation of these outcomes and impacts (Sections 4.2.1- 4.2.6). All intended TBT outcomes and impacts can be aligned directly with TLIF fund-level outcomes and impacts.

4.2.1 Findings related to improved leadership quality

A significant perceived outcome was greater confidence among senior leaders to lead whole-school behaviour management. Leaders often felt the underlying principles and processes learned throughout the TBT project (see Section 3) helped with their '*clarity of thought*' in relation to behaviour management. This prompted most senior leaders to convey they were more emboldened to fully commit to ensuring their vision became systematically implemented back at school:

> I felt more authoritative in my position at the same time. - *Senior leader, Secondary school*

There was also greater self-belief that revised styles of leadership and change management techniques could be effectively transferred to other aspects of leaders' roles:

> If I've got to do another big change, then yeah I'll have the skills to do it regardless of whether that's behaviour or one of my other roles... because I thought about everything first, and got the rationale, and I was ready and thought about all the objections and questions and prepared for that, so I think that really gave me confidence more than anything. - *Senior leader, Secondary school*

Participants believed that leadership practice had been enhanced. There was evidence that aspects of the TBT project had been taken on by senior leaders in order to positively affect the implementation of a culture conducive to good behaviour and effective behavioural strategies (see Section 3.2). Senior leader interviewees perceived that teachers benefited from their more explicit guidance on how they should respond if a pupil behaved outside of clearly-conveyed expectations:

> I think teachers will be… more secure that they understand what I want... Because it has prompted me to be more explicit, then I think they've got greater confidence…in knowing that they're doing, what is required of them. The danger is you can just say I want quiet lines or this that and the other, and I've been in schools where it's happened, and you almost guess what the expectations are. - *Senior leader, Primary school*

In a minority of instances, teachers expressed relief that involvement in the TBT project had acted as a catalyst for senior leaders to give pupil behaviour a greater share of their attention and to drive home the message that *'it's got to be the bread and butter of the school in the same way that maths or English or the individual*

subjects that you're teaching are' (Teacher, Secondary school). Particularly in instances where the school signed up for the Running a Room course, there was a sense that they had been given the 'green light' to allocate the time to address and '*explicitly talk about* [behaviour issues]*'.*

There was strong evidence that behaviour had become prioritised in participating schools and that behaviour policies had become more embedded across the whole school workforce, with staff practices becoming more consistent. Pupil focus groups tended to chime with staff views that whole-school behavioural expectations had been more transparently communicated, with most able to clearly communicate top-level behaviour policies. In certain instances, the cohesiveness of effective whole-school behaviour strategies even extended to more direct self-policing from pupils. This was indicative of particularly ingrained cultures, as is outlined in the extracts below:

> I have students that come to me and say 'Look, Miss, I'm not getting on in my lesson because… so-and-so and so-and-so are being disruptive', this is happening in this lesson, and this is happening in this lesson. So, they're taking responsibility for themselves and for others coming to report that. Then with that information I can then speak to those subject teachers and say, 'Just so you know, I've had a report from these students that there's been a bit of low-level disruption distracting this girl or boy in the class – can you keep an eye on it?' - *Teacher, Secondary school*

> I would say in Year 6 particularly we have seen pupils actually holding each other to account, so if they are talking in the corridors, they will be turning around and telling them 'You're talking' and to be quiet. - *Senior leader, Primary school*

In a separate instance, a non-teaching pastoral member of staff referred to increasing numbers of pupils who were now proactively seeking out school-led restorative approaches to resolve disputes or incidents before they escalated further:

> Some students come in and say, 'Can we have a meeting, I fell out with such-and-such on the bus this morning'. Yes, we'll do it period 5. - *Pastoral member of staff, Secondary school*

One school was trialling training their prefects to act as mentors and positive role models with the ultimate aim of getting to the stage where they would be given the authority to *deliver sanctions to other students* [when appropriate to do so] - *Senior leader, Secondary school*

The vignette below epitomises a school (please refer to Section 3.2 for details on School B's background and implementation) that had travelled a particularly significant distance in creating effective behaviour strategies that had become consistently embedded on a school-wide basis.

School B: Target school (Running a School)

Enhanced and sustained senior leader strategies to 'empower' staff to make more informed individual judgements were **beginning to have a transformative effect on the whole-school implementation of a coherent behaviour strategy**, to the extent that a culture of individual accountability was beginning to take hold in individual classrooms:

> This year I've taken on board what happened in the training, and I've really implemented it and it's worked really well, especially with Key Stage 3. I'm just really consistent with every class. They all know the expectations. And everything that I'm doing in those classes is from that training [delivered through staff at School B]. It's worked really well. - *Teacher*

And throughout the school more widely:

> I love it now, you say 'Shirt' to somebody, and they'll be like, 'You're like the fourth person who's told me this in five minutes!!' - *Teacher*

Although teaching staff were universally positive about the progress made and departing from an over-reliance of purely 'top-down' approaches, certain members did feel the absence of a formal whole-school behaviour policy did on occasion lead to 'confusion'. This was felt to be a possible next stage for senior leaders to address.

A number of schools had focused on creating a culture of mutual respect amongst staff. This tended to be facilitated through more effective and overt leadership **messaging of expectations**. Different members of school staff frequently said that situations in which only certain (typically senior) members of staff were able to meaningfully deal with behaviour (as was the case in the past) were increasingly rare. The following quotation evidences an example of a more egalitarian approach to whole-school behaviour management:

> I think the hierarchy a lot of the time used to be the problem. I think having certain staff who would deal with everything was

> part of the issue, because the student would only engage when it got to that stage, and their time was then being spent doing everything, when actually it's a whole staff concept. It doesn't matter if it's a learning mentor talking to you, it doesn't matter if it's Mrs ___. It doesn't matter who that person is. - *Teacher, Secondary school*

A senior leadership approach that was able to effectively affect this transition in whole-school culture was top-level messaging accompanied by more nuanced explanations. These included why it was necessary to operate in a more universal manner and what the adverse implications of not doing so were. The following vignette exemplifies how tightened senior leadership practices, and operationalising a deliberately utilitarian approach to school behaviour, resulted in individual practice both in the classroom and around the school becoming far more standardised and reflective of a whole-school behaviour culture.

School D: Non-target secondary school (Running a School)

This school was located in an inner-city context with significant proportions of pupil premium entitlement. Part of a national MAT, the school had travelled significantly in recent times, moving out of special measures. Following involvement in the TBT project, the senior leader interviewed described gains in realising a more effective whole-school behaviour culture. This was driven by the senior leadership team consciously putting limits on the amount of discretion individual teachers had in relation to enforcing the pupil behaviour management policy. Instead, emphasis was placed on the need for more consistent whole-school approaches that valued **collective staff accountability for behavioural policy enforcement:**

> I think the clarity of what our rules are and that those rules are the same whether they're being enforced by a TA or a principal. I think that's quite empowering to the staff body. I think having a shared language, having a clear sort of plan. When you have a **clear policy and road map about what behaviour is like, it takes a lot of pressure off the individual decision-making**. You're all answering to a higher power really – the students and the staff. So yes, I think that's had a massive impact on new staff, on existing staff. - *Senior leader*

Through improved leadership practices, senior leaders were able to make a compelling case for teachers to consistently follow the behaviour policy in order to **achieve better whole-school pupil behaviour** and generally **greater collegiality among colleagues** … **group consistency is a hard pill to**

> **swallow, if you're used to being good at something**. Because you're no longer the one… we all get a bit of a buzz from being the one teacher who can teach Year 8 set whatever. And that buzz comes from the fact that their maths teacher can't control them and their science teacher can't control them…You're the superhero…But what Tom is sort of driving at [what leadership practice mimicked] takes that away, because the things I'm doing in my classroom should make it easier for my colleague in maths to do their job. - *Senior leader*

According to one senior leader, these revised whole-school behavioural strategies were responsible for a more even adherence to the school rules across year groups.

I think that the culture of our current Year 11 is very much that they're… first among equals at best, but they're not different really to the Year 7s in terms of how they follow the school policies. And I think that's a big change from the past. – *Senior leader*

Greater efforts to ensure all staff directly received whole-school guidance and expectations relating to behaviour management practices were frequently noted. In certain instances, this marked a pronounced change from the past and meant that, for the first time, all staff were invited to receive behaviour management-related CPD:

> I think clear expectations and understanding. A really positive model of behaviour that is consistent throughout the school…I think everything is just [slicker]. Everybody understands, everybody is following… you can go from classroom to classroom to classroom and the expectations will be the same. It's clear, communicated expectations from the senior leadership team, **shared with all members of staff, whether that be support staff, midday supervisors, the people who provide lunch**. - *Senior leader, Primary school*

As was previously outlined in Section 3.3, it can be challenging to drive forward a whole-school behaviour policy in the context of so many different competing priorities, particularly in large secondary school environments. The following senior leader from School A (see Section 3.2 for details on their background and extent of

implementation) felt reassured that their school was heading in the right direction when they made the following observation:

> Pre-TBT, he [misbehaving pupil] would have just gone into another classroom and carried on with his work. Post… English decided to create their own restorative slip as they've called it… I haven't told English to produce this, so I find this reassuring and it makes me proud and makes me think we're doing the right job. *- Senior leader, Secondary school*

Finally, we return to School C, a school that had experienced significant previous difficulties (see Section 3.2 and 3.3.1 for further background), but that was now led by a confident and focused SLT that did not hesitate in following through their vision for school behaviour despite it being at odds with TBT Project Leads. This school's experience demonstrates an example of how taking a different pathway to that suggested by TBT Project Leads could lead to equally positive outcomes.

School C - Priority secondary school (Running a Room)

School C was reported to have successfully implemented reforms to their consequence system and running a single Internal Exclusion Room (IER), thereby confounding advice by TBT Project Leads to not do so. TBT Project Lead concerns about the IER environment becoming overwhelmed and unbefitting to a calm learning environment were claimed by senior leaders to have never materialised. The whole school staff accepted the focused changes to the consequence system, because they understood the logic and understood that it was a more effective deterrent. The following quotation showcases how pupils soon became aware it was not in their interests to persistently misbehave:

> It makes them question whether or not it's worth doing what they're about to do. So once [and I'm noticing] as a classroom teacher, it's now kids sometimes who will just stop and think – is this worth bothering? Because actually it's a directly applicable consequence and straight away, and if they do something again that day, then they've got even more to answer to… *- Teacher*

Subsequent to our case-study visit, an Ofsted report was published judging School C to be good overall, with Outstanding Leadership and Management. Within the Ofsted report, reformed behaviour policies drew particular praise, and were described as the feature that had improved most starkly since the previous Ofsted inspection.

4.2.2 Findings related to improved teaching quality

The majority of teaching staff interviewed outlined at least moderately **increased confidence** in their ability to employ effective behaviour management strategies in their classroom. ECTs and those experiencing the TBT project directly through the Running a Room course were most likely to demonstrate increases in confidence that could be directly attributed to the TBT project. Teachers' capacity to truly demonstrate improved teaching quality was clearly intimately linked to the effectiveness of senior leadership practice and how routinely whole-school behaviour policies were being implemented across the staff:

> I feel quite empowered actually that I can deal with issues at my door rather than taking them through to the office and someone else…. I can hold pupils to account and I know that [headteacher] is going to be backing me up 100% if I am going to make a call with regards to consequences…involving children. - *Teacher, Primary school*

The more effectively implemented whole-school behaviour management strategies tended to filter down to the classroom level, freeing up greater opportunities for teachers to showcase their pedagogical skill and application of high-quality teaching, as opposed to becoming side-tracked by occurrences of low-level pupil disruption (see Section 4.3.3 for further detail on school cultural shifts). Aided by increased confidence in whole-school approaches and being appropriately backed by senior staff, the thresholds for tolerating poor pupil behaviour were said to have reduced in the classroom. Over time, this was said to raise classroom-level behaviour expectations:

> I got to that point where I knew nothing would work so there was really not much point in me throwing somebody out anyway. But I have found that I do use the system more than I ever did, because my expectations are different. I'm not going to take the low-level stuff anymore. I am going to deal with it and make sure it's sorted out. - *Teacher, Secondary school*

A key overarching principle that teachers attempted to adopt into their teaching, and that was largely credited with improving teaching quality, was the conscious attempt to maintain 'dignity and learning'. This concept was operationalised in a variety of ways including the use of bespoke PowerPoint slides clearly outlining classroom expectations, along with an age-appropriate explanation of why it was important to 'preserve everyone's dignity and learning'; something that was then reinforced verbally as required:

> Praising a kid in public, and reprimanding in private, and things like that. Something that is easy to do, just – 'Why did you do that?' You know, but then having a quiet word with that kid. And it's all about if we change our attitudes then obviously the kid will likewise do the same. - *Teacher, Secondary school*

Related to the concept of 'dignity and learning', other teachers referenced applying a 'toolkit' of different de-escalation strategies and techniques to help defuse potentially stressful situations from spiralling. Applying techniques such as these were helpful for dealing with recurrently misbehaving pupils and meant reducing the likelihood of needing to send them out of class, something that would potentially diminish their learning potential:

> The smaller things that might have escalated, now they don't escalate. The staff know how to deal with those in a really productive and careful way…setting those routines out for members of staff. Something as simple as having a classroom [routine] [to settle] before you come in. [Headphones out], phones are switched off. Staff are aware of the best possible way of dealing with that, being consistent with the students and having that routine. - *Senior leader, Secondary school*

Where it was necessary to enforce the consequence system, the focus tended to be aimed at the restorative as opposed to the punitive, with teachers noting the need for pupils to understand why they had been punished and to ascertain if there were any underlying reasons for their misbehaviour that needed to be addressed. Many interviewee teachers suggested that modifying their own behaviour and openness to more nurture-based and restorative approaches had positively influenced their relationship with pupils. Pupil-staff relationships were further enhanced in schools that attempted to implement appropriate reward systems, and that encouraged more informal opportunities to relate to each other (see Section 4.2.3 below).

4.2.3 Findings related to school culture and staff satisfaction

Most participants referred to calmer, happier, more enriched and generally far more positive school cultures:

> It's really opened up a much nicer culture, much more positive. -*Teacher, Secondary school*

> So that general calmness around the place. The walking around the school I think is calmer. - *Teacher, Primary school*

> Kids smile. That's the big thing. - *Headteacher, Secondary school*

In most cases, positive changes to school culture were consistent with an existing trajectory of gradually consistent improvement. This, for certain schools, would have been incomprehensible two or three years ago:

> If I compare these Year 7s starting September and the Year 7s that started last September, I think they knew that the school didn't really have this culture and ethos…The way we do things is really different. And they're quiet, and they listen, and they want to learn and they will do the things that you want them to do. - *Senior leader, Secondary school*

Unsurprisingly, school cultural shifts of this magnitude had a profoundly uplifting effect upon staff morale, as is outlined in the following quotation:

> I mentioned the science teacher – telling us that they can teach. He sent an email saying that – he just said thanks; **I can't believe it, best teaching in 20 years**. They've been able to do those sorts of things. I would have hated to have been a classroom teacher here two years ago. It must have been horrific. - *Senior leader, Secondary school*

In a different example, the staff from a target primary school completed a local authority staff wellbeing survey following the implementation of initial reforms perceived to be in part inspired by the TBT project. The feedback received from the local authority was that wellbeing was higher than in other schools with 'similar issues' and that '*indicators around feeling in control and being supported*' were even higher than the local authority average.

More routinely, teachers attributed improved relationships with pupils as being related to clearer conveyance of behavioural expectations, in tandem with more consistent staff responses to misbehaviour and reducing the scope for pupils to feel aggrieved with sanctions. As one primary senior leader reflected *'these are the rules, so if you make the wrong choice, you can't argue… 'it's not fair'.*

A further reported feature, which underpinned improved school cultures, was a greater emphasis on nurture-based approaches, focused more on achievement recognition as opposed to solely punitive behaviour policy enforcement. This played out in various ways including:

- end of term/year rewards for pupils who had displayed good behaviour (for example being entered for a draw for cinema tickets)
- getting staff involved in running voluntary lunchtime/break time activities, clubs and societies ranging from darts to Deal or No Deal to knitting clubs:

  > They [pupils] enjoy seeing us [staff] not in just a teaching role, which is nice. - *NQT, Secondary school*

- setting up, reviving, or expanding house systems including competitive sporting competitions:

  > So, every week, it's… again it's about that culture of 'we're a team', 'we're a house team'. - *Senior leader, Primary school*

All of these approaches tended to enhance pupil and staff relationships, reduce opportunities for misbehaviour over break/lunchtimes and further incentivise pupils to voluntarily accept the school's culture as opposed to rebelling against it:

> Everyone is aspiring to get that praise. People [aren't] aspiring to be – [laugh at me – I'm a clown]. - *Teacher, Secondary school*

4.2.4 Findings related to CPD quality and staff engagement in CPD[13]

There was very limited evidence of any significant shifts in how CPD was valued or an upsurge in demand for it more broadly. However, one primary-based senior leader did directly refer to the TBT project having acted as a reminder that 'evidence-based practice' was *'really important'* and to be more mindful to apply this to '*other areas of running the school*'.

In addition, as outlined in Section 4.2.1, the TBT project did plug a gap for the CPD of non-teaching and support staff, thereby offering greater inclusivity (albeit often on a voluntary basis). It would be interesting to monitor whether this trend extends beyond the life of the project and to other areas of school life.

4.2.5 Findings related to retention and progression impacts

This section explores the extent to which the TBT project achieved its expected impacts in relation to teacher retention and progression (through analysis of teacher

[13] Please note there is insufficient data across the qualitative projects to justify separating these factors out as was appropriate for other TLIF projects with a survey dimension.

outcomes in the SWC[14]). It also explores participants' perceptions of the impact of the project on teacher retention and progression (through analysis of qualitative data).

Retention and progression analysis

The evaluation aimed to explore the impact of the TBT project on the fund-level goals to improve teacher retention and progression. As outlined previously, the TBT project intended to achieve both teacher-level and whole-school-level impacts. Therefore, this analysis was conducted on TBT participants and a matched comparison sample of teachers (teacher-level impacts), and on all teachers from TBT schools (whole-school impacts) and all teachers in a matched comparison sample of schools. The analysis matched participating teachers and participating schools to respective comparison groups using a range of key characteristics (see Appendix D Tables 23 and 24) to estimate what counterfactual retention and progression rates might have been with and without the TBT project.

As such, the findings are reported in two sections; one section reporting the impact the TBT project had on teacher-level retention and progression and another section using school-level data to explore the impact TBT had on school-level retention and progression.

Teacher retention was analysed in terms of:

- retention in the state-funded sector in England
- retention in the school
- retention in the same LA
- retention in challenging schools.[15]

Teacher progression was analysed in terms of:

- progression in the state-funded sector in England
- progression in the school
- progression in the same LA
- progression in challenging schools.

[14] This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

[15] Challenging' schools were defined as schools rated by Ofsted as 'requires improvement' or 'inadequate'. A teacher was defined as remaining in a challenging school if they either stayed within the school they were in at baseline, or moved to another school which was rated 'requires improvement' or 'inadequate'.

Teacher-level retention

The following sections discuss the findings of the SWC secondary analysis at the teacher level. The tables below summarise TBT's estimated impacts across the four retention measures analysed. We use the descriptor 'teacher-level' to describe analyses of all project participants, irrespective of their level of seniority. It is important to note that, since it was not possible to distinguish whether participants attended the TBT leader training course (Running a School) or TBT teacher course (Running a Classroom) in the teacher-level analyses, interpretations cannot be drawn out about the relative impacts of the different TBT courses for leaders and teachers.

Retention in the state-funded sector in England

Table 5: Difference in the estimated rate of retention in state-funded teaching in England between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated retention rate in state-funded teaching 1 year after baseline (%)	93.0	87.8	5.3	Yes
Number of teachers	**555**	**3930**		
Estimated retention rate in state-funded teaching 2 years after baseline (%)	85.1	81.3	3.8	Yes
Number of teachers	**517**	**3561**		
Estimated retention rate in state-funded teaching 3 years after baseline (%)	79.4	75.7	3.7	No
Number of teachers	**294**	**1832**		

Note: Estimated retention rates are the average predicted retention rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers

Analysis presented in Table 5 shows that the TBT project was associated with a statistically significant higher rate of retention within the state-funded teaching profession one and two years after the baseline but not three years after the baseline

data was collected., with treatment teachers being 5.3 percentage points more likely to be retained in teaching one year after baseline, 3.8 percentage points two years after baseline. This may suggest that the TBT project had some impact on teacher retention in the profession. However, the presence of a significant impact only one year after baseline indicates that there may have been systematic differences between the treatment and comparison samples that are not accounted for in this analysis. At one years after baseline participants would have only received a limited amount of the TBT training. As such, it is implausible to assume that such limited training would have had such a large impact on retention. It is more likely that systematic differences between the treatment and comparison samples that are not accounted for in this analysis would explain these findings. his limitation also applies to findings in Tables 6-8 below.

Retention in the school

Table 6: Difference in the estimated rate of retention in the same school between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated retention rate in the same school 1 year after baseline (%)	93.6	86.7	6.9	Yes
Number of teachers	**441**	**3094**		
Estimated retention rate in the same school 2 years after baseline (%)	82.8	79.2	3.6	No
Number of teachers	**411**	**2875**		
Estimated retention rate in the same school 3 years after baseline (%)	75.8	71.4	4.4	No
Number of teachers	**222**	**1439**		

Note: Estimated retention rates are the average predicted retention rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

Analysis presented in Table 6 shows that the TBT project was associated with a statistically significant higher rate of retention within the state-funded teaching profession only at one year after baseline, with treatment teachers being 6.9 percentage points more likely to be retained in teaching at that point. Again, as with

retention in the same school, such a finding was likely due to systematic differences unaccounted for in this analysis, rather than the effects of the TBT training.

Retention in the same local authority

Table 7: Difference in the estimated rate of retention in the same local authority district (LAD) between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated retention rate in the same LAD 1 year after baseline (%)	95.4	90.4	5.0	Yes
Number of teachers	**441**	**3094**		
Estimated retention rate in the same LAD 2 years after baseline (%)	87.0	85.0	2.0	No
Number of teachers	**411**	**2875**		
Estimated retention rate in the same LAD 3 years after baseline (%)	81.3	79.7	1.6	No
Number of teachers	**222**	**1439**		

Note: Estimated retention rates are the average predicted retention rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

Analysis presented in Table 7 shows that the TBT project was associated with a statistically significant higher rate of retention within the same LAD but only one year after baseline, when treatment teachers were 5.0 percentage points more likely to be retained in teaching. As previously discussed, at one year after baseline participants may have only received a limited amount of the TBT training. As such, it is implausible to assume that limited training would lead to such an impact and it is more likely that systematic differences between the treatment and comparison samples (that are not accounted for in this analysis) would explain these findings.

Retention in challenging schools

Table 8: Difference in the estimated rate of retention in challenging schools[16] between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated retention rate in challenging schools 1 year after baseline (%)	95.9	90.5	5.4	Yes
Number of teachers	**436**	**3050**		
Estimated retention rate in challenging schools 2 years after baseline (%)	87.4	84.9	2.5	No
Number of teachers	**402**	**2809**		
Estimated retention rate in challenging schools 3 years after baseline (%)	79.2	79.6	-0.4	No
Number of teachers	**218**	**1399**		

Note: Estimated retention rates are the average predicted retention rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

Analysis presented in Table 8 shows that the TBT project was associated with a statistically significant higher rate of retention within the state-funded teaching profession only at one year after baseline, when treatment teachers were 5.4 percentage points more likely to be retained compared to non-participating teachers.

As previously discussed, at one years after baseline participants would have only received a limited amount of the TBT training. As such, it is implausible to assume that limited training would lead to such an impact, and it is more likely that systematic differences between the treatment and comparison samples (that are not accounted for in this analysis) would explain these findings.

Across all of the retention measures, at teacher level, the largest impacts are observed in the first year after baseline. The impacts are only significant one year after baseline for three of the retention measures and in years one and two for

[16] For the purposes of this analysis, 'challenging' schools were defined as schools rated by Ofsted as 'requires improvement' or 'inadequate'. A teacher was defined as remaining in a challenging school if they either stayed within the school they were in at baseline, or moved to another school which was rated 'requires improvement' or 'inadequate'.

retention in the state-funded sector. At one year post baseline participants may have only received limited training, in addition one year is a very short timeframe for the project to have influenced retention. It is therefore implausible to associate impacts one year post baseline to the TBT training. It is more likely that the impact observed was due to selection effects, systematic differences between the treatment and comparison teachers that could not be controlled for in the matching and subsequent analysis. In addition, a key limitation of this analysis is that we could not accurately observe when teachers were recruited to the project or when they finished their training. As such the assignment of a baseline year was a best approximation of the year before teachers started the TBT training. It is possible that some teachers would have still been participating in the project when retention rates were compared one year after baseline. This provides further explanation as to why retention rates for treatment teachers were so high in the first year after baseline and why they then began to fall two and three years after baseline.

Progression in the state-funded sector in England

The tables below summarise the TBT impacts across the four progression measures analysed at teacher level. Progression rates are defined as the proportion of teachers who moved from either a classroom teacher to a middle/senior leader role, or a middle leader role to a senior leader role within one and two years of baseline

Table 9: Difference in the estimated rate of progression in state-funded teaching in England between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated progression rate in state-funded teaching 1 year after baseline (%)	9.6	7.6	2.0	No
Number of teachers	**346**	**2425**		
Estimated progression rate in state-funded teaching 2 years after baseline (%)	15.7	13.1	2.7	No
Number of teachers	**319**	**2253**		

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated progression rate in state-funded teaching 3 years after baseline (%)	28.4	16.8	11.6	Yes
Number of teachers	**158**	**1098**		

Note: Estimated progression rates are the average predicted progression rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

The analysis displayed in Table 9 shows that there were no significant differences in the progression rates of treatment and comparison teachers in state-funded teaching one or two years after baseline. However, a significant difference of 11.6 percentage points was observed three years after baseline. As the initial progression effect is small and grows over time, it is plausible that the impact on progression in state-funded teaching could be attributable to the TBT project.

Progression in the school

Table 10: Difference in the estimated rate of progression in the same school between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated progression rate in the same school 1 year after baseline (%)	9.6	6.8	2.8	No
Number of teachers	**319**	**2120**		
Estimated progression rate in the same school 2 years after baseline (%)	15.3	11.7	3.6	No
Number of teachers	**255**	**1808**		
Estimated progression rate in the same school 3 years after baseline (%)	21.7	14.8	6.9	No
Number of teachers	**113**	**801**		

Note: Estimated progression rates are the average predicted progression rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

The analysis displayed in Table 10 shows that there were no significant differences in the progression rates of treatment and comparison teachers who stayed in the same school, either one, two or three years after baseline. These findings suggest that the TBT project had no impact on progression in teaching in the same school. As the findings from table 9 suggest that the TBT project did have an impact on progression, teachers who did progress were likely to have done so by moving to a different school.

Progression in the same local authority

Table 11: Difference in the estimated rate of progression in the same local authority district (LAD) between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated progression rate in the same LAD 1 year after baseline (%)	10.1	7.1	3.1	No
Number of teachers	**327**	**2208**		
Estimated progression rate in the same LAD 2 years after baseline (%)	16.5	11.6	4.9	Yes
Number of teachers	**272**	**1938**		
Estimated progression rate in the same LAD 3 years after baseline (%)	25.4	15.2	10.2	Yes
Number of teachers	**122**	**898**		

Note: Estimated progression rates are the average predicted progression rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

The analysis displayed in Table 11 shows that there were no significant differences in the progression rates of treatment and comparison teachers in the same local authority one year after baseline. However, a significant difference was observed after two and three years. Since the initial progression effect was small and grows

over time, it is plausible that the impacts in years two and three could be attributable to the TBT project.

Progression in challenging schools

Table 12: Difference in the estimated rate of progression in challenging schools[17] between treatment and comparison teachers

	Treatment teachers	Comparison teachers	Difference	Statistically significant?
Estimated progression rate in challenging schools 1 year after baseline (%)	9.1	7.3	1.8	No
Number of teachers	**324**	**2180**		
Estimated progression rate in challenging schools 2 years after baseline (%)	15.4	12.5	2.9	No
Number of teachers	**263**	**1894**		
Estimated progression rate in challenging schools 3 years after baseline (%)	23.1	16.3	6.9	No
Number of teachers	**116**	**864**		

Note: Estimated progression rates are the average predicted progression rates from a logistic regression model for treatment and comparison teachers, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of this difference is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison teachers.

The analysis displayed in Table 12 shows that there were no significant differences in the progression rates in challenging schools of treatment and comparison teachers, either one, two or three years after baseline.

The findings on teacher-level progression are mixed. It can be suggested that the TBT programme had a significant impact on progression, and that those teachers who progressed did so at other schools, within the same LAD. Where significant changes were observed they occurred at two and three years after baseline.

[17] For the purposes of this analysis, 'challenging' schools were defined as schools rated by Ofsted as 'requires improvement' or 'inadequate'. A teacher was defined as remaining in a challenging school if they either stayed within the school they were in at baseline, or moved to another school which was rated 'requires improvement' or 'inadequate'.

Therefore it is plausible that these effects were due to the TBT programme itself rather than systematic differences in individual teachers.

Although there was significant evidence that the TBT project had contributed towards enhanced staff satisfaction and more positive school cultures (see Sections 4.2 and 4.3), most interviewees said it had not had a direct bearing on them being more likely to stay in the profession or having contributed to their career progression. The comparatively short timescale of the project intervention, combined with the lack of distance from programme end to qualitative data collection, meant that insufficient time had elapsed for participants to assess potential impact on retention and progression.

School-level retention

The following sections explore the findings from the SWC secondary analysis on retention at the school level (school-level impacts).

Retention in the state-funded sector in England

Table 13: Difference in rate of retention in state-funded teaching in England

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated retention rate in state-funded teaching 2 years before baseline	89.8	88.9	0.9	-	-
Estimated retention rate in state-funded teaching 1 year before baseline	88.3	89.2	-1.0	-	-
Estimated retention rate in state-funded teaching 1 year after baseline	91.1	90.0	1.1	1.1	No
Estimated retention rate in state-funded teaching 2 years after baseline	90.4	90.5	-0.1	-0.1	No
Estimated retention rate in state-funded	90.0	91.7	-1.8	-1.7	Yes

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
teaching 3 years after baseline					
Number of schools	**69**	**584**	-	-	-

Note: Estimated retention rates are the average predicted retention rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis shown in Table 13 reveals one significant finding. The difference between treatment and comparison schools three years after baseline was significantly different to before the project started. Before the project started treatment schools had lower retention rates in state-funded teaching than comparison schools by an average of only 0.05 percentage points. However, three years after baseline treatment schools were 1.8 percentage points less likely than comparison schools to retain teachers. This difference appears to have been caused by a notable increase in the retention rate in comparison schools in this year (rather than the difference being caused by a change within treatment schools).

Retention in the school

Table 14: Difference in rate of retention in the school

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated retention rate in the same school 2 years before baseline	91.5	90.5	1.0	-	-
Estimated retention rate in the same school 1 year before baseline	90.0	90.6	-0.6	-	-
Estimated retention rate in the same school 1 year after baseline	90.1	90.9	-0.9	-1.1	No
Estimated retention rate in the same	91.7	91.9	-0.1	-0.4	No

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
school 2 years after baseline					
Estimated retention rate in the same school 3 years after baseline	93.7	93.3	0.5	0.2	No
Number of schools	**69**	**584**	-	-	-

Note: Estimated retention rates are the average predicted retention rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis displayed in Table 14 did not demonstrate any differences in retention rate within the same school between treatment and comparison schools. This contrasted to the findings in Table 13. However, the findings in Table 13 were thought to be due to a notable increase in the retention rate in comparison schools in this year, rather than a change within treatment schools. Table 14 demonstrates that the pattern in retention rates within the same school was comparable between both treatment and comparison school.

Retention in the same local authority

Table 15: Difference in rate of retention in the same LA

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated retention rate in the same LA 2 years before baseline	93.5	93.2	0.3	-	-
Estimated retention rate in in the same LA 1 year before baseline	92.6	93.5	-0.9	-	-
Estimated retention rate in in the same LA 1 year after baseline	92.5	93.9	-1.3	-1.1	No
Estimated retention rate in in the same LA 2 years after baseline	94.1	94.3	-0.3	0.0	No

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated retention rate in in the same LA 3 years after baseline	95.2	95.5	-0.3	0.0	No
Number of schools	**69**	**584**	-	-	-

Note: Estimated retention rates are the average predicted retention rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis displayed in Table 15 did not demonstrate any differences in retention rate within the same LAD between treatment and comparison schools. This contrasted to the findings in table 13. However, the findings in Table 13 were likely due to a notable increase in the retention rate in comparison schools in this year, rather than a change within treatment schools. Table 15 demonstrated that the pattern in retention rates within the same LA was comparable between both treatment and comparison school.

Retention in challenging schools

Table 16: Difference in rate of retention in challenging schools[18]

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated retention rate in challenging schools 2 years before baseline	95.6	94.8	0.8	-	-
Estimated retention rate in challenging schools 1 year before baseline	94.7	94.8	-0.1	-	-
Estimated retention rate in challenging schools 1 year after baseline	94.7	95.1	-0.3	-0.7	No

[18] For the purposes of this analysis, challenging schools were defined as schools rated by Ofsted as 'requires improvement' or 'inadequate'. A teacher was defined as remaining in a challenging school if they either stayed within the same school, or they moved to a different school which was rated 'requires improvement' or 'inadequate'.

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated retention rate in challenging schools 2 years after baseline	95.4	95.7	-0.3	-0.6	No
Estimated retention rate in challenging schools 3 years after baseline	96.3	96.4	-0.1	-0.4	No
Number of schools	**69**	**584**	-	-	-

Note: Estimated retention rates are the average predicted retention rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted retention rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis displayed in Table 16 does not demonstrate any differences in retention rate within challenging schools between treatment and comparison schools. Again, this contrasts with the findings in Table 13. However, the negative findings in Table 13 are thought to be due to a notable increase in the retention rate in comparison schools in this year, rather than a change within treatment schools. Table 16 demonstrates that the pattern in retention rates within challenging school is comparable between both treatment and comparison schools.

The findings discussed above suggest that there may have been a negative impact on school level retention in state-funded teaching three years after baseline. However, for the other three retention measures, no impact (negative or positive) was observed. It was beyond the scope of the evaluation to collect qualitative data from non-participants on indicators of school-level retention.

School-level progression

The following sections explore the findings from the SWC secondary analysis on progression at the school level (school-level impacts).

Progression in the state-funded sector in England

Table 17: Difference in rate of progression in state-funded teaching in England

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated progression rate in state-funded teaching 2 years before baseline	7.9	5.8	2.2	-	-
Estimated progression rate in state-funded teaching 1 year before baseline	5.5	5.4	0.0	-	-
Estimated progression rate in state-funded teaching 1 year after baseline	5.9	5.3	0.7	-0.4	No
Estimated progression rate in state-funded teaching 2 years after baseline	4.7	4.7	0.0	-1.1	No
Estimated progression rate in state-funded teaching 3 years after baseline	5.0	4.2	0.8	-0.3	No
Number of schools	**69**	**583**	-	-	-

Note: Estimated progression rates are the average predicted progression rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis displayed in Table 17 does not demonstrate any differences in progression within state-funded schools between treatment and comparison schools.

Progression in the school

Table 18: Difference in rate of progression in the school

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated progression rate in the same school 2 years before baseline	6.7	4.9	1.8	-	-
Estimated progression rate in the same school 1 year before baseline	4.6	4.5	0.1	-	-
Estimated progression rate in the same school 1 year after baseline	4.8	4.3	0.5	-0.4	No
Estimated progression rate in the same school 2 years after baseline	3.8	3.8	0.0	-1.0	No
Estimated progression rate in the same school 3 years after baseline	3.9	3.4	0.5	-0.5	No
Number of schools	**69**	**574**	-	-	-

Note: Estimated progression rates are the average predicted progression rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis displayed in Table 18 does not demonstrate any differences in progression within the same school between treatment and comparison schools.

Progression in the same local authority

Table 19: Difference in rate of progression in the same LA

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated progression rate in the same LA 2 years before baseline	7.2	5.1	2.0	-	-
Estimated progression rate in in the same LA 1 year before baseline	4.7	4.8	0.0	-	-
Estimated progression rate in in the same LA 1 year after baseline	5.2	4.6	0.7	-0.3	No
Estimated progression rate in in the same LA 2 years after baseline	4.2	4.1	0.1	-0.9	No
Estimated progression rate in in the same LA 3 years after baseline	4.2	3.6	0.6	-0.4	No
Number of schools	**69**	**582**	-	-	-

Note: Estimated progression rates are the average predicted progression rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

The analysis displayed in Table 19 does not demonstrate any differences in progression within the same LAD between treatment and comparison schools.

Progression in challenging schools

Table 20: Difference in rate of progression in challenging schools[19]

	Treatment group	Comparison group	Difference	Difference-in-difference	Statistically significant?
Estimated progression rate in challenging schools 2 years before baseline	7.3	5.1	2.2	-	-
Estimated progression rate in challenging schools 1 year before baseline	4.9	4.8	0.1	-	-
Estimated progression rate in challenging schools 1 year after baseline	5.2	4.5	0.7	-0.5	No
Estimated progression rate in challenging schools 2 years after baseline	4.0	4.0	0.0	-1.1	No
Estimated progression rate in challenging schools 3 years after baseline	4.2	3.6	0.6	-0.6	No
Number of schools	**69**	**580**	-	-	-

Note: Estimated progression rates are the average predicted progression rates from a logistic mixed-effects regression model for treatment and comparison schools, controlling for observed characteristics. The difference in average predicted progression rates is the marginal effect. Statistical significance of these differences is assessed at the five per cent level. Due to rounding, some estimated marginal effects may not exactly equal the difference between treatment and comparison schools.

[19] For the purposes of this analysis, challenging schools were defined as schools rated by Ofsted as 'requires improvement' or 'inadequate'. A teacher was defined as remaining in a challenging school if they either stayed within the same school, or they moved to a different school which was rated 'requires improvement' or 'inadequate'.

The analysis displayed in Table 20 does not demonstrate any differences in progression within challenging schools between treatment and comparison schools.

The findings detailed above suggest that, at the school level, the TBT programme did not have an impact on teacher progression rates. It was beyond the scope of the evaluation to collect qualitative data from non-participants on indicators of school-level retention and progression.

4.2.6 Indicative findings related to pupil attainment

While there were no reports of direct impacts on pupil attainment, reported improvements in intermediate pupil outcomes may support longer-term improvements in attainment. Most schools referenced significant reductions in the number of in-school sanctions (e.g. lunchtime and after-school detentions) taking place:

> If I was to compare week-on-week the lesson removals for spring term and summer term, we had an 87% reduction. We don't even look at which staff are removing children now because there are so few children removed. - *Senior leader, Secondary school*

However, some interviewees mentioned, at least initially, a spike in the number of detentions, which tended to be attributed to the more consistent application of the school rules:

> The number of removals from lessons has actually gone up. But I think that's because teachers feel more confident to remove students, whereas before they would just have struggled on. Whereas now they will actually say no this is not acceptable, you're disrupting learning, and they're prepared to actually remove them from the lesson. - *Senior leader, Secondary school*

Participants were less comfortable sharing any hard data on exclusions and attainment monitoring data or speculating on intermediate shifts - feeling it was unrealistically early to expect an impact. However, where they were willing, the trends were universally positive, both in terms of exclusions and attainment:

> We've reduced fixed-term exclusions this year in comparison with this time last year, by 50 per cent. - *Senior leader, Secondary school*

> Our results have gone up quite dramatically…we've now got a positive Progress score whereas before it was very low…. parents, the community, have got more faith in what the school has to offer, certainly. - *Headteacher, Secondary school*

Of course, we cannot attribute these perceived outcomes to the TBT project at this stage, but there is evidence to suggest that negative views held across local communities about certain target schools were beginning to lift. For example, some school leaders reported pupils that had moved to other schools previously were now applying to return, and in some cases Year 7 intakes were now over-subscribed.

4.3 Summary of outcomes and impacts

It is difficult to robustly disaggregate the relative contribution that involvement in the TBT project made to the outcomes outlined in Section 4.2. However, in light of the contextual information provided in Section 4.1, the TBT project appears to have facilitated a range of positive improvements in schools as well as changes in leadership, teaching and wider staff practices, and indications of some positive changes in pupil behaviour to a level at least in line with original hypothesised expectations.

There was some evidence to suggest that the TBT project may have had a positive impact on **teacher level retention** in the state-funded sector two years after baseline and **teacher-level progression** in state-funded schools at year three and in the same LAD at years two and three. We can only speculate as to the reasons for this, given no follow up interviews with teachers were undertaken beyond the project.

However, based on qualitative findings obtained shortly after the conclusion of the programme, it is not unreasonable to suggest revised behaviour management strategies, might have had a positive effect on behaviour management at the classroom and wider school level, making participants more committed to staying within the profession.

5. Sustainability

The abiding impression given by interviewees was that there was a genuine commitment to not only sustain existing learning, but to actively monitor, review and evolve in whichever directions were needed to maintain and improve pupil outcomes. The following quotation exemplified how many school interviewees were aware of the importance to guard against complacency:

> I think once you've got a system and you think, well that's it, well you're already losing. You need to keep developing and keep changing it and keep working out what's best for the students. - *Senior leader, Secondary school*

This mindset among leaders was particularly encouraging, suggesting an awareness of the importance of continued reinvention as opposed to treating behaviour policies as static documents, rarely to be re-visited. Tom Bennett articulated just how important he felt this particular theme was to realising sustained positive outcomes:

> Something we made explicit in our theory of change and our original application, which was that cultures are only sustained as long as somebody sustains them. They must be constantly created. They are acts of constant creation... Frequently we see very negative cultural change happen in schools, because key members of staff leave, because they are the people that tend to embody the programmes that they support. And that particularly from a leadership perspective, one of the ways in which you create a culture which is meaningfully sustainable is by creating structures, which survive your demise or departure…So, their behaviour policy…should actually be an embodiment of what it means to be them, with types of things that they do, and the types of behaviour they expect and how they're taught, and who teaches them, and who's held accountable for them… the behaviour policy should be the vertebrae of the behaviour of the school, and should not only be an embodiment, but a constant invitation to revitalise their processes. - *TBT Project Lead*

Senior leader interviewees, in particular, were able to convey the next stage of the plan. For example, one senior leader was proud of the progress their school had made in largely eliminating disruptive behaviour in class, but indicated that their next goal was to address a core number of pupils' apathetic approach to learning. Another outlined their commitment to ensuring new staff were quickly brought up to speed with key reforms and to understand how behaviour management worked within the school:

> Any new staff that come in, it's going to be so important that they're trained up on that and that we run sessions with them…We've tasked deputy heads with the assistant head to be really explicit in devising the next evolution of the behaviour approach and strategy. - *Senior leader, Secondary school*

A key predictor of whether momentum could be maintained appears to be the ability to keep staff (not just teachers) from across the school integrated into future CPD. Usually, TBT Project Leads were confident that changes in school practices and cultures that had resulted in positive gains to behaviour management systems and pupil behaviour would endure. However, there were a limited number of scenarios where they were less assured of this. These included schools where questions remained over the stability of the existing SLT:

> There was a strong sense of nobody knew if the things from this programme would be implemented because a new academy trust was taking over. So, there was a high level of uncertainty there…that was one of the few ones where I thought to myself, we've had impact here, we've had as much impact as we could have… because of the uncertainty. - *TBT Project Lead*

A small number of schools specifically suggested a further check in session with TBT Project Leads around a year after the formal project ended would be a useful enabling strategy for sustaining learning.

6. Evaluation of the TBT project theory of change

In this section, we draw on the findings presented earlier and reflect on the extent to which the project's ToC (Appendix A) was validated. Please refer back to Section 1.1 for a discussion of the theory underpinning anticipated change.

- For the most part, the project delivered activities (inputs) as intended, with the exception of the online platform (see Section 3) that lacked fidelity to the original design but did not adversely affect outcomes, and the Running a Room course for self-referring teachers which did not run (See Section 2.3). Most schools involved in the Running a Room course decided to invite non-teaching staff to attend, therefore there is an argument that the inputs (project activities) should be differentiated in order to maximise the impact for this group of attendees. The project successfully met or exceeded most of the key intended outputs but failed to recruit sufficient participants from target schools (see Section 2).
- Our data strongly suggests that the TBT project had at least a moderate impact on achieving the key ToC outcomes connected to pupil behaviour (see Section 4), but had less of an impact in terms of the retention and progression fund-level outcomes and impacts that were less directly related to pupil behaviour. A number of factors are likely to be responsible for why some outcomes were met more effectively than others. The underlying rationale and evidence in the ToC is very focused on affecting positive pupil behaviour at a school and classroom level and was delivered by a pupil behaviour specialist. Participants and schools are likely to have been drawn to the project for those reasons primarily, as opposed to non-behaviour specific outcomes such as intention to stay within the profession. In the context of limited contact time with participants, it is perhaps unrealistic to expect CPD marketed at pupil behaviour to meaningfully influence an outcome such as retention in the profession, which can be influenced by many other factors. Therefore, it may be necessary to streamline the number of non-behaviour specific outcomes and impacts in the ToC. However, it is important to note that most interviewees were from schools that had already embarked on behaviour reforms and tended to report low instances of highly challenging pupil behaviour. Therefore, although our judgement is that the evaluation largely validates the ToC for TBT (aside from a small number of outcomes not directly related to pupil behaviour), we need to acknowledge that it was predominantly based on schools from a particular context.
- Analysis of data from the SWC data connected to any impact TBT might have had on improved teacher retention and progression was mixed and inconclusive. For example, although not a consistent across all years, there was some evidence to suggest that the TBT project may have had a positive

impact on **teacher level retention** in the state-funded sector two years after baseline and **teacher-level progression** in state funded schools at year three and in the same LAD at years two and three.

7. Learning about effective CPD for schools in challenging circumstances

7.1 Recruiting and engaging schools

The use of social media and free provision of CPD were found to be effective in recruiting schools in challenging circumstances. TBT's significant existing social media presence was considered a '*useful way of getting around the door*' of target schools, requiring only one member of staff or an associate of the school to pick up on the opportunity. This, along with it being free, was considered particularly important for engaging target schools whose managerial and administrative structures were sometimes 'overloaded' responding to immediate crises and lacking the necessary bandwidth to be more outward facing.

A further enabler to recruiting and constructively engaging schools in challenging circumstances was the flexible, no-blame and non-dictatorial approach employed. There was pragmatism to the TBT offer, which did not seek to impose prescriptive 'one-size-fits-all' solutions to behaviour management challenges, acknowledging instead the importance of applying key learning and strategies to schools' own contextual circumstances. In many cases, target schools had been through significant upheaval (and in a minority of cases were still in very unstable circumstances) and, as such, the morale of staff was often extremely low and their confidence fragile. Therefore, the positivity and non-judgemental pitch of the TBT Project Leads was something that particularly resonated with schools:

> People were so happy to finally get some sustained CPD about behaviour management that wasn't 'It's your fault, be better'.
> - *Senior leader, Secondary school*

7.2 Designing effective CPD

Coe (2020) has drawn together an evidence-informed list of practical implications for the design of CPD. These are based on the broad congruence of evidence found in reviews about the characteristics of effective CPD that support changes in teachers' classroom practice which, in turn, are likely to lead to substantive gains in pupil learning. These are set out in Appendix G. The first purpose of this section is to highlight key features of the TBT project which appeared to lead to positive outcomes indicative of effective CPD that align with Coe's list. The second is to identify any key features of the TBT project that appeared to lead to positive outcomes indicative of effective CPD, which are not included in Coe's list.

Alignment of key features of the TBT project with existing evidence on effective CPD[20]

The specific features of the TBT training design that were found to be important in leading to positive outcomes in schools in challenging circumstances, which align with the generic existing evidence on effective CPD across all schools are detailed below.

1. Evidence-based content, sequencing and differentiation

    - Content was founded on a secure evidence base - the majority of which is published in the *Creating a culture: How school leaders can optimise behaviour* report (Bennett, 2017).
    - Content was not a prescriptive or 'one-size-fits-all' approach, but instead contextually relevant and differentiated as appropriate.

2. Duration and frequency

    - CPD was delivered over a two-term period (six months) (but as a low-input training project, frequency was limited).

3. Opportunities for participants

    - Emphasis was on sharing/modelling examples of good practice informed by the deliverers' close contact with current school practice (3c).
    - Practical techniques were used that teachers could directly apply within their classrooms (for example de-escalation techniques) (3d).
    - Opportunities were incorporated into the delivery structure for changes to be trialled in practice and then to be revisited, discussed and potentially adjusted following advice from Project Leads (e.g. Booster days) (3d).

4. Learning and implementation environment

    - Opportunities were provided for collaboration and reflection with peers, sharing challenges and plans and experiences of implementation (4a).
    - School leaders were challenged so that they were ultimately responsible for creating the environment for teachers to flourish (4b).
    - Project Leads strongly advocated that participants created an environment where the expectations of a behaviour management system were clear to all stakeholders (e.g. senior leaders, teaching staff, wider staff, pupils and parents) (4c).

[20] Numbers relate to the section in Coe's list - see Appendix G.

Other key features of the TBT project that appeared important to achieving positive outcomes in schools in challenging circumstances

Two key features of the TBT project that were perceived as significant factors in leading to positive outcomes in schools in challenging circumstances, but are not as strongly emphasised in Coe's list or the current evidence base on effective CPD are:

1. Characteristics of the lead deliverer

    - A Renowned expert in the field directly delivering inputs.
    - The charisma and presentational skill of the lead deliverer.
    - As noted earlier, these features are unlikely to be replicable at scale.

2. Acceptance of school circumstances and pragmatic, empathetic and non-judgemental approach to supporting change that is possible in those circumstances

    - Empathetic approach that treated participants with a professional courtesy and non-judgemental tone.
    - Pragmatic approach that did not become fixated on delivering the project in solely optimal conditions:

        > You work with the people you have in the circumstances you've got, to achieve as many of the outcomes…you're looking for [as possible]
        >
        > There are practical considerations which mean that the non-ideal model is still ideal for that circumstance.

Replicating these features of the TBT project more widely may support more effective CPD for schools in challenging circumstances.

7.3 Summary

Overall, the TBT project offered a set of evidence-informed pupil behaviour management strategies and change management techniques that led to moderate positive outcomes related to pupil behaviour. This aligned with a charismatic and non-judgemental presenting style that pragmatically encouraged participants to apply learning in contextually appropriate ways, and enabled it to contribute to positive outcomes for teachers, schools and pupils in challenging circumstances. However, given TBT's reliance on Tom Bennett to lead on delivery, and participant feedback about the central importance of this, it is unclear whether a significant scale-up of the project in its current form would be viable.

In terms of improvements to the TBT project, positive outcomes may be further enhanced by scheduling an additional Booster roughly a year on from the end of the project to enable participants to share experiences, reflect on implementation, benefit from the expertise of Project Leads to troubleshoot if necessary and plan next steps. More precise initial targeting of 'priority' schools is also necessary to ensure KPI targets for recruitment are fully met.

Other organisations designing CPD for schools in challenging circumstances may wish to consider adopting TBT's pragmatic, empathetic, non-judgemental approach and willingness to deliver the project in sub-optimal conditions that appears to have supported the achievement of positive outcomes in these types of schools.

A final learning point, for organisations commissioning CPD projects for schools in challenging circumstances, is the need to provide sufficiently long lead in times. This enables schools to both incorporate the CPD into their school calendar and allows providers sufficient time to implement strategies to recruit the schools most in need of the CPD and to plan delivery so that schools’ preferences for scheduled delivery dates can be accommodated.

References

Bennett, T. (2017) *Creating a culture: How school leaders can optimise behaviour*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/602487/Tom_Bennett_Independent_Review_of_Behaviour_in_Schools.pdf

Carter, A (2015), Department for Education, Carter review of initial teacher training (ITT), (2016) A framework of core content for initial teacher training (ITT), Developing behaviour management content, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/399957/Carter_Review.pdf

Coe, R. (2020). 'The case for subject-specific CPD', *Paper presented at the Institute of Physics summit to discuss a long-term approach to subject-specific continuing professional development (CPD)*,The Institute of Physics, London, January 2020. Available at: https://www.iop.org/about/publications/the-case-for-subject-specific-cpd (Accessed: 14 June 2022).

Deans for Impact (2016) *Practice with purpose: The emerging science of teacher expertise* https://deansforimpact.org/wp-content/uploads/2016/12/Practice-with-Purpose_FOR-PRINT_113016.pdf

Lemov, Doug. (2011). Teach Like a Champion Field Guide. Hoboken, NJ: Jossey-Bass

Marzano, R., Marzano, J. & Pickering, D. (2003) *Classroom management that works: research-based strategies for every teacher* (ASCD, Alexandria).

NFER, 2013, Teacher Voice Omnibus: June 2013 Survey: Pupil Behaviour

NFER 2016, Teacher Voice Omnibus, September 2016: Pupil Behaviour

Ofsted (2014) *Below the radar: low-level disruption in the country's classroo*ms https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/379249/Below_20the_20radar_20-_20low-level_20disruption_20in_20the_20country_E2_80_99s_20classrooms.pdf

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, 6(4), 342-353.

Appendix A Project theory of change

Rationale and Evidence: The 'Below the Radar,' Ofsted, 2014 and the Teachers Voice Omnibus 2013, 2016 reports outline the scale of the behavioural problem in the UK, and also describe the impact this has on student attainment at multiple levels. The Carter review (2015) called for significant improvements to teacher training in behaviour management. The 2017 report 'Creating a Culture' recommends to "Ensure school leaders have access to training in a range of behavioural strategies and examples of best practice in the school system, by the creation of an optional training scheme. The program will be effective because it applies several proven strategies for classroom and behavioural management (Marzano et al, 2003), as well as proven training techniques (Smith et al, 1978; Deans for Impact, 2016,).

Project activities

Six month rolling programme involving three key inputs from TBT:

- **Two day course**
- **One day booster**
- **Final one day booster**
- **Access to online community portal with resources**

Programme model of delivery:

A) Senior Leader(s) only

B) Teacher (s) only

C) A/B - whole school

Outputs

- 720 participants from 100 schools (primary and secondary). Across delivery model A-C.
- **Spread measure**: Sourced from Level 5 and 6 LA areas. Category 3 or 4 Ofsted schools. Requirement for 70% in priority schools.
- **Retention measure:** Recruit the best candidates and engage them throughout so that 95% of schools participants? complete the project
- **Satisfaction measure:** 80% of participants rate the project as good or above overall (end of project internal survey)

Softer outcomes

Leaders e.g.

- Improved confidence in ability to lead on behaviour management in school
- Changes in leadership practice related to effective behaviour management
- Heightened level of satisfaction
- Improved senior leader perceptions of school and their academic and civic identity

Teachers e.g.:

- Improved confidence in ability to employ effective behaviour management strategies
- Changes in classroom practice related to effective behaviour management
- Effective teacher-student relationships
- Level of satisfaction
- Improved teacher perceptions of school and their academic and civic identity.
- Motivation/likelihood to stay in school/profession

Pupils e.g.:

- Improved pupil wellbeing, behaviour and discipline (e.g. via pupil attendance, exclusions data

Schools e.g.:

- Effective strategies used across the whole school to improve behaviour.
- Valuing of CPD

Impacts: Schools

- Improved school culture.
- Increased engagement in /demand for CPD
- Behaviour management strategies sustained over time
- Improved quality of teaching (underpinned by improved behaviour strategies)

Impacts: System

- Improved academic and civic identity
- Improved learning about how behaviour management programmes can support improvements in schools in challenging circumstances?
- Improved CPD offer (quality/ availability)
- Increased demand for CPD

Impacts:

- Improved leader/teacher progression
- Retention (Teacher/SL)
- Increased pupil attainment
- Improved pupil social mobility via exploring the attainment of pupils eligible for free school meals and reduction in percentage of NEETs

Contextual issues

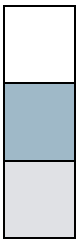
- Specific needs of individual schools and participants (scale of existing issues with BM)
- Prior knowledge, experience, beliefs and attitudes of participants
- School readiness and capacity for change
- Model of delivery signed up to

Appendix B Achieved sample of schools

AEA	Ofsted	Priority*	School phase	School type	FSM	School size	Mode of TBT uptake
3	Null[21] (4)	Medium priority	Primary	Academy sponsor led	1-10%	401-600	Running a school
6	3	High priority	Primary	Community	1-10%	201-400	Running a school
1	2	Lowest priority	Primary	Community	11-20%	601-800	Running a school
N/A	N/A	Medium priority (NQT)	Primary	N/A	N/A	N/A	Running a room
5	3	High priority	Primary	Community	11-20%	201-400	Running a school
6	N/A	Medium case	Primary	Free school	N/A	N/A	Running a school
4	2	Medium priority	Secondary	Academy converter	11-20%	1401-1600	Running a school
6	2	Medium priority	Secondary	Voluntary Aided	11-20%	601-800	Running a school
3	1	Lowest priority	All through	Academy sponsor led	41+%	1601-1800	Running a school
5	4	High priority	Secondary	Academy converter	31-40%	601-800	Running a school
5	4	High priority	Secondary	Community	11-20%	801-1000	Running a school
6	2	Medium priority	Secondary and 16 to 18	Academy converter	11-20%	1801+	Running a school
6	3	High Priority	Secondary	Academy	21-25%	601-800	Running a school

[21] The last Ofsted rating of the school that became a new academy sponsor-led school was 4 (and therefore was used as an indicator for sampling purposes). However, it is important to state at the time of signing up for TBT, the new academy sponsor-led school had never received an Ofsted rating.

AEA	Ofsted	Priority*	School phase	School type	FSM	School size	Mode of TBT uptake
6	2	Medium priority	Secondary	Academy converter	11-20%	601-800	Running a room
6	4	High priority	Secondary	Academy sponsor led	31-40%	1201-1400	Running a room



Single telephone interview

Telephone case study

School case study visit

* See Appendix C for description of priority schools

Appendix C TBT qualitative sampling: key principles

Sampling for case studies and telephone interviews undertaken at the school level.

1. The process was conducted over time to take account of rolling recruitment. At each sampling point, only schools that completed all intended face-to-face inputs (ideally having had at least one half-term since the last TBT input to offer greater opportunity for implementation) considered for sampling.

2. Over-sampling of high and medium priority schools (see high, medium and low categories definition below). Some low priority schools included to maximise learning re: CPD in the most challenging circumstances. Priority[22] has been determined as follows:

    - Highest priority: Ofsted category 3-4 & AEA 5-6.
    - Medium priority: Either Ofsted category 3-4 or AEA 5-6.
    - Least priority: Ofsted category 1-2 & AEA 1-4.

3. A mix of purposive and random sampling of schools undertaken to create a sample that:

    - includes only schools with at least one senior leader that attended every TBT session run.
    - is balanced as above (point 2) re: high/medium/low priority schools.
    - ensures that the different modes of delivery are represented (with more schools selected from the most frequently deployed mode/s of delivery) (project-level factor).
    - includes a balance across primary and secondary that reflects the pattern of recruitment (school-level factor).
    - includes a geographical spread across the TLIF target area (area-level factor).
    - does not include more than one school from the same MAT.
    - no repeat interviews or case studies.

4. In addition, as far as is possible, the sample will ensure some variation in the following other school-level factors:

[22] While the only contractual KPI was 70% priority schools, the intention (as shown in the ToC - Appendix A) was to source schools from AEA 5/6 areas.

- school-type: maintained/ non-maintained.
- Ever6[23] FSM entitlement.

[23] EVER6 FSM entitlement ensures that schools receive Pupil Premium funding for any pupil that has been eligible for FSM at any point during the last 6 years

Appendix D SWC matching and comparison group construction

Data sources

The main data source used for the retention and progression analysis was the School Workforce Census (SWC). The SWC has been collected annually on the first Thursday of November since 2010 and it observes teaching staff and their characteristics from all state-sector schools in England. The key teacher characteristics recorded in the SWC and used for the analysis comprised gender, age, qualification date and role, while key school characteristics comprised school phase, type and region.

Each teacher in the SWC is assigned a unique identifier, which enables analysis of the same individual over multiple censuses. This allows observation of key pieces of information about teachers' careers, such as whether they leave state-sector teaching, move school/ area, or progress into a more senior role.

The SWC records the school in which each teacher is employed, meaning it is also possible to identify teachers who move to different schools, LAs and regions.[24] However, since the SWC does not include teachers in private sector schools or schools outside of England, any teachers who move to one of those schools will appear to have left teaching, even though, in reality, they may not have.

The data quality and response rates to the SWC are very high, so the data has good coverage and few gaps. However, it has some gaps due to schools not submitting returns or individual teachers missing from submitted returns, so to minimise the influence of errors and data gaps, and improve the reliability of the retention outcomes, records were imputed where gaps or errors were evident.[25] While this is unlikely to have completely eliminated all instances of SWC data gaps it is unlikely to affect the interpretation of the findings as they are very likely to affect treatment teachers/ schools in a similar way to comparison teachers/ schools.

[24] Teachers may have contracts in multiple schools, but the file that we used for this evaluation contains one record per teacher per year of the 'main school' that a teacher is working in. The school changes that we observe are therefore changes in the 'main school', as recorded in the SWC.
[25] Cases where data gaps are obvious include the observations in which a teacher is not recorded in a school in a year after which the SWC records them as having started in a particular role. For example, if the SWC shows a particular teacher is working in a school in the 2017 census year and they are recorded as having started in their current role in the 2016 census year, where they have no SWC record, then the missing record for 2016 is imputed. In these cases, it is assumed they were teaching in the same school as in 2017, and their time-variant characteristics are imputed as appropriate (reducing their observed age, experience, etc. by one year). School-level characteristics and teacher-level characteristics that do not vary by time (i.e. gender, ethnicity), are set to their observed value in 2017. This imputation affects relatively few records and does not apply to any records in which role start date is not observed.

In addition to the teacher-level variables, school-level data was used for the analysis including region, phase, Ofsted rating and Achieving Excellence Area (AEA) category, all data which is published by the DfE.[26]

The final data source consisted of the management information (MI) data collected by the TLIF providers on the teachers participating in each project, and collated by DfE. The MI data observes teachers' personal details, participation in TLIF projects, along with the provider, the name of the school in which the teacher participated in the training and, for some projects, the training start and end dates.

Each teacher in the MI data was linked to their SWC records using their name, Teacher Reference Number (TRN) and birth date. Across all TLIF projects, 97 per cent of teachers in the MI data were matched to at least one record in the SWC. Match rates varied somewhat across the different projects, although were generally very good, even after accounting for teachers in the MI data who linked to multiple teachers in the SWC, or did not link to an SWC record in the year in which they were recruited to the project.[27]

Table 21 shows that the match rate for teachers listed in the MI data as participating in the TBT project was 78 per cent to an SWC record in the year in which, according to the MI data, they were recruited to the project.

Table 21: Matching teachers to the SWC

Stage of matching	Frequency of teachers
Total TBT participants identified in the MI data	754
Total TBT participants matched to at least one SWC record	613
Total TBT participants matched to an SWC record after removing SWC inconsistencies and records with missing baseline information	587
Match rate (%)	78

Table 22 shows that the match rate for schools in the MI data as participating in the TBT project was 92 per cent.

[26] The latest data is available here: https://www.get-information-schools.service.gov.uk/
[27] Cases such as these where the match was clearly wrong were removed from the analysis.

Table 22: Matching schools to the SWC

Stage of matching	Frequency of schools
Total TBT schools identified in the MI data	75
Total TBT schools with at least one participant who matches to an SWC record	69
Match rate (%)	92

Methodology

Each of the methodological steps in the analysis were performed separately for evaluating the project effects at the individual teacher and the whole-school level. After linking the MI data to the SWC, the group of comparison schools/teachers was derived whose retention and progression outcomes were compared to TBT-participating schools/teachers.

For each treatment and comparison teacher/school, a baseline year was defined, relative to which subsequent retention and progression outcomes were observed. For TBT participant teachers, this was defined as the year in which the teacher was recruited to the project. For any teachers with multiple observed recruitment dates, the first observed date was used as baseline. For schools, the baseline year was defined as the most common recruitment year for participant teachers in that school. For example, if the majority of teachers in a particular school were recruited to the project in 2017, then 2017 was assigned as the baseline year for that school.

With this full set of potential comparator teachers/schools, a statistical technique called *propensity score matching* was used to ensure that the treatment and comparison groups were highly comparable in observable characteristics. This was done similarly but separately for teachers and schools. For teachers, the probability (propensity score) that a particular teacher with given characteristics was part of the treatment group was estimated. TBT participant teachers were then matched with up to ten of their 'nearest neighbours' – comparison teachers with the most-similar likelihood of being in the treatment group, and therefore with the most similar observed characteristics. For schools, the propensity score was estimated with the observed characteristics of the school, rather than individual teachers.

When propensity score matching is able to match on all of the variables that influence selection into the treatment group then the only remaining difference between the treatment and matched comparison group is the effect participating in the project had. However, variables can only be included in the matching if they are observed in the data. If other unobserved variables influence selection into the

treatment group, and also affect retention, then this may partially explain some of the differences in outcomes between the two groups. The potential for this 'selection bias' means caution should be exercised about interpreting the differences between the groups as only representing the causal impact of the project.

The characteristics we used for matching differed between the teacher and school-level analyses. At the teacher level, both teacher and school characteristics (observed at the baseline year) were used as variables in the matching. The teacher characteristics included age, gender, years since qualification,[28] full-time/part-time status, post and baseline year. The school characteristics used for matching included Ofsted rating, phase, quintile of free school meal (FSM) eligibility, quintile of attainment,[29] AEA and region.

At the school level, the following school characteristics (observed at the baseline year) were used as variables in the matching: school phase, Ofsted rating, quintile of free school meal (FSM) eligibility, quintile of attainment,[30] pre-baseline year retention rates and an indicator of whether the school was participating in any other TLIF projects.

The quality of the match was assessed by examining cross-tabulations of the matching variables across the treatment and comparison groups. Where the variables are balanced – meaning the distribution of characteristics is similar between the treatment and comparison groups – the propensity score matching can be said to have performed well (see Tables 23 and 24 for the matching output).

As all of the outcome variables are dichotomous (i.e. yes or no), the differences in retention and progression outcomes between the two groups were estimated using logistic regression modelling. Retention and progression are considered separately from four different perspectives:

1. Within the same school one, two and three years after baseline[31]

2. Within the same LAD one, two and three years after baseline

[28] We used years since qualification as a stand-in for experience as the variable observing year of entry into the profession (which was used to calculate years of experience) had a substantial amount of missing observations.
[29] Attainment was measured as the proportion of pupils in the school that met the minimum requirements in Reading, Maths and Science at Key Stage 2 (for primary schools) or GCSEs (for secondary schools). Schools were assigned to an attainment quintile based on this proportion.
[30] Attainment was measured as the proportion of pupils in the school that met the minimum requirements in Reading, Maths and Science at Key Stage 2 (for primary schools) or GCSEs (for secondary schools). Schools were assigned to an attainment quintile based on this proportion.
[31] While in principle outcomes three years after baseline were observed, sample sizes at this stage are too small to be statistically reliable, so only outcomes one and two years after baseline are reported.

3. Within the profession as a whole one, two and three years after baseline

4. Within a ‘challenging’ school one, two and three years after baseline.

A teacher was considered to have been ‘retained’ in the same school/LAD if they were teaching in a particular school/LAD in a given year, and were then recorded as teaching in the same school/LAD (based on URN and LAD codes) one, two or three years later. Similarly, a teacher was considered to have been ‘retained’ in the profession if they were recorded as teaching in a state-sector school in England in a given year, and then were also teaching in a state-sector school in England one, two or three years later.[32]

‘Challenging schools’ were generally defined as schools that were rated by Ofsted as ‘requires improvement’ or ‘inadequate’. However, it was also assumed that all TBT participant teachers were teaching in a ‘challenging school’ when they were recruited to the project at baseline, even for the relatively few teachers that were in a ‘good’ or ‘outstanding’ school (see observed characteristics in the matched sample - Table 23). This is because the school had been deemed challenging enough to be targeted by the TBT project, despite having been rated as ‘good’ or ‘outstanding’ by Ofsted in its last inspection.

Retention in a challenging school was defined at the teacher level. That is, a TBT participant teacher was considered as having been retained in a ‘challenging school’ if they either stayed in the same school they were in at baseline, or had moved to a different school which was rated ‘requires improvement’ or ‘inadequate’ in the year they moved. It should be noted that this same definition also applies to comparison teachers (including those in ‘good’ or ‘outstanding’ schools not targeted by the TBT project), but the results of the statistical matching (see Table 23) ensure that the observed characteristics of the ‘good’ and ‘outstanding’ schools in the comparison group are similar to the observed characteristics of the ‘good’ and ‘outstanding’ schools within the treatment group.

As a concrete example, a TBT teacher in a ‘good’ school who stayed in the same school, or a non-TBT teacher in a ‘requires improvement’ school who moved to an ‘inadequate’ school would both be considered to have been ‘retained in a challenging school’. Similarly, any teachers who moved to another school with a ‘good’ or ‘outstanding’ rating were considered to have moved to a ‘non-challenging’ school, regardless of the rating of the school they were in at baseline.

[32] To reiterate, since the SWC only observes teachers in state-sector schools in England, any teacher who moves to a private school or to a school outside of England will be considered to have left the profession.

Progression was defined according to three broad role categories – classroom teachers, middle leaders, and senior leaders. Middle leaders were defined as teachers in a "Leading Practitioner", "Excellent Teacher", "Advanced Skills Teacher", or "Advisory Teacher" post, or who received a Teacher Leadership Responsibility (TLR) payment of £100 or more in a given year.[33] Senior leaders were defined by those in an "Executive Head Teacher", "Head Teacher", "Deputy Head Teacher" or "Assistant Head Teacher" role in a given year.

A teacher was considered to have 'progressed' if they moved from a classroom teacher role to either a middle or senior leadership role, or a middle leadership role to a senior leadership role one, two or three years after baseline. Progression within a school/LAD/challenging school is defined as those teachers who remain within the same school/LAD/a challenging school and progressed from classroom teacher to middle leadership or middle leadership to senior leadership.

Eight different regression models were estimated, one each for retention and progression within the same school/the same LAD/challenging schools/the profession. This was done using separate regression models for the teacher-level and the school-level analysis.

For the teacher-level analysis, a logistic regression model was used to estimate the likelihood of retention/progression in each of the eight models. As independent variables, all of the variables from the propensity score matching were included – in order to control for any remaining imbalances in the matching variables between the treatment and comparison groups after matching – as well as the treatment indicator and year dummy variables to account for specific time period effects (e.g. the impact of Covid-19 on the 2020 data). Senior leaders were excluded from the sample estimating the effect on progression as, based on the definition above, they are not able to progress any further and therefore progression outcomes are 'did not progress further' by definition.

To compare the differences between the two groups, the probability of 'retention' or 'progression' was estimated if every teacher had been involved in the project, and then again if every teacher had not been involved in the project. The average of these predicted probabilities is the average estimated retention/progression rate for treatment and comparison teachers, respectively. The difference between treatment and comparison teachers is the estimated 'marginal effect', which is presented in the tables in section 4, with the accompanying odds ratio estimates in Appendix E. Standard errors for the marginal effect estimates are calculated using the delta method and statistical significance is assessed at the five per cent level.

[33] This is a definition of middle leader that has been used by DfE in the past. See Footnote 14 in https://www.gov.uk/government/statistics/teachers-analysis-compendium-2017

For the school-level analysis, the models were estimated using teacher-level data in a logistic mixed-effects regression model. As independent variables, all of the variables from the propensity score matching, as well as the treatment indicator, census year and an interaction between these variables were included. School was included as a random effect.

To compare the differences between the two groups, the model estimated the probability that each teacher in the matched sample would have been 'retained' or 'progressed' if they had been involved in the project, and then again if they had not been involved in the project, in each of the five census years. The average of these predicted probabilities was then taken to find the estimated retention/progression rate, with and without the treatment. The difference between these estimated retention/progression rates is the estimated 'marginal effect', which is presented in the tables in Section 4. The difference-in-difference testing was then performed to compare the difference between treatment and comparison, between pre-baseline and each post-baseline year. For each post-baseline year, the treatment vs. comparison difference was compared to an average of the pre-baseline differences. The same difference-in-difference estimates are also presented as odds ratios in Appendix E. Statistical significance is assessed at the five per cent level.

Statistical Matching

Table 23 below highlights the sample characteristics for the full treatment and comparison groups for the teacher-level analysis. In the unmatched samples, treatment teachers were slightly more likely to be male, younger and less experienced than in the unmatched potential comparison group. Similarly, the schools that treatment teachers were in were more likely to be rated 'requires improvement' or 'inadequate', had lower attainment, higher proportions of pupils eligible for free school meals, and were more likely to be an AEA category 5 or 6 school than in the unmatched comparison group.

After matching, the proportions of comparison teachers in each of the key matching characteristics were much more closely aligned with treatment teachers. While some small differences between treatment and comparison teachers still existed after matching, including the matching variables as covariates in the logistic regression modelling ensured that the final estimates controlled for any of these outstanding differences.

Focussing on the subset of potential comparison teachers who were the most similar to treatment teachers necessarily involved discarding some potential comparison teachers from the matched sample, when there were no sufficiently similar treatment teachers with which to match. Of the 1,380,905 potential comparison teachers, only 3,930 were matched to a treatment teacher, highlighting how potential comparison

teachers were still fairly dissimilar to teachers recruited to the TBT project (at least in observed teacher and school characteristics. 32 potential treatment teachers were also discarded from the matched sample, as these teachers have no sufficiently similar counterpart in the potential comparison teacher sample.

Table 23: Characteristics of treatment and comparison teachers before and after matching in the full sample

Characteristic	Treatment teachers (%)	Potential comparison teachers (%)	Matched treatment teachers (%)	Matched comparison teachers (%)
Male	37.5	24.2	36.6	37.6
Female	62.5	75.8	63.4	62.4
Aged under 30	24.4	22.8	24.5	25.2
Aged 30-49	59.6	59.7	59.1	59.3
Aged 50 or older	16.0	17.5	16.4	15.5
Within 5 years of qualifying	29.5	24.3	29.5	28.5
Between 5 and 9 years since qualifying	17.7	20.1	16.2	18.4
Between 10 and 19 since qualifying	31.0	30.4	31.7	30.2
20 years or more since qualifying	14.3	21.1	14.8	14.2
Unknown years since qualification	7.5	4.1	7.7	8.7
Classroom teacher	56.9	68.3	57.5	59.0
Middle leader	23.0	18.3	21.8	21.6
Senior leader	20.1	13.5	20.7	19.3
Full-time	88.6	76.3	88.1	88.8
Part-time	11.4	23.7	11.9	11.2
Ofsted outstanding	2.6	22.0	2.7	2.5
Ofsted good	38.2	61.0	40.4	39.8
Ofsted requires improvement	27.4	10.9	29.0	27.9

Characteristic	Treatment teachers (%)	Potential comparison teachers (%)	Matched treatment teachers (%)	Matched comparison teachers (%)
Ofsted inadequate	26.2	3.4	22.2	23.5
Ofsted rating unknown	5.6	2.8	5.8	6.3
Primary school	21.6	53.2	22.9	21.0
Secondary school	78.4	46.8	77.1	79.0
AEA category 1	4.1	18.0	4.3	4.3
AEA categories 2-4	13.1	50.7	13.9	13.8
AEA category 5	26.2	15.9	22.7	23.2
AEA category 6	56.6	15.4	59.1	58.8
FSM highest 20%	40.9	19.5	43.1	39.1
FSM middle-highest 20%	19.6	20.2	20.2	19.8
FSM middle 20%	30.0	20.1	26.7	32.2
FSM middle-lowest 20%	< 9.0*	19.9	< 9.0*	7.1
FSM lowest 20%	< 2.0*	20.3	< 2.0*	1.7
Attainment highest 20%	2.2	20.9	2.3	2.3
Attainment middle-highest 20%	8.5	23.5	9.0	8.1
Attainment middle 20%	14.7	21.8	15.5	17.8
Attainment middle-lowest 20%	49.6	19.4	51.9	48.8
Attainment lowest 20%	10.9	7.7	11.5	11.1
Attainment unknown	14.1	6.7	9.7	11.8
East of England	24.4	11.4	25.6	28.0
East/West Midlands	3.4	19.3	3.6	3.3
London	11.1	16.5	11.7	10.9
North East	0.0	4.7	0.0	0.0

Characteristic	Treatment teachers (%)	Potential comparison teachers (%)	Matched treatment teachers (%)	Matched comparison teachers (%)
North West	0.0	13.0	0.0	0.0
South East	45.7	16.0	47.7	46.5
South West	15.5	9.6	11.4	11.3
Yorkshire and the Humber	0.0	9.6	0.0	0.0
Baseline year 2017	50.3	33.3	53.0	49.8
Baseline year 2018	43.3	33.2	40.2	42.3
Baseline year 2019	6.5	33.5	6.8	7.9
Number of teachers	**587**	**1380905**	**555**	**3930**

Note: * indicates proportion has been rounded due to small sample sizes.

In addition to the full matched sample, a second matched sample was derived, with which to estimate the differences in career progression and retention within the same school/same LA/a challenging school. This sample was only used for the teacher-level analysis and not the school-level analysis. Given that career progression or retention within the same school/same LA/a challenging school for teachers who left the profession is not observed for teachers who leave the profession, this additional matched sample consisted of a subset of teachers in the full sample who did not leave the profession in the three years after baseline. Characteristics of teachers in the matched sample of non-leavers were very similar to the full matched sample.

Table 24 below highlights the sample characteristics for the treatment and comparison groups for the school-level analysis. In the unmatched samples, treatment schools were more likely to be rated ‘requires improvement’ or ‘inadequate’ and had lower attainment and higher proportions of pupils eligible for free school meals.

After matching, the proportions of comparison schools in each of the key matching characteristics were much more closely aligned with treatment schools. While some small differences between treatment and comparison schools still existed after matching, including the matching variables as covariates in the logistic regression modelling ensured that the final estimates controlled for any of these outstanding differences.

Table 24: Characteristics of treatment and comparison schools before and after matching

Characteristic	Potential comparison schools (%)	Treatment schools (%)	Matched comparison schools (%)
Nursery	2	0	0
Primary	77	40	43
Secondary	15	60	54
16 Plus	0	0	0
Special	6	0	2
East of England	12	50	47
East Midlands	9	0	4
West Midlands	11	0	0
Inner London	5	0	3
Outer London	7	0	4
North East	5	0	0
North West	15	0	0
South East	15	20	22
South West	11	20	19
Yorkshire and the Humber	10	0	0
AEA Category 1	15	0	3
AEA Category 2	15	10	14
AEA Category 3	17	0	4
AEA Category 4	19	0	6
AEA Category 5	17	30	28
AEA Category 6	17	50	45
FSM lowest 20%	19	0	4
FSM middle-lowest 20%	18	20	19
FSM middle 20%	18	30	26
FSM middle-highest 20%	18	30	26
FSM highest 20%	18	20	22
FSM unknown	9	0	3
Attainment lowest 20%	15	10	17

Characteristic	Potential comparison schools (%)	Treatment schools (%)	Matched comparison schools (%)
Attainment middle-lowest 20%	17	30	24
Attainment middle 20%	18	20	19
Attainment middle-highest 20%	18	20	23
Attainment highest 20%	16	0	5
Attainment unknown	16	10	13
Ofsted Inadequate	3	10	10
Ofsted Requires Improvement	10	30	32
Ofsted Good	64	30	43
Ofsted Outstanding	20	10	8
Ofsted Unknown	3	10	7
Number of schools	**21675**	**69**	**567**
Number of teachers	**500032**	**6685**	**24847**

Note: Matching was performed at a school level so these percentages are also at a school level e.g. 10 per cent of schools not 10 per cent of teachers. Comparison school percentages are rounded to the nearest 1 per cent. Treatment school percentages are rounded to the nearest 10 per cent. The rounding is to ensure data are not disclosive.

Appendix E Outcomes of SWC impact analysis

Table 25: School level: Odds ratios from the retention and progression outcome analysis

	1 year after baseline	2 years after baseline	3 years after baseline
Retention in state-sector teaching	1.9 (1.4 - 2.8)	1.3 (1 - 1.8)	1.3 (0.9 - 1.8)
Retention in the same school	2.3 (1.6 - 3.5)	1.3 (1 - 1.7)	1.3 (0.9 - 1.9)
Retention in the same LA	2.3 (1.5 - 3.7)	1.2 (0.9 - 1.6)	1.1 (0.8 - 1.7)
Retention in a challenging school	2.6 (1.7 - 4.3)	1.3 (0.9 - 1.7)	1.0 (0.7 - 1.5)
Progression in state-sector teaching	1.3 (0.9 - 2)	1.3 (0.9 - 1.9)	2.3 (1.4 - 3.8)
Progression in the same school	1.5 (0.9 - 2.4)	1.5 (0.9 - 2.2)	1.8 (1.0 - 3.4)
Progression in the same LA	1.6 (1.0 - 2.4)	1.6 (1.1 - 2.4)	2.3 (1.3 – 4.0)
Progression in a challenging school	1.3 (0.8 - 2.1)	1.3 (0.9 – 2.0)	1.8 (0.9 - 3.2)

Note: Figures in brackets represent the 95 per cent confidence interval of the odds ratio estimate.

Table 26: Teacher level: Odds ratios from the retention and progression outcome analysis

	1 year after baseline	2 years after baseline	3 years after baseline
Retention in state-sector teaching	1.1 (1.0, 1.3)	1.0 (0.9, 1.1)	0.8 (0.7, 1.0)
Retention in the same school	0.9 (0.8, 1.0)	1.0 (0.8, 1.1)	1.0 (0.9, 1.3)
Retention in the same LA	0.8 (0.7, 1.0)	1.0 (0.8, 1.2)	1.0 (0.8, 1.2)

	1 year after baseline	2 years after baseline	3 years after baseline
Retention in a challenging school	0.9 (0.7, 1.0)	0.9 (0.7, 1.1)	0.9 (0.7, 1.1)
Progression in state-sector teaching	1.0 (0.8, 1.1)	0.8 (0.7, 1.0)	1.0 (0.8, 1.3)
Progression in the same school	0.9 (0.8, 1.2)	0.8 (0.7, 1.0)	1.0 (0.7, 1.2)
Progression in the same LA	1.0 (0.8, 1.2)	0.9 (0.7, 1.1)	1.0 (0.8, 1.2)
Progression in a challenging school	0.9 (0.8, 1.2)	0.8 (0.7, 1.0)	0.9 (0.7, 1.2)

Note: Figures in brackets represent the 95 per cent confidence interval of the odds ratio estimate.

Appendix F Analysis of Management Information for the Teaching and Leadership Innovation Fund: Tom Bennett Training

Introduction

The Teaching and Leadership Innovation Fund (TLIF) was a DfE fund through which 10 providers offered support to schools in a variety of areas from behaviour management to phonics and STEM teaching. The aim of the fund was to create and develop a sustainable market for high-quality Continuous Professional Development (CPD). This is a summary of Management Information (MI) data submitted by all ten providers receiving TLIF funding and **does not** assess project impact. The data was submitted in February 2020 and covers the schools and participants recruited, as indicated by the providers. Comparable national figures in this report are based on the 2018 School Workforce Census covering teaching staff in state-funded schools, and Ofsted as at the most recent inspection. The 2018 School Workforce Census was chosen in order to align with the most schools across programme cohorts between 2017 and 2020. The school level analysis refers to all schools that were recruited by providers to participate in the project, including those that withdrew. Schools may have been recruited by more than one provider and participants may have been registered for more than one project.

Targets: Background

Each provider had a number of Key Performance Indicators (KPIs). These were broken down into three different categories:

- **geography**: whether specific areas were targeted by providers (e.g. regional targets, Opportunity Areas, priority areas) and whether particular schools should be targeted by providers (e.g. based on Ofsted rating)
- **schools**: the minimum number of schools
- **participants**: the minimum target number of participants

All providers had a geography target and either a participant or a school target, but not necessarily both.

In the context of the TLIF evaluation, a priority area is defined as Achieving Excellence Areas (AEAs) 5 or 6 (Opportunity Areas fall within this category), and a priority school is defined as a school with an Ofsted rating of Requires improvement (Ofsted grade 3) Or Inadequate (Ofsted grade 4).

Note: there are some discrepancies between the overall numbers from providers and those in the data set sent to us. The provider numbers cannot be broken down in school/area type etc. so analysis will not be conducted on this data, however headline figures will be presented where available.

Targets: Breakdown

Tom Bennett Training delivered behaviour management training for teachers and school leaders. The aim of the project was to improve capability to deal with challenging behaviour.

Tom Bennett Training had the following KPI targets:

Geography Level:

- A minimum of 70% of participants were to come from priority schools.
- Recruitment was targeted at Southern England.

School Level:

- The programme aimed to recruit a minimum of 100 schools
- The project recruited from both primary and secondary schools

Participant Level:

- A minimum of 720 participants were to be recruited during the programme
- Participants were to be recruited from all teaching and leadership levels

Note: During the first year, Tom Bennett Training changed the approach to recruit larger groups of participants from schools. This resulted in more participants being recruited and fewer schools. This approach was agreed by DfE with a target of 75 schools.

Total school numbers

A total of 75 schools participated over three cohorts. However, removing schools where participants withdrew reduces this to a total of 72 schools.

Note: The initial target was 100, however a revised target of 75 was agreed by DfE.

Total participant numbers

- The total number of teachers that participated in the course was 756.
- 109 teachers withdrew, leaving 647 participants who completed the course.
- The target number of participants was 720.

Note: Tom Bennett Training's own data puts the number of participants at 720, which would be exactly on target.

- Of all participants (including those who withdrew) 59% were from priority schools.

- Of the participants that completed the programme 68% were from priority schools.
- The target was for 70% to be from priority schools

Note: 2 schools had no Ofsted rating data and were excluded from this analysis.

Schools by Phase

Tom Bennett Training recruited from primary, secondary, and special schools

Of all Tom Bennett Training participant schools (including withdrawals):

- 39% of schools were primary schools,
- 59% were secondary schools,
- 3% were special schools.

Compared to the national distribution of schools, secondaries are over-represented in the Tom Bennett Training participant population.

Schools by Region

Tom Bennett Training recruited from schools in five of the eight RSC Regions.

The programme was aimed at "Southern England", which covers most of the regions recruited from, with:

- 47% of schools located in the East of England and North East London,
- 23% in South East and South London,
- 19% in the South West,
- and 9% in South Central and North West London.

The arguable exception is the East Midlands and the Humber region, where around 3% of recruited schools were located.

Schools by AEA Category

AEA categories are DfE classifications of Local Authority Districts (LADs) by educational performance and capacity to improve, introduced in 2016. It splits areas into six categories from "Strong" Category 1 areas to "Weak" Category 6 areas.

Tom Bennett Training recruited schools from all AEA Categories, with a very clear focus on Category 5 and 6 areas.

Of all schools recruited (including withdrawals), 77% were in Category 5 and 6 areas

Schools by Index of Multiple Deprivation Decile

The Index of Multiple Deprivation (IMD) is a "neighbourhood" measure of deprivation produced by the Ministry of Housing, Communities and Local Government. Each neighbourhood is placed into a decile with decile 1 containing the most deprived areas and decile 10 containing the least deprived.

Tom Bennett Training recruited across all of the areas, but just over half (55%) of all participating schools were recruited from deciles 1-5, covering the most deprived areas.

Participants by role

Roles were provided in TLIF Management Information as free text and matched to a standardised leadership level. Below these have been compared to national figures taken from the 2018 School Workforce Census Publication.

Tom Bennett Training recruited participants from all teaching and leadership levels, with almost half (48%) being classroom teachers.

Senior leadership roles are over-represented, with 17% of participants being senior leaders, compared to 10% nationally.

A further 15% of participants were middle leaders (compared to 28% nationally), and 3% were headteachers (compared to 5% nationally).

Although not covered in the national data, around 9% of participants were Non-Teaching Staff and 6% were Teaching Assistants.

Appendix G Practical summary of the evidence about effective CPD (Coe, 2020)

CPD that aims to support the kinds of changes in teachers' classroom practice that are likely to lead to substantive gains in pupil learning should:

1. Focus on promoting the teacher skills, knowledge and behaviours that are best evidenced as determining pupil learning. Such content should be appropriately sequenced and differentiated to match the needs of participants.
2. Have sufficient duration (two terms) and frequency (fortnightly) to enable changes to be embedded.
3. Give participants opportunities to:
    a) be presented with new ideas, knowledge, research evidence and practices
    b) reflect on and discuss that input in ways that surface and challenge their existing beliefs, theories and practices
    c) see examples of new practices/materials/ideas modelled by experts
    d) experiment with guided changes in their practice that are consistent with these challenging new ideas and their own context
    e) receive feedback and coaching from experts in those practices, on an ongoing basis
    f) evaluate, review and regulate their own learning
4. Create/require an environment where:
    a) participants can collaborate with their peers to support, challenge and explore
    b) school leadership promotes a culture of trust and continuous professional learning
    c) teachers believe they can and need to be better than they are
    d) the process and aims of the CPD are aligned with the wider context (e.g. accountability)

Source: Coe, R., (2020) The Case for Subject-Specific CPD. Report for the Institute of Physics January 2020

Department for Education

Reference: DFE- RR1259

ISBN: 978-1-83870-401-8

For any enquiries regarding this publication, contact us at:
www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications