Talk With Tales for Children (TWiTCH) programme

Pilot Evaluation Report

July 2025

Nadine Fitzfussell, Vic Menzies, Paivi Eerola, Johny Daniel, Kelly Burgoyne, Xiaofei Qi, Helen Cramman.

Co-funded by:

St Edmund's Early Years Stronger Practice Hub, Yorkshire and Humber

The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, nurseries, and colleges to improve teaching and learning for 2- to 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence**—reviewing the best available evidence on teaching and learning and presenting it in an accessible way;
- **Finding new evidence**—funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socioeconomically disadvantaged backgrounds; and
- **Putting evidence to use**—supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were re-endowed with an additional £137m, allowing us to continue our work until at least 2032.

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

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

0207 802 1653

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The project was independently evaluated by a team from Durham University: Victoria Menzies, Helen Cramman, Johny Daniel, Paivi Eerola, Nadine Fitzpatrick, and Xiaofei Qi assisted by Kelly Burgoyne at the University of Manchester.

The lead evaluator was Victoria Menzies.

Contact details

Victoria Menzies
School of Education, Durham University
Confluence Building
Stockton Road
Durham, DH1 3LE

Email: victoria.menzies@durham.ac.uk

Executive summary

The project

The Talking With Tales for Children (TWiTCH) programme is a structured continuous professional development (CPD) programme for nursery practitioners aiming to support their use of dialogic pedagogy in their practice. Specifically, the programme aims to enhance complex language interactions between children and practitioners during daily story-time and related activities to improve preschool children's language and communication skills. TWiTCH is targeted at settings in areas of disadvantage where there is greatest need to improve children's language skills.

Practitioners deliver the TWiTCH programme to three- and four-year-old children in their setting through nine three-week story cycles. Each cycle is based around a storybook of a traditional fairy tale likely to be familiar to practitioners and children—such as Little Red Riding Hood. Sessions are designed to be delivered to all children in small groups of five to eight children with each session taking around 15 minutes. In Week 1 of each cycle, practitioners read the storybook to children daily and use specified dialogic reading prompts to stimulate discussion about different aspects of the story. In Week 2 of each cycle, practitioners deliver three sessions of story-based, choice-based language activities to promote reasoning language. In Week 3, continuous provision activities based on the storybook are provided in the setting, allowing children to engage further with the story content. Practitioners also provide targeted additional support to specific children identified as being 'in need'. Beyond the nine cycles, practitioners are encouraged and supported to choose their own book for a tenth TWiTCH cycle. The TWiTCH programme uses a coach training model to provide nursery settings with on-site training and ongoing coach support. A TWiTCH champion in each setting coordinates and supports delivery. Champion network meetings also encourage champions to share practice.

The pilot evaluation used a mixed-methods approach based around a contribution analysis framework to evaluate the acceptability of the TWiTCH programme, implementation fidelity across early years settings, evidence of promise, and the programme's readiness for trial. The focus for the evaluation was on practitioner-level outcomes.

Twenty-five settings across West and South Yorkshire (12 maintained and 13 private, voluntary, and independent, 'PVI') were recruited to take part in the pilot evaluation. The evaluation collected quantitative data through a practitioner survey and audio recordings of storytelling practice conducted before and after TWiTCH delivery to look at changes in practitioner confidence, knowledge, and behaviour. TWiTCH champion practitioner and developer team interviews as well as observations and interviews with setting staff (practitioners, champions, and SLT) in five case study settings provided data about views of the programme, TWiTCH delivery, and perceived impact. A post-training practitioner survey, observations of training, and administrative data were also analysed as part of the evaluation.

TWiTCH delivery and evaluation took place between May 2023 and July 2024. The TWiTCH programme was developed and delivered by a team at Sheffield Hallam University (SHU) led by Fufy Demissie. As part of the Department for Education's Early Years Recovery Programme, the Education Endowment Foundation (EEF) is working with Stronger Practice Hubs across England to fund early years settings' access to evidence-informed programmes and study the programme's influence on practice and children's outcomes. This initiative aims to support education recovery following the pandemic while also developing our understanding of effective professional development in the early years. The EEF has worked with St Edmund's Early Years Stronger Practice Hub to fund settings' access to TWiTCH and evaluate the programme through a pilot study.

Key findings

Table 1: Summary of pilot findings

Area of research	Key finding
Acceptability	The TWiTCH programme was generally seen as acceptable to settings. Although coach training and support to settings was delivered inconsistently by different coaches, training and coach support was seen by practitioners as adequate and effective at improving practitioner knowledge and confidence for delivering TWiTCH for the majority of practitioners. Settings were positive about the programme, however, Week 2

	activities were often considered to be pitched at too high a level for children, especially at the beginning of the programme.
Fidelity	Almost all settings (92%) were able to implement at least six cycles of the programme during the year, with 38% completing all nine and 29% completing eight—some with time remaining to complete a ninth cycle at the time of data collection. The programme was delivered with a varying degree of fidelity for different elements across settings. Practitioners attended training and engaged with coaching. The group of children who received TWiTCH varied across settings with some settings not delivering TWiTCH with their full cohort of preschool children due to logistical challenges related to space and insufficient staffing numbers to facilitate smaller group work for all children. Children attending nursery part-time were also excluded, sometimes due to not being in nursery for all TWiTCH sessions. In the delivery of TWiTCH, Week 1 sessions of shared book reading were delivered with high fidelity to the programme, however, the Week 2 choice-based language activities were delivered with more variability due to perceived difficulty of the activities and subsequent adaptations made by practitioners. Targeting specific 'at need' children during Week 3 of the cycles was often not done. Practitioners did not always see the value of targeting specific children with TWiTCH activities over their usual practice of targeting 'at need' children more generally.
Evidence of promise	The evaluation showed evidence of promise, with practitioners' quantitative and qualitative data suggesting that the TWiTCH programme contributed to their increased knowledge and confidence in supporting children's language development, and to changes in practice both during storytelling and general practice. Practitioners reported observing improvements in children's communication, language, vocabulary, use of reasoning language, and confidence to contribute in a group setting, which they attributed to the programme. Practitioners felt that taking part in TWiTCH led children to have a better understanding of—and ability to retell—stories.
Readiness for trial	The TWiTCH programme is close to being ready for a larger-scale trial. Readiness for trial is conditional on updating training and materials for Week 2 to ensure practitioners can deliver the content to the target age and ability of children, tailoring training and materials to make sure delivery is suitable in larger group sizes, and refining training for coaches with the aim to achieve greater consistency in delivery of training and coaching across settings. The contribution analysis found that the TWiTCH theory of change was mostly supported and concluded that the expected TWiTCH outcomes occurred and could be attributed to the programme. For a future trial, consideration needs to be made of the target group for the programme given the inconsistency of delivery to the full year group in the pilot.

Additional findings

The trained coaches model was effectively used to allow scaled-up delivery of the TWiTCH programme to a greater number of settings, as required for this pilot. However, training by coaches was inconsistent and sometimes omitted some of the training content. Additional training and support, as well as a quality assurance process for delivery of training to settings, would be required to ensure consistency across coaches during trial delivery.

TWiTCH was facilitated by practitioners being given agency in the delivery of the programme. TWiTCH was delivered more successfully when practitioners adapted the activities to the specific needs of the children and their interests, however, not all practitioners did this. Practitioner confidence and experience, or support from the TWiTCH champion, were key to successful adaptation. Greater focus on adapting the programme could be given during initial training and in the materials, which could better support practitioners with making successful adaptations.

Settings incurred very few additional monetary costs to implement the programme. A few practitioners reported printing costs for additional resources to use in the sessions. The time required for practitioners to attend the six hours of training was the biggest resource required.

Introduction

Background evidence

Importance of early language and communication

The importance of developing early language and communication skills as a foundation for future learning and life outcomes has been widely acknowledged. Early language skills at age four can predict language outcomes throughout primary school and the persistence of low language skills is linked with lower academic achievement (Snow, Burns and Griffin, 1998; Eadie et al., 2021). Preschool vocabulary skills can predict reading comprehension after three or four years of school even when controlling for factors such as parent literacy and education and the child's early literacy skills (Sénéchal, Ouellette and Rodney, 2006).

Studies have shown that children from disadvantaged backgrounds enter school with lower levels of language and early reading skills than their more socioeconomically advantaged peers (Tymms et al., 2014) and this trend persists throughout primary school (Merrell, Little and Coe, 2014) where lower language skills affect children's ability to understand, learn through spoken language, and express ideas and emotions (Oxford University Press, 2018).

Attendance at preschool and the quality of preschool provision have been shown to predict children's scores on oral language at entry to school (Sammons et al., 2004; Justice, Jiang and Strasser, 2018), reading and maths assessments at Key Stage 1 (Sylva et al., 2004), maths and science achievement at Key Stage 2 and Key Stage 3 (Sammons et al., 2011), and quality of GCSE results (Sylva et al., 2014). This indicates justification for the provision of interventions that improve the quality of the early years learning environment as a way of improving language and other academic outcomes later.

Dialogic pedagogy and sustained shared thinking

'Dialogic pedagogy' is the use of conversation to engage students, stimulate their thinking, and to help them to understand and expand ideas (Alexander, 2017). It uses open-ended questions and discussion between teachers and students to support students' learning. The benefits of dialogic pedagogy in the classroom have been described in the literature as developing deeper conceptual understanding, higher cognitive processes, and the ability to express views (Laird-Gentle, Kanasa and Grootenboer, 2023). An EEF trial of dialogic teaching in primary school found that its use had positive effects on attainment scores for children in Year 5 (Jay et al., 2017).

The concept of 'sustained shared thinking' (SST) is used more frequently in the early years and is defined by Sylva et al. (2004, p.36) as:

> *[A]n episode in which two or more individuals 'work together' in an intellectual way to solve a problem, clarify a concept, evaluate activities, extend a narrative, etc. Both parties must contribute to the thinking, and it must develop and extend.*

In an early years setting this usually means an adult practitioner and children having an in-depth discussion or exchange where the adult and child both develop the ideas in the conversation and joint thinking and learning takes place (in contrast to the transmission model of teaching where the adult shares knowledge and the child absorbs this without questioning). SST supports deep-level learning, encouraging children to think critically and contribute to their own learning. SST between an adult and child can support children's learning on three levels: first, by encouraging children in their learning, second, by modelling more advanced language structures, and third, by developing and extending children's thinking on a topic (Brodie, 2014).

Dialogic approach-based interventions in the early years have been found to have positive effects on children's learning and development. A systematic review investigating the link between dialogic approach-based early childhood interventions and children's learning found that six out of eight studies included found positive effects of interventions for children's learning and development (Garcia-Carrion and Villardon-Gallego, 2016). The quality of practitioner-child interactions was strongly and consistently associated with children's academic and language skills. One U.K. study included in the systematic review found that in preschool classrooms where adult-child interactions involved 'sustained shared thinking' (Taggert et al., 2006) there were still positive impacts on children at age 11 and at age 14.

Episodes of SST have been found to be relatively low in early years settings at around 10% of classroom interactions and have been observed more frequently in interactions between qualified teachers and children rather than those involving unqualified teachers (Siraj-Blatchford and Manni, 2010; Meade et al., 2013). Early years practitioners recognise the importance of high-quality adult-child interactions but they have limited CPD opportunities and consequently lack skills and confidence in developing children's language and communication and their use of sustained shared thinking (Sylva et al., 2004). It seems likely, therefore, that professional development to support early years practitioners in developing this practice could lead to improved impacts for children. Interventions which provide professional development (PD) support to early years practitioners to improve the quality of conversational exchanges with preschool children in the classroom have led to increased use of dialogic strategies and improved quality of classroom talk as well as an increase in expressive language outcomes for children (van der Wilt, Bouwer and van der Veen, 2022; Piasta et al., 2012). The use of dialogic pedagogy in the primary school setting has also been linked to higher learning outcomes and more positive attitudes to school (Hardman, 2019).

Sharing books

Exposure to books is a well-established means to developing children's vocabulary and knowledge of language and print (early literacy skills) (McKeown and Beck, 2006; Sénéchal and LeFevre, 2002) and interventions focused on adults and children reading books together have been shown to impact on language and early literacy skills. For example, a meta-analysis of studies which looked at the impact of interactive and structured shared book-reading in kindergarten (Mol, Bus and Jong, 2009) found that shared book reading impacted on oral language skills (including vocabulary) and on early literacy skills, while a meta-analysis of parent-child book sharing interventions also found positive impacts on children's expressive and receptive language (Dowdall, et al, 2020). A further meta-analysis of home shared book-reading programmes found that it was only dialogic shared reading methods—specific structured approaches where the adult facilitates dialogue about the book asking questions of the child to promote deeper understanding—that had positive effects on children's outcomes (Barone et al., 2019) showing that the dialogic nature of the intervention is important. A meta-analysis by Flack, Field and Horst (2018) also demonstrated that how shared reading is done influences the number of words learnt, with dialogic reading styes having a greater impact that shared book-reading alone, and repeated readings leading to greater learning of words. However, not all shared book-reading interventions have shown large effects and some studies have found that effects are small and variable across socioeconomic status (Barone et al., 2019) and some shared reading programmes have not found effects even when dialogic pedagogies are embedded (for example, Burgoyne et al., 2024) so it is not clear that shared book-reading is always effective for everyone.

Philosophy for Children (P4C)

Philosophy for Children (P4C) is a philosophical approach to learning and teaching, originally developed by Lipman and Sharp (1978), which aims to enable children to think with others and for themselves and to develop children's reasoning abilities. The approach has been adapted since its inception and is used in schools with children of different ages in the U.K. and around the world. The P4C programme highlights the value of collaborative reasoning about social concepts and ideas that matter to children and provides children with the tools to question, identify inconsistencies, and reason through discussion supported by a practitioner (SAPERE, n.d.). The P4C programme has a focus on social concepts which are 'contestable' (that is, not universally agreed on) and which reflect societal values or beliefs. P4C advocates the use of picture books which are rich in these concepts and accessible, relevant, and already familiar to children and practitioners.

A number of systematic reviews and meta-analyses which have looked at the impact of P4C have found the programme to have positive impacts on children (Trickey and Topping, 2007; Garcis-Moriyon, Rebollo and Colom, 2005; Ad Wahab, Zulkifi and Razak, 2022), including on reasoning skills, thinking skills and language skills, however, none of the studies included in these reviews were with early years children. The EEF have also commissioned two effectiveness trials of the P4C programme looking at the impact of the programme on primary school children. The first trial found that P4C had a positive impact on reading and maths at Key Stage 2 and a positive impact on cognitive abilities (Gorard, Siddiqui and See, 2015) while the second larger trial did not find that P4C impacted on children's reading or maths scores or social and communication skills, although teachers did report positive outcomes for pupils and there was no negative impact of the P4C programme on any outcomes (Lord et.al., 2021).

The TWiTCH programme

Talk With Tales for Children (TWiTCH) is a CPD programme supporting nursery practitioners to use dialogic pedagogy and sustained shared thinking (SST) in their teaching through a structured, cyclical programme. Drawing on the principles of P4C and dialogic reading principles, TWiTCH is embedded within the usual practice of story-time, which is a common feature in the U.K. early years classroom. The programme teaches practitioners how to use story-time as a daily opportunity for complex language interactions between children and adults. It supports the use of adult-child talk through language games, providing opportunities for SST, which encourage children's thinking and reasoning skills. These interactions aim to give children the language to express complicated thoughts and ideas. The programme includes targeted reading aloud and book discussion with children, explicitly extending pupils' spoken vocabulary using structured questions to develop reading and language comprehension and the use of purposeful curriculum-focused dialogue and interaction.

As a communication and language approach drawing on several evidence-based strategies highlighted as effective in supporting children's language development and future literacy in the EEF's Preparing for Literacy guidance report (EEF, 2018)—including dialogic shared reading, promotion of high quality interactions with children through sustained shared thinking, and activities to develop reasoning—the TWiTCH programme was seen to have the potential to improve children's language through changing the practice of practitioners. The project was expected to provide learning about the practice of early years practitioners around dialogic reading approaches and how these could be implemented where there was not much existing research.

The TWiTCH programme has previously been delivered in a small pilot of five nurseries prior to this larger pilot evaluation, however, this research was disrupted by Covid-19 and two settings were unable to engage at all with the early research. Nevertheless, findings from three settings, one PVI and two school nurseries, (Demissie, 2019) in this early pilot indicated that TWiTCH was delivered fully in two of these settings and partially in one setting, despite the Covid-19 restrictions. TWiTCH was generally seen as acceptable and valued to participating practitioners. In one setting where observation of practice using the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale (Siraj, Kingston and Melhuish, 2015) had been completed before and after the TWiTCH programme, the programme seemed to improve the quality of practitioner interaction and practice. Interviews with the headteacher and lead practitioner at one setting found that they felt the programme had significant impact on children's language development and thinking skills and that it encouraged children to think beyond what they would normally have been encouraged to do. Assessments conducted in this setting also supported this with a significant number of children starting nursery below age-related expectations (65%) and between 65% and 81% (depending on subject) moving to age-related expectations or one step below during the year. Children were progressing well against most of the Tower Hamlets' Progression in Language Structures (Tower Hamlets, 2009) statements showing particular development of retelling and argumentation language (for example, the use of 'yes because' and 'no because' statements.

Although these changes could have been developmental rather than due to the programme, the specificity of improvement related to the elements of the programme provides some initial evidence of promise for the programme and justification for testing it further.

The TWiTCH programme has been further developed based on feedback from the early pilot, and the training model has been scaled up to allow delivery and support to more settings through a 'train the trainer' model.

The delivery of the TWiTCH programme in this evaluation was through the U.K. government's Early Years Stronger Practice Hubs. This funding facilitated the delivery of TWiTCH to three cohorts of nursery settings with the second cohort being the focus of the evaluation reported here. The first cohort (Phase 1) began delivery of the programme in April 2023 and the delivery of TWiTCH for this cohort was internally evaluated by the developer team with feedback from Phase 1 leading to the update of the programme. The second cohort (Phase 2) reported on here delivered TWiTCH beginning in September 2023. A third cohort (Phase 3) is delivering the TWiTCH programme beginning in September 2024 and is ongoing at the time of this report. Updates to the TWITCH programme for the third cohort have been made as a result of feedback from this evaluation. These changes for Phase 3 delivery are also described in this report although the third cohort does not have an evaluation element.

This evaluation reports on the piloting of the TWiTCH programme across a larger number of settings to the initial pilot and investigates the feasibility, the acceptability, and fidelity of the programme delivery as well as evidence of promise to justify scaling up to an impact trial.

Intervention

Name

Talk With Tales for Children, 'TWiTCH'.

Why (theory/rationale)

As discussed above, TWiTCH builds on the principles of dialogic pedagogy, shared dialogic book reading and sustained shared thinking which have previously been found to have positive impacts on learning. TWiTCH is also influenced by the principles of the P4C programme and in particular promoting children thinking and reasoning language through the discussion of the moral concepts found in traditional tales. The TWiTCH programme structures activities which are designed to stimulate practitioner/child interaction around concept-rich traditional fairy tales told in the UK (for example, 'Jack and the Beanstalk', 'Little Red Riding Hood'). The use of storybooks which are rich in concepts have also been found to provide opportunities for high-quality interactions which explore ideas (Haynes and Murris, 2012).

Through TWiTCH training, it is expected that practitioners learn how to use a clearly structured intervention that builds on familiar early years (EY) practice (story-time and familiar daily activities) to create a language-rich environment that improves the quality and quantity of adult-child interactions, leading to improvements in the vocabulary and complexity of children's language and their ability to express opinions. TWiTCH uses the dilemmas and moral concepts in stories as the basis for discussion and activity. Drawing on dialogic pedagogical methods, the adult uses facilitative rather than directive questioning approaches to develop children's language and thinking skills. For example, practitioners might ask children to discuss 'Do you need to have superpowers to be a hero?'.

Who—recipients

The programme is for nursery practitioners working with three- to four-year-old children in the year before they start reception year at school.

Early years education settings that offered funded preschool places for three- to four-year-old children—including private voluntary and independent (PVI) settings and maintained nursery settings—across West and South Yorkshire took part in the project. Settings were asked to have at least three practitioners working in the preschool as the programme involved one member of staff, the 'champion', providing support to other members of staff in the setting. Settings currently taking part in the Early Years Professional Development programme were excluded from taking part in this study due to the overlapping elements of the programme and the demands of both programmes.

Practitioners working with children aged three to four in these settings were invited to take part in the programme and to deliver the TWiTCH activities in their classrooms.

One practitioner in each EY setting was selected by the setting to be the TWiTCH champion. In this role they were the key contact between the developer team and the setting; they received additional training and were expected to support the delivery of the programme across their nursery. Guidance was given to settings that the TWiTCH champion should be experienced in working in early years settings with this age group and be someone interested in supporting the development of other practitioners in the setting.

What—procedures

Training for coaches

Coaches attended two three-hour sessions of training as a small group. The first session introduced coaches to the theoretical aspects and research base for TWiTCH as well as describing the structure and core principles of the programme. The session provided time for coaches to discuss the programme and to engage intensively with the practitioner and coach manuals as well as the resources. Session two provided further training on practitioner-child interactions, providing examples of good practice and asking coaches to evaluate the use of different types of prompts. Practitioners were also asked to role-play the TWiTCH activities they would be asking practitioners to do in their training. The coach role and expectations for coaches were also discussed in this session alongside talk about programme fidelity.

A coach WhatsApp group was set up for coaches and the developer team to share their experiences and to ask for support. Coaches also met with the developer team three times during the year to feedback on their experiences of working with settings. This communication and feedback were important to the programme delivery as this was the key way that the developer team was able to learn about how nurseries were responding to and delivering the TWiTCH programme.

Initial meeting between coach and setting senior leadership team

Coaches were expected to arrange a meeting with a member of the senior leadership team in each of their settings in advance of the first TWiTCH practitioner training to give the setting an overview of how TWiTCH works and what would be expected of the setting, as well as to learn about the context of the setting.

Training and support for all practitioners in settings

Coaches were asked to provide two three-hour in-person training sessions held at each participating setting (with at least a two-week gap between sessions). Each coach supported five settings. This training was designed to be attended by all practitioners working with three- and four-year-olds (scheduled at a time convenient for the setting). The first session aimed to introduce practitioners to the programme and some of the techniques, to establish the relationship between the coach and practitioners, and to find out about the setting's starting point and current practice. The training covered the programme's three-week structure, research evidence for TWiTCH, principles of interactive book reading, and focused on the Week 1 activities building in time for practitioners to explore the manual and to role-play the Week 1 activities. Coaches guided practitioners on how TWiTCH could fit into their current context. The second training session aimed to encourage practitioners to reflect on their understanding of TWiTCH so far, and to support practitioners' understanding and practice of the Week 2 and Week 3 activities. The training covered the importance and benefit of facilitative interaction and using big ideas for discussion, the choice-based language games that are part of Week 2 activities, and the expectations for Week 3 activities. The training provided opportunities for practitioners to role-play the activities and receive feedback from the coach and time for planning and discussion.

Training and support for TWiTCH champions

A two-hour online training session for the TWiTCH champions was hosted by the developer team in November 2023 (practitioners had the option to attend one of two sessions hosted at different times on the same day). This session aimed to provide the TWiTCH champion with additional information about the role of the Champion, and deeper knowledge of the underlying theory and techniques used in TWiTCH. The session focused on dialogic pedagogies and facilitating discussion and reflection about TWiTCH delivery up to this point.

Three one-hour TWiTCH champion network events held throughout the delivery period brought TWiTCH champions from different settings to meet online together (champions had the option to attend one of two sessions hosted at different times on the same day). These sessions encouraged champions to reflect on the delivery of TWiTCH in their setting and to ask for and receive feedback from both the developer team and the other TWiTCH champions.

Ongoing support from coaches

Settings were supported in their delivery of TWiTCH throughout the programme through a schedule of coaching visits—three face to face, two-hour coach visits to the setting and three online coaching contacts, which could include email or online meetings. Coaching visits were spread throughout the programme delivery, alternating between face to face and virtual. In the face-to-face visits, one practitioner was observed by the coach to review and plan for the next stage of delivery. Online contact between the coach and the champion reflected on how the programme was going and any issues the setting was experiencing. Coaches were expected to act as a critical friend and support for the setting, talking to practitioners about their experiences of the TWiTCH programme, observing, and offering constructive feedback.

The TWiTCH champion role

The TWiTCH champion was expected to be a liaison for the programme and to support the delivery of TWiTCH in their setting. The champion role included:

- liaising with the coach prior to visits to clarify aims and expectations;
- overseeing fidelity of delivery by liaising with other practitioners about how they are implementing the programme and providing support, suggestions, and ideas for better implementation;

- coordinating availability of storybooks and language activity resources; and
- actively initiating peer reflection opportunities for colleagues.

Delivery of TWiTCH programme in settings

TWiTCH is based on a three-week story cycle which settings are encouraged to implement after the second training session. The three-week cycle involves:

Week 1

The storybook for the three-week cycle is introduced to the children in small groups of five to eight and read to them at each daily session,[1] drawing out sequence, concepts, and characters. Each day's reading has a different focus:

- day one—overview of the story;
- day two—characters in the story;
- day three—sequencing;
- day four—listening to and enjoying the story; and
- day five—prediction.

There are prompts for practitioners about what to ask children in each session.

Week 2

Week 2 involves continued reading of the storybook to the whole class or group as required or requested. Over three days, three choice-based language games should be played in the small groups based on the ideas, dilemmas, or concepts and characters as a vehicle for dialogue and sustained shared thinking. These language activities encourage the children to make a choice, give reasons, and reflect on their reasons through language such as 'I agree because' or 'I disagree because'. Practitioners are encouraged to use specific techniques to facilitate children's reflections including:

- repeating and checking what the child says (REPEAT/CHECK);
- asking children to give reasons for their choices and what they say (REASONS); and
- asking 'what if…' questions (IMAGINE).

Initially, Week 2 activities are mainly conducted between the practitioner and a rabbit puppet to expose the models of language used and build the practitioner's and children's confidence in engaging with the activities. Children are not expected to do much at first, but as the children become more used to the format and the practitioner more confident with delivery, they are expected to more actively involve the children.

While practitioners are encouraged to use the rabbit puppet throughout the programme to engage the children with the story and to model the language responses that they would like to see from the children, interactions between the practitioner and rabbit are key for Week 2 activities where a script for practitioners and for rabbit introduces the children to the choice-based language activity games and the language of argument and reasoning. Scripted interaction between the rabbit and the practitioner provides the opportunity for practitioners to practise their target language prompts (REPEAT/CHECK, REASONS, IMAGINE—'RCRI') and build confidence in using these interaction styles with the children. Practitioners adapt their interactions with rabbit based on responses from the children in the room. It is expected that practitioners will move away from reliance on the script as they grow more confident with the TWiTCH programme and the children become more familiar with the language games.

[1] This was later revised to small groups of up to ten children due to some settings' difficulties in staffing, timetabling, and space for smaller groups.

Week 3

The story is used as the basis for activities offered across familiar daily activities as well as in story-time. These activities could be offered to the whole group or in smaller groups. Practitioners are encouraged to intentionally focus their support and engagement on children with greater needs in language during this period. Practitioners are asked to incorporate at least three of the suggested activities during this week. Activities could include story-scribing (where a child tells a story and a practitioner writes the story down word for word), one to one discussion, or playing alongside to narrate and recast children's language.

The three-week cycle is repeated across the year using different stories in order to enable both practitioners and children to become familiar and confident with the format.

Settings were expected to deliver nine cycles of the programme and were given nine storybooks and materials. It was not required for all practitioners in the setting to complete the nine cycles—just that the setting delivered each of the nine cycles. This could mean that one practitioner completed all of the cycles or that multiple practitioners completed all of the cycles with different groups or that practitioners shared the delivery of the cycles. Towards the end of the nine cycles, settings were expected to be supported by their coach to choose their own book and come up with their own activities for a tenth cycle, which they should then deliver.

What—materials

- TWiTCH manual for coaches covering the background and further detail about the programme, the role of the coach, training for practitioners, and a logbook template of information that should be captured during each interaction between a setting and a coach.
- TWiTCH coach CPD training slides for two sessions.
- Practitioner CPD training slides and accompanying notes and timings for two sessions.
- TWiTCH manual for practitioners including TWiTCH background, details for each of the nine book cycles of what each session should contain, and a bank of activities to choose from for week three).
- Resource pack for each three-week cycle including:
    - the storybook;
    - wooden puppet sticks;
    - laminated printed cards with key phrases and symbols related to activities;
    - rabbit puppet; and
    - carrot soft toy.
- Video recording of TWiTCH practice.
- Audio recording of dialogue between practitioner and children illustrating dialogic pedagogy.

How—format

Training and support for nursery practitioners was a combination of face-to-face sessions in the setting (training and some coaching sessions) and online sessions (TWiTCH champion training and network meetings and some coaching sessions). The delivery of the TWiTCH activities was designed to be as part of normal preschool practice with the activities designed to fit with what settings are already doing. Practitioners were asked to work with groups of up to eight children for the delivery of the activities.

Tailoring—adaptation

Practitioners were expected to incorporate the programme into their normal practice in the way that best suited their setting and children and were given flexibility in who delivered the sessions, the size of the groups, and when the sessions were delivered. The nine cycles could be delivered in any order. Settings were given particular flexibility in Week 3 of each cycle to choose activities that fitted the needs of their children and the resources available in the setting. Settings were also encouraged to focus and tailor the Week 3 activities to the needs of specific children, especially those that need some additional support. Settings were provided with some guidance in the training to help them select

which children might benefit from this support; this could include, but was not limited to, children who are part-time and missed Week 1 or Week 2 sessions, children who were reluctant to communicate in the small group, and children who might benefit from additional language support. Coaches were expected to provide feedback and suggestions to practitioners to support their specific needs.

When and how much—dosage

All nursery practitioners who would be delivering TWiTCH were expected to attend the two initial CPD training sessions.

Following training planned for October 2023, practitioners were expected to deliver nine three-week cycles of the TWiTCH programme between October 2023 and June 2024 using the materials provided. A tenth cycle was then expected to be developed and delivered by the setting in July 2024. There were different expectations for each of the three weeks of the cycle.

- in Week 1, five sessions of reading the story and related discussion should be delivered;
- in Week 2, all three language activities need to be delivered during the week along with story reading as needed or requested; practitioners should aim to include all children in each activity at least once; and
- in Week 3, practitioners choose at least three activities from the selection provided to offer during the week, as well as focusing on the needs of specific children through working in smaller groups.

All activity is done as part of usual nursery practice and should be incorporated into the normal daily routine. The expected length of each session was around 10 to 15 minutes. The TWiTCH champion is also expected to attend additional training and all networking events as well as engaging in the six coaching visits.

Who—provider

The Sheffield Hallam University TWiTCH developer team provided all training and materials to the coaches and practitioners of the TWiTCH programme. Five trained coaches employed by the developer team each supported five settings in their delivery of TWiTCH including delivering their initial training, virtual and in-person coaching visits, and ad hoc ongoing support to the setting and the TWiTCH champions.

The TWiTCH champion coordinated and supported the delivery of TWiTCH in their setting. Practitioners then used the resources to implement TWiTCH in their setting with the children in the three- to four-year-old room.

Where—location

TWiTCH was delivered in both maintained and PVI early years settings recruited from the West and South Yorkshire region with a focus on areas of high socio-economic deprivation.

Theory of change

The theory of change model developed at the start of the programme is included in Appendix C1 along with the associated contextual and causal assumption logs in Appendices C.2 and C.3. The narrative description below describes the expected causal pathway between the programme's inputs and activities and the expected outputs and subsequent impact. This includes the additional level of the coach 'train the trainer' model, which was not described in the originally developed theory of change but has been added into an updated version (Appendix C.4). A summary of the theory of change is given in Figure 1.

The TWiTCH theory of change operates over multiple levels. It involves:

- coaches: the programme developers using a 'train the trainer' approach to recruit and train up the coaches to train and support programme in preschool nurseries;
- practitioners: the TWiTCH practitioners—including the TWiTCH champion—taking part in training and delivering the programme; and
- children: the children taking part in the TWiTCH activities being expected to benefit from the programme.

For each level there are programme inputs, expected outputs, and expected short term and longer-term outcomes. Assumptions are made as to how the inputs will lead to the expected outputs and impact through the programme. In the narrative below, the different aspects of the theory of change are described along with any relevant assumptions from the assumption logs.

Figure 1: Summary of TWiTCH theory of change

Coach level

The programme is reliant on the recruited coaches to transmit the programme knowledge and information and to provide support to nurseries for them to deliver the programme. Therefore, the recruitment of coaches with required skills and experience to deliver the training for settings and to support early years settings to adapt the programme to their needs [input]—an important causal pathway for the programme. Coach training and support provided by the developer team [input] provides coaches with the programme-specific knowledge that they need to understand the TWiTCH programme. Following the training, it is expected that coaches feel knowledgeable and confident to deliver the programme [short-term outcome] so that they can effectively and consistently deliver training in line with the programme expectations and provide support to the nurseries during the programme [output].

Coaches are provided with resources [input] to facilitate their role and to support consistency and fidelity of their training and to support delivery. This includes a coach handbook with details about the programme, expectations of coaches, useful resources, and a coach logbook template to enable the recording of information about each setting they visit. Coaches are also provided with a resource pack identical to that which settings receive. Slides to use during training

and detailed notes on the content and timings for training delivery were also provided with the expectation that training would be done in a consistent manner across coaches and that all aspects of the training were covered.

Access to support for coaches from the developer team and other coaches [input] throughout the programme through the shared WhatsApp group and through occasional meetings [input], allows coaches to clarify their understanding of the programme [output], to share practice and to better support their own settings while maintaining fidelity to TWiTCH [short-term outcome].

Practitioner level

At practitioner level, the programme provides two sessions of on-site training for the preschool team of nursery practitioners delivered by the coach allocated to their setting [input]. This training provides practitioners with the background to TWiTCH and use of dialogic pedagogy and information about the three-week cycle of storytelling and activities. It provides practitioners with the video modelling of TWiTCH techniques so that they understand what TWiTCH looks like in practice; it gives practitioners the opportunity to practice the TWiTCH activities and techniques and to plan with a coach how the programme will work in their setting. It is expected that following the training, the practitioners will have been equipped with sufficient knowledge, skills, and confidence to deliver the programme [output]. It is assumed that practitioners lack confidence to support young children's language and sustained shared thinking [assumption] and that the TWiTCH programme training will improve this. It is assumed that delivering the training on-site will allow more nursery practitioners to be able to attend than had it been delivered offsite [assumption] and that delivering training specifically in the setting with practitioners only from that setting will allow the coach to adapt the training to the context of the setting [assumption]. It is also expected to facilitate the establishment of a working relationship between the coach and the practitioners [output].

The TWiTCH programme provides nurseries with resources to facilitate delivery [input]. This includes a handbook which details exactly what should be done in the programme across the three-week cycle for each of the nine TWiTCH storybooks. Accompanying delivery resources are also provided along with a rabbit puppet for modelling language practice as described in the manual; a carrot soft toy is also provided for practitioners to use to facilitate turn-taking during TWiTCH sessions. It is expected that practitioners will access and use these resources in their delivery of the programme as described in the handbook and at their training sessions [output].

Following training, nursery practitioners are expected to deliver the nine cycles following the schedule set out in the handbook with children in their setting [output]. There is an assumption that practitioners will start out sticking very closely to the handbook but that over time they will grow in confidence and skill using TWiTCH and will tailor the programme better to the needs of their children [assumption].

During the year of delivery there is expected to be ongoing support provided by the coach to the setting [input]. This support consists of termly coaching visits to the setting to observe TWiTCH practice and provide feedback [input] as well as remote contact with the TWiTCH champion to discuss how TWiTCH is going. This support is expected to develop practitioner skills and confidence in delivering TWiTCH and, in the later sessions, support practitioners with moving on to use TWiTCH practices with their own storybooks [output]. This support is expected to facilitate reflection in practitioners' practice and practitioners are expected to take action following these visits to develop their own practice [output].

The inclusion of the TWiTCH champion role within each nursery [input] as part of the TWiTCH programme is intended to facilitate the communication between the nursery and the developer team, to support other staff with delivering TWiTCH, and raise the profile of the TWiTCH programme within the setting. The provision of an online TWiTCH champion training session [input] accessed by all TWiTCH champions together is expected to provide these practitioners with additional knowledge about the programme and to support practitioners in how they can support other staff in the setting. The training and guidance about the role are expected to lead to the TWiTCH champion liaising with the coach and taking responsibility for monitoring and improving TWiTCH practice in the setting following the visit [output]. In this role it is expected that the TWiTCH champion will increase in their confidence for supporting other staff to reflect and improve on their practice [outcome]. Attending facilitated online networking meetings with other TWiTCH champions and the programme developer team is expected to allow practitioners in these roles to reflect on their delivery of TWiTCH and to share their experiences [output]. These meetings are expected to allow TWiTCH champions to share any challenges or concerns and to receive support directly from the developer or to get ideas of how other settings have delivered or extended TWiTCH and overcome any issues.

For all early years practitioners taking part in the programme, the training and subsequent delivery of TWiTCH during the year is expected to leave practitioners well prepared and confident to deliver the TWiTCH programme [short-term outcome] and ultimately to improve practitioners' skills, knowledge, and confidence for supporting children's language development [longer-term outcome]. Staff are expected to be able to deliver TWiTCH effectively and be able to adapt the programme and books to work in their own context [short-term outcome] and to use the TWiTCH principles of dialogic pedagogy and sustained, shared thinking across all their practice beyond the TWiTCH programme [longer-term outcome].

Child level

It is expected that all the three- and four-year-old children in the nursery setting would access the sessions of the TWiTCH programme, focusing on one storybook over the three-week cycle, doing activities related to the story each day using a dialogic approach facilitated by the practitioner and continuing this for around ten cycles [inputs]. The repeated readings will increase the children's exposure to each story and are expected to improve their understanding of them, making them better able to engage with the other TWiTCH activities. Children will access these sessions in small groups of up to eight so that all children will have the opportunity to contribute and practitioners will be better able to support the needs of all children in the group [input]. It is expected that children will be engaged in the TWiTCH sessions and will contribute to them [output] and that during the sessions they will reflect and reason about the concepts in the stories through their dialogic interactions with the practitioner and other children in the group [output].

Children are also expected to use talk to express their ideas and to make links between what is happening now and what may happen [output] during the TWiTCH sessions. As a result of taking part in TWiTCH, children are expected to increase their story comprehension skills and their ability to recall and retell the stories [short-term outcome]. They are also expected to have a wider range of vocabulary, which they can use and understand related to the stories [short-term outcomes]. The programme is expected to build children's communication and language skills and to increase their confidence in using these skills [short-term outcome]. Specifically, it is expected that children will be able to use more complex sentences and language structures to express their thoughts as a result of the activities they take part in during TWiTCH [short-term outcome]. Improving children's skill and confidence in language and communication is expected to lead to impact on the Early Years Foundation Stage (EYFS) scores for communication, language and literacy [longer-term outcome], improved readiness for school [impact], and, ultimately, improved literacy achievement in primary school [longer-term outcome].

Research questions

The evaluation aimed to investigate the acceptability of the TWiTCH programme to early years settings and the extent to which it was delivered with fidelity, specifically regarding the CPD, implementation of the programme, and changes to practice. This was done with focus on the practitioner and setting, rather than impact and experience for the child. It was expected that any changes to child outcomes would be dependent on changes to practitioner knowledge and practice.

Research questions were developed across four areas aligning to the evaluation aims: feasibility of trial, programme acceptability, programme fidelity, and evidence of promise.

Feasibility of trial

1. To what extent does the recruitment strategy recruit and retain targeted disadvantaged settings and practitioners throughout the study?
    a. What are the characteristics of settings that are recruited to the project?
    b. To what extent are settings recruited from targeted areas of disadvantage?
2. How feasible is the delivery of the programme within the delivery period?
    a. Delivery of training.
    b. Delivery of nine cycles of TWiTCH.
3. To what extent is the TWiTCH programme defined (manualised) so that training and support is delivered with consistency between settings and across different trainers?

Acceptability of programme

4. Is the TWiTCH programme acceptable to settings and the individual practitioners working within them?
    a. Is the training at the appropriate level and accessible to practitioners in each setting?
    b. Is the level of support provided by the developer team sufficient and how are the different elements viewed (training, coaching visits, online support)?
    c. How is the programme received by practitioners in different roles and different settings?
    d. Is the intervention suitable for all children in the setting?
    e. How does the level of agency/decision-making given to the practitioner in the programme affect delivery and views of the programme?
5. What costs and resources are required by settings implementing TWiTCH and is this feasible and acceptable to settings? What cost would settings be prepared to pay for the intervention?

Programme fidelity

6. Is the TWiTCH programme delivered with fidelity in settings?
    a. To what extent do all practitioners attend the training? What factors affect whether practitioners can attend?
    b. To what extent is the programme delivered in classrooms with fidelity? How does this vary between settings? How consistent is delivery across practitioners and settings? What are the reasons for variation in delivery?
    c. What role does the TWiTCH champion play in supporting the intervention delivery?
    d. How is the programme adapted by those delivering it?
7. What challenges/barriers are faced by settings in delivering the intervention?
8. What factors support the delivery of the programme?

Evidence of Promise

9. How different is the TWITCH programme to (a) other CPD programmes the settings have participated in, (b) existing knowledge, and (c) usual setting practice?
10. Does practitioners' practice change (and if so, how) as a result of taking part in the TWiTCH programme (measured through self-report and audio observations)?
11. How does practitioner confidence and knowledge for supporting children's language development change (i) immediately following training and (ii) throughout delivering the programme as a result of the TWiTCH programme?
    a. Does this vary across settings, practitioners, and roles?
12. What impact do practitioners feel the programme is having on children's language development? Are there any unintended outcomes?

Success indicators for the project and delivery of the TWiTCH programme were developed at the beginning of the project with the developer team and were included in the study plan (Table 2). These indicators were designed to inform an overall assessment of the pilot implementation dimensions (evidence of promise, feasibility of implementation, and readiness for trial).

Table 2: Pilot success indicators

Pilot criteria	Success Indicators	How to assess this?	Research question
Evidence of promise	Changes in practitioner practice in the majority of settings.	Audio recordings, survey, case studies, interviews, coach feedback.	10
	Increased practitioner confidence and knowledge in supporting children's language development across the majority of the practitioners.	Surveys, interviews.	11
	Practitioners in the majority of settings perceive TWiTCH to have a positive impact on children's language development.	Interviews, case studies.	12
Feasibility of implementation	Programme is delivered in settings of different types with medium to high fidelity as judged by the developer team.	Developer team interview, coach feedback, interviews with TWiTCH champion and practitioners.	6
	Initial training sessions are attended by minimum of three practitioners per setting for majority of settings and are accessible to practitioners.	Training attendance records, post-training surveys.	6
	At least six TWiTCH cycles completed by majority of settings.	Coach feedback, surveys, interviews.	2
	Absence of any major barriers to delivery for majority of settings.	Interviews, case studies.	7
	TWiTCH programme seen as acceptable and accessible to majority of practitioners.	Surveys, interviews, case studies.	4
Readiness for trial	TWiTCH training and support is delivered consistently across different settings and trainers.	Observations.	3
	TWiTCH programme sufficiently described to enable consistent delivery.	Interviews, coach feedback.	3

Ethical review

Ethics approval for the evaluation was sought from the Durham University School of Education Ethics Committee and was originally approved on 22 March 2023. Several updates were made to the original ethical approval during the project for changes to methodology and approval of research instruments. Ethics approval specifically for the delivery of the TWiTCH programme of activities was sought by the developer team from the Sheffield Hallam Ethics Committee.

Nursery setting agreement to take part in the project was obtained through the setting leadership signing a memorandum of understanding document (see Appendix A.1), which provided full details about the project.

Practitioner agreement to take part in the trial was sought in different stages. Settings were provided with a practitioner information sheet when they signed up to the project, which they were asked to circulate (see Appendix A.2). For collecting audio recording data at the beginning of the project, a specific information sheet was provided to those practitioners, and they signed a participation agreement form for taking part in this element (see Appendix A.3). The practitioner information sheet was provided again to practitioners directly when collecting the baseline survey data and agreement to participate in the project obtained before progressing onto the survey (Appendix A.2). For interview data collection, participants were provided with a short, interview-specific information sheet (see Appendix A.4 for champion

interview settings and Appendix A.5 for case study setting) and permission to take part in the interviews was obtained by email in advance of the interview and confirmed orally before starting the interview.

The study did not involve the recruitment of children to the project, however, the audio from children was included in the audio recordings of storytelling practice, which involved a nursery practitioner reading a storybook to a small group of children with the children responding to the story and to any prompts given by the practitioner. Settings were asked to circulate a brief parent information sheet (Appendix A.6) to parents in advance of the audio recordings and parents were given the opportunity to opt their children out of taking part in the audio recordings by contacting the setting. Durham's ethics approval for using opt-out consent for children for the audio recordings was contingent on children's data not being transcribed, analysed, or processed further.

Data protection

Participants were provided with a separate privacy notice as well as the information sheets described above. The privacy notice (included in Appendix A.7) described the purpose of the project and data processing to evaluate how the TWiTCH programme was delivered and the impact it had on nursery staff's confidence, knowledge, and practice. It described how some personal data would be collected directly by the evaluation team at Durham University and some would be collected by the Sheffield Hallam University team and shared with Durham—both teams would act as data controllers. It described that the project would collect personal data through documents completed as part of the project, surveys, training and delivery records, audio recordings of story-time, interviews, observation of TWiTCH training, mentoring visits, and TWiTCH practice. This data would be stored securely at both of the above universities and on approved survey software websites, transcription services, and audio recorders. It set out that the lawful basis relied on for this project was 'public task' (GDPR Article 6 (1) (e): the processing is necessary for an activity being carried out as part of the university's public task, which is defined as teaching, learning, and research and that the project was carrying out research. It informed participants that the retention period for the data would be five years after the end of the project (April 2030). It also described how, during the project, personal data would be shared between the SHU team and the Durham University team, and with the evaluation researcher who was based in the University of Manchester, for the purposes of delivering the project.

A Data Protection Impact Assessment (DPIA) was conducted at the start of the project, describing why the collection of personal data is necessary, the data that will be collected, and the ways that risks will be minimised. A data management plan was also in place to set out the procedures for the team to follow to ensure data confidentiality throughout the project. A data sharing agreement was put in place between the developer team at Sheffield Hallam University and the evaluation team at Durham University to ensure that both teams handle data correctly and that data can be shared securely and appropriately.

Project team

Developer team

The developer team was responsible for the recruitment of settings to the project and for the delivery of the TWiTCH programme.

Fufy Demissie (SHU) Fufy was the TWiTCH project director who developed the programme and materials and oversaw its delivery. She was responsible for training and managing TWiTCH coaches, providing champion training, and facilitating the network meetings during the project.

Sally Pearse (SHU) Sally was the project's strategic lead feeding into the original bid and advising on the development and delivery of the programme and the recruitment of settings.

Louisa Pinder (SHU) Louisa was the project manager responsible for managing the day-to-day delivery of the TWiTCH programme. Louisa was the main contact for settings during recruitment and delivery.

TWiTCH coaches

TWiTCH coaches provided training, mentoring, and support to five nursery settings each during the project. Coaches were:

Katarzyna Fleming (SHU).

Sarah Gordon (SHU).

Kate Halliwell (external consultant).

Ann Dawson (external consultant).

Topsy Page (external consultant).

St Edmund's Early Years Stronger Practice Hub supported with recruitment to the project and the retention of settings throughout the project.

Evaluation team

Vic Menzies (Durham University): Vic was the principal investigator on the project and with overall responsibility for the project design and protocol, project delivery, and reporting and project management.

Johny Daniel (Durham University): Johny was responsible for development of the audio recording coding measure and analysis of this data. Johny also contributed to evaluation design, programme definition and description, and reporting.

Xiaofei Qi (Durham University): Xiaofei provided advice to the team as a specialist in early years practice as well as contributing to the research design and the development of the research tools and measures.

Helen Cramman (Durham University): Helen provided advice and expertise on theory-based evaluation, contributing to the design of the research and the use, interpretation, and reporting of contribution analysis.

Kelly Burgoyne (University of Manchester): Kelly provided specialist advice in early years language development and early years programme development throughout the project, contributing to the development of the audio recording coding measure and other research tools as well as the reporting of the project.

Nadine Fitzfussell (Durham University): Nadine was the project's research associate and managed the day-to-day delivery of the evaluation, drafted data collection tools, conducted data collection (observation, interviews, surveys), analysed data, and contributed to the report writing. Nadine was on adoption leave between March 2024 and August 2024.

Paivi Eerola (Durham University): as research assistant, Paivi contributed to the project during Nadine's adoption leave (March to August 2024) by coordinating and undertaking data collection, analysing data, and contributing to the report.

Methods

Recruitment

The initial target for recruitment was 25 nursery settings. This number allowed for possible 20% attrition as experienced in some early years trials. The developer team was responsible for recruiting maintained and PVI nursery settings to the project. This was done through establishing contact with local authority leads in the recruitment areas, through an advert on the EEF website, through the personal networks of the developer team, and through the St Edmunds Early Years Stronger Practice Hub. Settings responded initially with an expression of interest for taking part; if they fitted the recruitment criteria they were invited to complete a full memorandum of understanding. Settings were recruited from areas of high socio-economic disadvantaged from West and South Yorkshire.

Eligibility criteria were that settings:

- offered funded places for three- and four-year-olds;
- had a minimum of three practitioners working with preschool cohorts;
- were not participating in other programmes which involved CPD to support children's language development; and
- were not participating in the Early Years Professional Development programme.

It was aimed to recruit around 50% PVI settings and 50% maintained settings. Recruitment prioritised those with higher levels of disadvantage judged by the postcode decile ranking using the Indices of Multiple Deprivation (Ministry of Housing, Communities and Local Government , 2019).

Twenty-six settings were initially recruited to take part in the pilot evaluation but one school-based setting withdrew from the project before the recruitment process had closed. A second setting (a PVI) withdrew a few months into the project due to difficulties with having the right space for delivering the programme. Of the remaining 24 settings, 12 were PVI nurseries and 12 were maintained nurseries. The design allowed for 20% attrition (specified within the study plan) and attrition was well within the anticipated attrition rate from the study plan.

Participants in the evaluation were nursery practitioners working with preschool (aged three-to-four years) groups in the recruited settings. Consent was sought from practitioners at each point of providing data for the evaluation. Children in the nursery settings were not directly involved in the study other than their voices being part of the audio recordings of storytelling.

Data collection

Design

Contribution analysis is used as the underpinning evaluation framework for the study (Mayne, 2008; TASO, 2022). This approach is used to draw robust conclusions as to the contribution of the activities stated within the theory of change towards achieving the intended outputs and impacts of the model.

Contribution analysis explores attribution through assessing the contribution that a programme is making to observed results (Mayne, 2008). To infer causality, four conditions must be met (Mayne, 2012):

Step 1: the programme is based on a reasoned theory of change;

Step 2: the activities within the programme were implemented as planned;

Step 3: the expected chain of results has occurred; and

Step 4: other contextual factors influencing the programme have not made a significant contribution.

For Step 1, the developer team created an initial theory of change model for the TWiTCH programme before the involvement of the evaluation team. This theory of change included a causal assumptions log and an indication of the

strength of existing evidence for these assumptions and also a contextual assumption log highlighting the context assumptions that would need to be met for the TWiTCH programme to be implemented in nurseries and to deliver the expected outcomes. This original theory of change model and logs were then interrogated through a series of initial workshops and set-up meetings between the EEF, the developer team, and the evaluation team with the evaluation team taking ownership of the documents and revising in line with the discussions (see Appendix C.1, C.2, and C.3).

For this pilot project, the causal pathway focus was at the practitioner level focusing on whether the activities within the theory of change were implemented as planned (Step 2) and whether the specified outputs and outcomes for practitioners occurred as part of the project (Step 3). More robust data collection was designed to make sure this was captured. We did also aim to capture evidence indicating the causal pathway for children but this was less robust when looking at children's outcomes.

Based on the theory of change and assumption logs, a framework of data collection methods was then developed to gather evidence covering elements of Step 2, Step 3 and Step 4. This included mixed methods data collection using a pre-post design along with some longitudinal aspects of data collection allowing the comparison of patterns of implementation across the participating settings. In a smaller number of purposely sampled settings, more detailed case studies of the delivery of the TWiTCH programme were also conducted, using the setting as the case level to understand the delivery specifically to each case.

Once the data collection was completed, the original theory of change model was then revisited by the evaluation team to assess whether the data collection throughout the project supported the theory of change expected pathway, including evaluating the evidence for each aspect of the theory of change and highlighting the extent to which each was supported by the collected evidence. A series of tables providing a narrative review of the evidence supporting each aspect were then created and are included in Appendix C.5, C. 6, and C.7. Examination of the data collected around the contextual factors allowed us then to make claims about the extent to which the TWiTCH programme contributed to the changes to outcomes seen as part of the project or whether other factors are likely to have significantly contributed. As a result of this synthesis and analysis, suggestions are made to improve the theory of change model as a result of the evidence collected (Appendix C.4).

Methods

Table 3 summarises the planned data collection and research methods used in this study. Further detail about the sample for each aspect actually collected is detailed in the description below.

Table 3: Summary of data collection and analysis

Research methods	Data collection methods	Evaluation dimension	RQ	Planned sample size and sampling criteria	Actual sample size	Data analysis methods
Administrative data	Training attendance and delivery records	Fidelity	2, 6	As available from developer	Provided for all events	Descriptive statistics
	Audit/feedback, fidelity records from coaching visits to settings	Fidelity	2, 6, 7, 8	As available from developer	Logbooks received for 24 settings (1 incomplete)	Deductive coding
	Data available about nurseries from online sources	Context/moderators	1	As available—data sought for all settings	As available	Descriptive statistics
	Setting context and current practice from leadership—from MoU sign-up form	Context; programme differentiation	1, 9	25 nursery managers/head-teachers (all)	25 received	Descriptive statistics, deductive coding
Usual Practice Surveys (pre/post)	Online questionnaire to capture practitioner characteristics as well as usual practice, confidence in supporting children's language development, knowledge and use of	Programme differentiation; perceived impact; context/ moderator, cost	2, 4, 6, 9, 10, 11, 12	All nursery practitioners of 3- to 4-year-olds in 25 settings	Pre intervention n=84 Post intervention n=34	Descriptive statistics, inductive of open text responses

	TWiTCH techniques, other similar training, changes in practice and confidence—specific questions if TWiTCH champion				Matched pre/post responses n=32	
Training Survey	Short online/paper questionnaire after training sessions to capture programme aligned knowledge gains, confidence in programme delivery, acceptability/appropriateness of the training, feedback on training	Responsive-ness; perceived impact	4, 11	All nursery practitioners who attended training	N=49 from 25 settings	Descriptive statistics, subgroup regression analyses
Observation	Semi-structured observation of coach training	Fidelity, adaptation	3	Initial coaches training session/s	Observations of 2 sessions	Deductive coding
	Semi-structured observation of twilight training sessions to capture engagement with training and consistency of delivery	Fidelity; quality; responsive-ness	3,4	5 case study settings (one from each coach) selected to represent a variety of setting characteristics	Observations of training sessions in 5 case study settings	Deductive coding
	Semi-structured observation of mentoring sessions	Fidelity; quality; responsive-ness;	3, 6	5 case study settings	14 observations of 5 case study mentoring sessions	Deductive coding
	Semi-structured observation of TWiTCH champion online network events	Fidelity; responsive-ness; adaptation	6, 7, 8	4 online meetings and those that attend	Observation of one session from each of 4 meetings	Deductive coding
Audio observation (pre/post)	Structured and coded observations of practitioner interaction practice during story time, audio recorded by practitioners' settings. Baseline to take place June/July to be comparable with June/July at post-intervention.	Fidelity; evidence of promise	10	2 x 10–15-minute recordings of practice from one practitioner in each of 25 settings at 2 timepoints. Setting selects 1 practitioner at beginning of the project willing to do recordings and who is likely to be around 1 year later.	Pre-intervention 37 recordings received Post-intervention 41 recordings received Matched pre and post recordings (practitioner or setting) 44 recordings coded	Coded for balance of child/adult participation, use of dialogic tools, use of storytelling pedagogy frequency counts
Interviews/ focus groups	Semi-structured interviews (two timepoints) phone/online	Fidelity; adaptation; responsive-ness; perceived impact; cost; context/moder ators	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12	TWiTCH champion in 12 settings chosen to represent a variety of setting characteristics.	1st interview n=10 (discussing 12 settings) 2nd interview n=9 (discussing 11 settings)	Both deductive coding based on RQs and inductive coding will be used for understanding the data within RQs and for new emerging themes
	Semi-structured interviews (three timepoints) in person during case study visits	Fidelity; adaptation; responsivenes s; perceived impact; cost; context/moder ators	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12	Short interview with up to 4 staff in each setting: nursery manager/headteacher, TWiTCH champion, room lead, practitioner at each of 5 case study settings	1st timepoint n=15 2nd timepoint n=11 3rd timepoint n=15	

	Semi-structured interview	Fidelity; quality; responsive-ness; adaptation	2, 3, 4, 5, 6, 7, 8, 10	2 with developer team—2nd including coaches	1 coach focus group (n=5) and 2 x developer team (n=3 people interviewed at each point).	
Cost workshops	Semi-structured workshops to establish cost to deliver intervention	Cost	5	2 with developer team (one early to establish ingredients model after intervention finalised and encourage collection of cost data and one towards the end to capture ongoing costs)	2 undertaken	Cost table for ingredients of intervention

Practitioner pre-post surveys

Baseline survey

Before any TWiTCH training was received, all practitioners working with three-to-four-year-old children in recruited settings were asked to complete a baseline survey (online before training, or a paper copy at the beginning of the training session: see Appendix B.1). The survey was used to establish the confidence and current practice across all participants before the programme began to allow comparison with data collected after completing the TWiTCH programme. Eighty-four practitioners from 24 settings completed this survey. This survey collected measures of practitioners' confidence and practice in supporting early language development generally, practice during story-time, and their knowledge of specific TWiTCH techniques. It also collected practitioner background information including level of early years qualification, length of experience working in early years, and training they had received over the previous two years.

The survey included a version of the Practitioner Confidence and Skills in Supporting Children's Language Measure (Lindorff, Sylva, Ereky-Stevens and Joseph, 2022) adapted so as to maximise the relevance of the measure to the TWiTCH programme, omitting items not relevant to the TWiTCH theory of change model and adding in a number of additional items specifically relevant to storytelling practice and TWiTCH pedagogies. These adaptations were made by the evaluation team in consultation with the developer team. The survey was piloted with five practitioners from four pre-pilot nursery settings in the North East of England in June/July 2023 to make sure the new items worked and that the survey was appropriate and acceptable to practitioners.[2] Pilot settings included one PVI, two integrated school nurseries, and a standalone maintained nursery.

The survey provided three scales used in our analysis: (1) practitioners' confidence in supporting children's language skills, (2) practitioners' usual practice when interacting with children in the nursery, and (3) practitioners' usual practice during storytelling sessions. For items in each scale, participants responded using a five-point Likert rating to rate level of confidence for scale one and frequency of activity for scales two and three. Scale reliability was calculated using the baseline data (N = 84)—see Appendix D.2 for summary of scales—with each scale showing high reliability of 0.88 or above. Three additional items assessed practitioners' knowledge and understanding of three elements of the TWiTCH programme: 'interactive book reading and book talk' (dialogic reading), 'sustained shared thinking', and 'choice-based language activities'—again using a five-point Likert scale response.

Post-intervention surveys

Following the intervention, practitioners were asked to complete a second version of the baseline survey (June/July 2024: see Appendix B.3). In addition to the confidence, practice, and knowledge items described in the baseline survey above, practitioners were asked about their role in TWiTCH, any other training since starting TWiTCH to support

[2] In the study plan we had aimed to pilot the survey with ten practitioners, however the nursery settings recruited for piloting were smaller than anticipated and only one member of staff was available to complete the survey in three of the settings.

children's language development, the TWiTCH books they had delivered, and about the perceived impact of the programme. This survey was completed by 34 practitioners across 23 settings. This survey had lower numbers of participants compared to the baseline survey due to the fact that not all practitioners working with three-to-four-year-olds who completed the baseline survey then attended the TWiTCH training or delivered the programme: the post-intervention survey was only completed by those involved with TWiTCH. The survey was completed by at least one practitioner in all except one nursery.

Analysis comparing baseline and post-intervention surveys used only the sample that completed both surveys (N = 32), which gave the most stringent conditions for comparisons and maximised power in the analysis.

Post-training surveys

Practitioners were asked to complete a short online questionnaire specifically about the TWiTCH training they had received (Appendix B.2) in November 2023 for those that had completed training in September and October, and in January 2024 for those who completed training in November and December. It was expected the practitioners would have had time to deliver at least one TWiTCH cycle at the time of completion. The survey aimed to gather feedback about the pitch and accessibility of the training and to capture knowledge and confidence in using the TWiTCH strategies. Most items used Likert scale or closed responses but a few provided opportunities for qualitative feedback about the training sessions and what could have been improved. At least one staff member from each participating nursery responded to the survey (n = 49) included 15 nominated TWiTCH champions, five non-champion senior leadership members, and 34 non-champion nursery practitioners.

Audio-observed story-time practice (pre-post)

Audio recordings of story-time sessions before and after the intervention were used to look at whether taking part in TWiTCH changed the practice of nursery practitioners when reading stories to the preschool children. This was seen as a more objective way to gather this data, which could be triangulated with practitioners' self-report of practice change in the surveys. The focus on practitioner speech rather than on child response was both a purposeful and pragmatic choice. TWiTCH as a CPD programme was intending to change practitioner interactional behaviour when reading storybooks to children and without this change in behaviour the programme would be unlikely to yield the expected impact on children: it was important, therefore, to look specifically at practitioners' speech. Methodological constraints limited coding of child responses: audio quality often made it difficult to distinguish individual voices, especially in group settings with overlapping speech, risking misinterpretation of participation levels and potential bias in interactional counts. Analysing practitioner talk alone provided a more consistent and valid measure of interactional change aligned with the intervention's focus on practitioner behaviour. Pragmatically, it would also have been an additional burden and challenge to also obtain opt-in consent from children taking part in the recordings in order to process their data and use it for the research and would have made it less likely to have obtained this data.

All participating settings were asked to audio record two sessions of a practitioner sharing a storybook with a group of up to eight children in the setting before any training or activities for the TWiTCH programme—June and July 2023—and again at the end of programme—June and July 2024. To do this, settings were provided with an audio recorder which they placed near them while doing story-time. Settings were given the flexibility to choose which practitioners completed the audio recording but it was asked that, where possible, the same one completed recordings at the beginning and end of the programme. At the end of the programme, it was specified that one recording should be of a TWiTCH book while the other should be of any storybook. Practitioners were asked to capture both the story reading and any discussion about the book in advance or following the story. The recordings were with children who were starting school the following September meaning that different children participated at both timepoints. However, this timing allowed control for the age of the children and the ways that practitioners interact with them; it may be that as children get older and more proficient with language practitioners use different strategies in their interactions. Practitioners completed a proforma to go with the recording providing their name, the name of the story, the number of children taking part, whether the story was familiar to the children and whether the recording was typical of their usual story telling practice (Appendix B.4).

The procedure for audio recording story-time in nurseries was piloted with one practitioner in each of the four settings recruited for piloting the baseline survey. This ensured that the use of the audio recorders would be acceptable to nurseries, that the placement of the audio recorders would work to pick up the practitioner and children's voices clearly enough to allow coding, and that our guidance to practitioners was clear enough so that the recordings were what we

wanted. Piloting of this procedure highlighted some areas for improvement but indicated that the audio recording process would work for the study.

A structured coding sheet was developed to observe the frequency of practitioner use of specific dialogic tools during story telling (Appendix B.5). This coding sheet was based initially on the ELLECCT tool (Weadman, Serry and Snow, 2022), which provides a structured framework to score early years practitioner's shared book-reading practice. This tool was adapted for use with audio data only and to capture additional pedagogical strategies specifically promoted by TWiTCH. It specifically covers practitioners use of prompts (open-ended, closed-ended, distancing), TWiTCH related teaching pedagogical tools, print instruction, use of phonological awareness strategies, and vocabulary instruction. The coder records the frequency of use of these strategies during the recording. The length of the session and duration (and percentage) of practitioner talk during the story reading and engagement related to the story was also captured to be used as a proxy for the engagement of children in the session with a lesser percentage of teacher talk indicating more child engagement.

The tool was developed and trialled through an iterative cycle of tool development, comment, and feedback from experts on dialogic reading and language and literacy development in the evaluation team and piloting the tool with data collected from the four practitioners during the audio-recording pilot phase. This was done to make sure the items that were being coded were appropriate, the scale was feasible to code for, and that reliability between coders could be achieved.

Thirty-seven audio recordings from 20 settings were received before the intervention delivery and 41 recordings from 22 settings were received at post-intervention.[3] Thirty-two were completed by the same practitioner at baseline (16) and post-intervention (16); the coding of these audio recordings was prioritised due to these providing the most high-quality comparison. An additional 12 audio recordings were coded where baseline (six) and post-intervention recordings (six) were available from the same setting. This resulted in the coding of 44 audio recordings in total (22 baseline and 22 post-intervention) with data from 17 nursery settings, and 28 practitioners with some settings having two recordings at baseline and post-intervention. Recordings were received from three settings rated 'outstanding', 11 rated as 'good', and one rated as 'requires improvement' and two rated as 'inadequate' on their latest Ofsted reports. The project role and setting type of the participants for the audio-recordings analysed is included in Table 4, split by baseline and post-intervention recordings.

Double coding was conducted on seven of the 44 coded audio recordings (15.91%) to assess inter-observer agreement (IOA). Using the item-by-item agreement formula, the average IOA across all items was calculated to be 96.38%, indicating a high level of reliability in the coding process.

Table 4: Project roles and setting types for audio recordings analysis (n = 44)

	Baseline recording				**Post-Intervention recordings**			
	Champion	**Senior leader**	**Practitioner**	**Unknown**	**Champion**	**Senior leader**	**Practitioner**	**Unknown**
Role in project	7	2	8	5	10	2	8	2
	PVI		Maintained		PVI		Maintained	
Practitioner's setting type	8		11		8		11	

Case studies

Case studies were conducted in five purposively sampled settings, one from each coach, reflecting a range of setting types, levels of deprivation, and Ofsted ratings. These were chosen before settings began the TWiTCH programme with

[3] Some settings were not recruited to the project until July 2023 so it was not possible for these settings to complete pre-delivery audio recordings before closing for their summer break. Two settings that opened for the first time in September 2023 were also not able to complete audio recordings.

the aim of capturing the in-depth experience of delivering the intervention across different contexts and to gather a more holistic view of programme delivery.

In these settings, TWiTCH training and mentoring visits were observed and interviews with staff delivering the programme were conducted at three timepoints for four settings and two timepoints for one. Staff interviewed during the visits included the TWiTCH champion, at least one other practitioner, and a member of the senior leadership team (when not the TWiTCH champion). An observation of TWiTCH practice was usually conducted on each visit, which was observed by the evaluation team and the TWiTCH coach. Case study settings were provided with an additional incentive of a £200 voucher for an online early years equipment provider to compensate them for the additional burden of these elements.

Observations of TWiTCH training and ongoing support

Coach training

Semi-structured observations of the two sessions of coach training were completed by one member of the team. This allowed the evaluation team understanding of the programme and what was expected from the coach delivery during the programme.

Training and support for settings

As noted, semi-structured observations were conducted of the CPD provided by each of the coaches in the five case study settings (observation schedules can be found in Appendix B.6 and B.7). Observations were completed of the initial training sessions as well as the three in-person mentoring sessions in each setting.[4] Of the 25 sessions observed, six were observed by two researchers allowing for discussion and agreement to be reached in the coding.

These observations investigated whether the TWiTCH programme was adequately defined to allow training and support to be delivered in a consistent manner across the different coaches and different types of setting. Observations noted the extent to which the CPD sessions were consistent with the initial coach training and the pre-specified plan for the session. Observations also captured the engagement of participants with the training and mentoring.

Training and support for TWiTCH champions

Observations were conducted of the two online training events for TWiTCH champions and one session of each of the three network meetings (two sessions were held at different times of the day for each network meeting). These observations captured the engagement of participants in these sessions, how closely the sessions stuck to the plan, and the reflections and questions asked by the participants.

Practitioner interviews

Semi-structured interviews with practitioners at two (remote interview settings) or three (case study settings) timepoints through the TWiTCH delivery period investigated how the setting has delivered the programme through the year, fidelity of delivery, any difficulties with delivery as well as gathering practitioner views of the acceptability and impact of the programme. Interview schedules can be found in Appendices B.8 to B.16.

Remote interviews

Twelve settings were selected for remote TWiTCH champion interviews immediately after recruitment, reflecting a range of setting contexts considering setting type (PVI, school-based, children's centre), size, level of deprivation, location, and Ofsted rating. Telephone interview settings received an additional £100 voucher for early years equipment as a thank you for taking part in the interviews.

These interviews took place by phone or using online video conferencing with the TWiTCH champion in the 12 pre-selected settings with all interviews recorded and transcribed. Interviews were expected to last between 15 and 20 minutes; the average interview length was 22 minutes. Table 5 shows the characteristics of the selected settings: three of these settings were part of a preschool academy chain with the same TWiTCH champion for all settings; one longer

[4] Only two mentoring sessions were observed in one setting due to train strikes making travel impossible.

interview (around 30 minutes) at each timepoint discussed all three settings. In total, ten interviews took place in January 2024 and nine in May and June 2024 (one TWiTCH champion had left the setting between interviews).

Table 5: Remote interview setting characteristics

Setting type	**PVI**	**Maintained**		
	7	5		
Level of deprivation	**High**	**Moderate**	**Low**	
	5	2	5	
Ofsted Rating *	**Outstanding**	**Good**	**Requires improvement**	**Inadequate**
	2	7	1	1

* One settings had not been inspected by Ofsted at the time of writing

Case study interviews took place in person during setting visits at the time of the coach mentoring visits (timing varied depending on the coaching visits and when the setting had started TWiTCH). Interviews were conducted with the TWiTCH champion, a member of the senior leadership team (when not the TWiTCH champion), and at least one practitioner who was also delivering TWiTCH (where available). Table 6 shows the number of participants interviewed at each timepoint: participants were the same at each timepoint when available.

Table 6: Number of case study participants interviewed at each timepoint.

	TWiTCH champion (n)	**SLT (non-champion) (n)**	**Practitioner (n)**	**Total (N)**
T1	5	4	6	15
T2	4	3	4	11
T3	5	4	6	15
Total (N)	14	11	16	

Developer team interviews and focus group

Two semi-structured interviews with the developer team were undertaken in December 2023 and September 2024 to capture the developer team's views of the TWiTCH programme delivery across the project, any challenges faced, any changes they made to the programme during delivery, and how actual delivery compared to their expectations. They also covered the essential TWiTCH components and those that could be adapted to feed into a future measure of fidelity. These interviews were audio recorded and transcribed.

A focus group with four TWITCH coaches was also undertaken in January 2024 to gather the views and reflections of the coaches on their training and how settings were delivering TWiTCH. One member of the coaching team was unable to attend the focus group but shared their views via email.

Administrative records

Records of practitioner attendance at CPD and support sessions were provided by the developer team and were used to measure the extent to which settings complied with the training and CPD elements of the programme.

Coach logbooks for each of the 24 settings remaining in the project were provided by the coaches containing notes from each of the training and mentoring sessions undertaken at each setting although for a few settings these were not complete for all visits. It was expected that these logbooks would provide detailed information about the delivery of TWiTCH in each of the settings, which would allow the creation of a measure of fidelity. However, the logbooks were completed in significantly different ways by each of the coaches and did not capture the quality or detail of the delivery for most settings. Although coaches were provided with a template for the logbooks in the TWiTCH coach handbook, which gave categories of information to fill in at each timepoint and questions to ask practitioners, the level of detail

provided for each heading was variable and not all headings were filled in by coaches for each visit. The logbook template also did not specifically ask for details on the ways that settings were delivering TWiTCH, and this was often not provided by coaches. All logbooks were analysed to look at details of those attending the training and details of the mentoring sessions delivered, however, not for fidelity of the setting to the TWiTCH approach.

Cost evaluation

Cost data was collected in two cost workshops with the delivery and evaluation teams. The first semi-structured workshop was conducted early in the programme (November 2023) to explore the cost ingredients of delivering the programme. This was organised after the TWiTCH delivery model had been finalised by the developer team. A second workshop, later in the project (September 2024), revisited the cost ingredients and aimed to capture ongoing costs beyond the evaluation and those associated with delivering TWiTCH as a three-year model. Both workshops established the known costs to the developer team as well as the estimated costs of delivering the programme.

Costs associated with time and resources were explored for the initial set-up of the programme in a typical nursery setting taking part. This includes the training of coaches and of practitioners (including the costs claimed for the cover to staff attending training), the resource packs required to deliver the programme, and the time for settings to plan and adapt for the delivery of TWiTCH in their unique settings. Additionally, ongoing costs for practitioner support and TWiTCH champion training were estimated and the ongoing communication between coaches, nursery settings, and the developer team. The workshops also explored the costs that settings would incur for participating in the programme.

Costs faced by nurseries in implementing TWiTCH were explored during the interviews with nursery staff as well as the price that a nursery may be prepared to pay for the programme outside of the research project.

Data analysis

Practitioner pre-post surveys

Analysis was undertaken using the 32 survey responses where matched data was available at baseline and post-intervention. This excluded two post-survey responses where the participants had joined the nursery during the year and had not completed the baseline survey.

Descriptive statistics are presented under the relevant research questions for the scores on the three scales of practitioner confidence:

1. practice when interacting with children;
2. story-time practice; and
3. for the three items indicating knowledge of the TWiTCH techniques—interactive book reading, sustained shared thinking, and choice-based language activities.

Data is explored by setting type (PVI or maintained), role in the project (champion, SLT, or practitioner), and level of experience categorised into three levels: 0–6 years, 7–12 years, and over 12 years. Baseline and post-intervention scores on the quantitative scales are compared using paired sample t-tests to look for significant differences between the scores. Exploratory hierarchical multiple regression was conducted with the dependent variable the difference in scores between post-intervention and baseline for each of the scales. The predictor of setting type (PVI or Maintained) was entered first into the model due to the expectation that the intervention would run differently in each type of setting and that the level of training staff received around language development would be higher in maintained settings. The predictors of length of early years' experience (continuous variable) and staff qualification (dichotomised so that level 3, level 4, and level 5 were in as one group and level 6 and QTS were a second group) were entered in a second step. The regression analysis was underpowered with only 32 matched responses to the survey (when 20 per predictor would be the recommended minimum number to conduct multiple regression with multiple predictors: Harrell, 2015) so results should be interpreted cautiously. Analyses were conducted using IBM SPSS Statistics version 29 (IBM Corp., 2023); the syntax is included in Appendix D.1.

Qualitative survey responses were coded inductively using thematic analysis. This was done in Excel with a researcher first reading through all responses to a question three times to understand the data and then identifying the key points in each response and assigning a descriptive code to these key points. Once all responses had been coded, a review of the codes was carried out and the responses to each considered to check that the code was appropriate.

Consideration was then made of whether to merge codes which covered similar themes. Review of the codes with a second member of the evaluation team was then completed to finalise the coding.

Post-training survey

Descriptive statistics are presented for the responses to the post-training survey. Qualitative responses were coded inductively using thematic analysis.

Audio-recording data

Three measures of duration were taken per audio recording: the full duration of a recording, practitioner talking time during storytelling, and practitioner talking time around storytelling (that is, before and after storytelling, combined). The average duration was calculated (in seconds) across all 44 recordings for baseline and post-intervention recordings, for each of these three duration types. An effect size calculation was performed using Cohen's d and confidence intervals calculated.

The number of occurrences of different question types and dialogic reading practice codes was averaged across the 44 audio recordings, for baseline and post-intervention audio recordings. Exploratory analysis explored the average use of each code by type of nursery setting, practitioner experience in early years (in three groups), and by practitioner qualification (two groups). Frequency recording is one of the most used measures in applied behaviour analysis to observe behaviour (Cooper, Heron and Heward, 2014).

Interview and focus group data

All data collected in interviews and focus groups was analysed using both deductive and inductive coding in a multi-step process as described by Bingham (2023). Transcripts were initially coded deductively using pre-developed codes based on the project's research questions to organise the data. Where data did not fit within these codes, new codes were generated. Inductive open coding was then used within the deductive coding categories to add more nuanced coding within the pre-existing codes and to support understanding of the data. Themes within the research questions were then generated through further inductive analysis of the codes generated. Qualitative coding and analysis of interview data was conducted using NVivo v15 (Lumivero (2024).

Observation data

Observation data from training and mentoring sessions was analysed deductively to look at level of consistency in delivery of training sessions between coaches, fidelity to the programme, and engagement of participants in the training and mentoring sessions.

Timeline

Table 7 outlines the timeline for the project.

Table 7: Timeline

Date	Activity
May 23–Jul 23	Recruitment of evaluation settings
May 23–Jul 23	Piloting evaluation tools (audio recording and surveys)
Jun 23–Jul 23	Observation of coach training
Jul 23–Sep 23	Baseline audio recordings
Sep 23	Baseline usual practice surveys
Sep 23–Nov 23	First twilight training session observation Second twilight training session observation
Oct 23	Delivery of TWiTCH programme begins
Nov 23, Jan 23	Post-training survey
Dec 23	First developer interview and cost workshop
Nov 23–Mar 23	Observation of first coaching visit (5 settings) First case study visits (5 settings)
Jan 24	First telephone interviews with TWiTCH champions (12 settings)
Jan 24–Mar 24	Observation of second coaching visits (5 settings); second case study visit (5 settings)
Feb 24	First feedback workshop between evaluator and developer
May 24–Jun 24	Second telephone interviews with TWiTCH champions (5 settings)
Jun 24	Second feedback workshop between evaluator and developer
Jun 24	Observation of third coaching visits (5 settings); third case study visit (5 settings)
Jun 24–Jul 24	Post-intervention audio recordings of practice Post-intervention usual practice surveys
Aug 24–Nov 24	Data cleaning, coding, and analysis
Sep 24	Second developer interview/cost workshop
Oct 24–Dec 24	Report writing
Dec 24	Draft report submission

Findings

Participants

Twenty-five nursery settings began in the project although one medium-sized outstanding PVI setting withdrew part way through. Setting characteristics are included in Table 8.

Table 8: Setting characteristics

Decile score for deprivation *	**1–2**	**3–4**	**5–6**	**7–8**
Number of settings	9	7	4	4
Nursery type	PVI	Maintained		
Number of settings	12	12		
Ofsted Rating	**Outstanding**	**Good**	**Requires improvement**	**Inadequate**
Number of settings **	3	15	1	2
Location **	**Urban**	**Semi-rural**	**Rural**	
Number of settings	18	1	3	
Intake	**Mean**	**Minimum**	**Maximum**	**SD**
Number of children registered in pre-school year at beginning of project	36.5	11	100	21.3

• From Index of Multiple Deprivation (Ministry of Housing, Communities and Local Government, 2019) using nursery postcode (lower decile indicates high deprivation).
** Ofsted and location data not available for two settings.

Of the 24 nurseries with continued participation in this study, half were PVI and half were maintained; the majority of settings had been given a 'good' Ofsted rating. Most nurseries were in in urban areas although four were rural or semi-rural. Half of the nurseries taking part were from the most deprived areas (lowest three deciles), while eight experienced moderate deprivation (middle three deciles) and four were from less deprived areas (deciles seven and eight). Nurseries ranged notably in size from 11 to 100 children registered to start in their preschool year at the start of the project: 29% had 20 or fewer preschool children registered, 46% had between 21 and 40 children registered, and 25% more than 40.

While 84 practitioners responded to the baseline survey, only 34 completed the post-intervention survey; many of the practitioners responding to the baseline survey did not go on to attend TWiTCH training or deliver TWiTCH in the setting.[5] Thirty-two respondents completed both the baseline and post-intervention survey and the analysis in this report is based on the data from these 32 respondents. Participant characteristic data for both the full baseline survey sample and the matched baseline and post-intervention sample is included in Appendix D.3. This data indicates a similar pattern of experience and qualification in the matched subgroup to the full baseline sample although for qualification level the matched group excludes those with the lowest level of qualification. There was a range of early years' experience for

[5] All staff working in the preschool (three- and four-year-old) room(s) were asked to complete the baseline survey, however, nurseries did not find it feasible to release all of these staff to attend the TWiTCH training. This was sometimes due to needing to pay staff extra to attend for training outside of their normal hours or difficulty in covering nursery provision. Some settings also just planned TWiTCH to be delivered by a small number of practitioners and did not see the need for all staff to attend the training.

practitioners with around half of the sample having a high level of experience (12 or more years) in early years settings. The largest group of participants were qualified to level three or four although just over a third of the sample were qualified to degree level. Around half of participants were based in the participating PVI settings and half in maintained settings of which 13% were based in nurseries in children's centres rather than in schools.

The majority of participants indicated that they had not taken part in any training to support children's language development since 2021, before doing TWiTCH.

TWiTCH champions

The survey data provided background characteristics for the 22 practitioners nominated as TWiTCH champions (Table 9). They came from a range of roles across the nursery and levels of education and represented a range of experience working in the early years.

Table 9: Characteristics of TWiTCH champions

Role in setting	EY practitioner	Middle leadership	Nursery teacher	Senior leader (SLT)		
Number of practitioners *	7	3	7	3		
Highest level of EY Qualification	Level 3 or 4 Childcare	Level 5	Early years teacher	Qualified Teacher Status	Level 6	Level 7
Number of practitioners	6	4	3	7	1	1
	Range	Mean	SD			
Early years experience	0-23 years	10.41 years	7.07			

* Role in setting data missing for two participants who completed the survey.

Findings related to research questions

Feasibility of trial

RQ1 To what extent does the recruitment strategy recruit and retain targeted disadvantaged settings and practitioners throughout the study?

a. What are the characteristics of settings that are recruited to the project?

b. To what extent are settings recruited from targeted areas of disadvantage?

Data used to answer this question has been synthesised from the developers' administrative records (for example, the MoU sign-up form), data available about settings from online sources (such as the nursery website or an Ofsted report), our own observations of each setting's participation in the TWiTCH programme, and the developer interviews.

The project's recruitment strategy was successful in recruiting the required number of settings to the project. The characteristics of those recruited are described in the Participants section above. Around half were from highly disadvantaged areas and another 17% from moderately disadvantaged areas. One setting withdrew from the project, however, this setting was not from a highly disadvantaged area; all others were retained throughout the project. Most of those recruited were Ofsted-rated 'good' or 'outstanding' with only three rated 'requires improvement' or 'inadequate': this may indicate a challenge with the recruitment strategy of engaging these latter types of settings with the project. There was engagement with all settings, except one, with the evaluation throughout the project; one setting did not engage with data collection at post-intervention and did not provide any reason for their lack of engagement. There was some attrition of practitioners from those who had originally been trained to do TWiTCH, throughout the project, with

practitioners getting new jobs in other settings or moving rooms in the nurseries and not continuing to deliver TWiTCH. However, this was only in a small number of settings.[6] This had the biggest impact when the practitioner leaving was the TWiTCH champion. In the evaluation, attrition between the post-training survey (for those that we know had engaged in the TWiTCH training) to endpoint reduced from 49 to 34 participants (attrition of 31%), however, we maintained representation from 23 of 24 settings in this data; we do not know how many of the 49 participants who responded to the training survey actually took part in delivering TWiTCH.

The developer team felt that recruitment to the Phase 2 pilot trial had been achievable through drawing on their own sector connections and through the support of the Stronger Practice Hub. They had additional interest from settings wanting to take part in the project from other less deprived settings but it was not possible to include these in the project. The team were of the opinion that for settings with a poor Ofsted rating, responding to Ofsted's feedback would take priority over participating in a new project.

To summarise, the project's recruitment strategy was successful in recruiting the required number of settings, with half from highly disadvantaged areas. Most of those recruited were Ofsted-rated 'good' or 'outstanding' with only three settings rated 'requires improvement' or 'inadequate'. Settings with low Ofsted ratings may have higher priorities than participating in a new project.

Delivery of training

RQ2 How feasible is the delivery of the programme within the delivery period?

a. Delivery of training.

b. Delivery of nine cycles of TWiTCH.

Data used to answer this question has been synthesised from the developers' administrative records (for example, training attendance and delivery records), training observations, coach logbooks, the post-training survey, post-intervention survey, and interview data.

Delivery of training

The delivery of the training programme was feasible within the delivery period. Both days of training were successfully delivered to all participating settings, though the time of delivery varied due to the availability of coaches and practitioners. It was hoped that training delivery would have been completed at the start of term (September-October). The first day of training was delivered to 18 settings in September, five in October, and two in November. Delays in the initial training were due to difficulties in communication between the nursery and the coach as well as challenges with finding suitable dates for the coaches and the setting. The second day of training was delivered to 15 settings in September, eight in October, and one in November.

Delivery of nine cycles of TWiTCH

The success criteria state that at least six TWiTCH cycles need to be completed by the majority of settings, which has been achieved. The average number of cycles completed by the June-July 2024 post-intervention survey was 7.9 (SD = 1.22); by then, the full nine cycles of TWiTCH had been completed by nine settings and seven had completed eight and so could theoretically complete a ninth cycle by the end of term. Seven settings had completed seven or fewer cycles—seven cycles: n = 3, six: n = 3, five: n = 1—and thus not likely to be able to complete the full nine cycles by the end of term.

When champions were interviewed in May and June 2024, six of the 11 interview settings had already completed nine cycles, with two of these already adapting non-TWiTCH stories and delivering these within their settings. Champions in two settings indicated they were planning to complete all nine cycles before the summer break.[7] One champion reported that their setting was on their fifth cycle and did not anticipate making it through all nine cycles. Champions in two

[6] We were informed of three settings where the TWiTCH champion or practitioners left the setting or changed room, however we didn't systematically collect this information.
[7] In one of the 12 interview settings the champion left the setting during the project, and it wasn't possible to interview another member of staff from this setting therefore interviews only covered 11 settings at the second interview timepoint.

settings were not sure how many cycles they had completed at the time of the interview. Reasons champions gave for not getting through all cycles of TWiTCH were:

- delays getting started—for example, not able to schedule the training until late in the year;
- not using TWiTCH books during celebration periods such as Christmas and Easter due to prioritising other texts and activities; and
- in some cases, practitioners did not feel that certain TWiTCH books were appropriate, particularly the case for Hansel and Gretl, which some settings chose to miss out due to the themes in the book particularly around the 'evil step-mother'. (Only 15 of the 23 settings responding to the post-intervention survey had delivered the Hansel and Gretl TWiTCH cycle and only 14 had delivered the Princess and the Frog cycle; other books were delivered by almost all settings.)

In summary, the delivery of the training and of the nine TWiTCH cycles was feasible within the delivery period, though not all settings achieved this. The pre-specified success criteria, 'at least six TWiTCH cycles completed by the majority of settings', was achieved.

RQ3 To what extent is the TWiTCH programme defined (manualised) so that training and support is delivered with consistency between settings and across different trainers?

Data used to answer this question has been synthesised from case study observations of training and mentoring visits and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

Training

All sessions of the TWiTCH training for practitioners were delivered by coaches and broadly followed the guidance for training delivery. Time spent on the different sections of the training varied, for example, not all the coaches gave practitioners the same time to practice the different tasks, and session length varied from one and a half to three hours. While coaches were provided with training slides (which included videos and audio clips) along with detailed notes and what to cover on each slide along with suggested timings, these were not always followed: coaches consistently missed out some items such slides on 'background research and evidence' and the detail of the specified content for each of the five days during Week 1 of the TWiTCH cycle. Sometimes this was due to coaches assuming that the practitioners already had this knowledge. An unexpected lack of IT facilities in nurseries also led to coaches making adaptations to the training—for example, not using AV and sitting around a laptop, or no slides at all. Video and audio activities were not consistently shown due to a lack of IT facilities. Providing coaches with guidance on how they should respond to IT issues as part of the coach training would support a more standardised approach to how the training is delivered in these circumstances.

Coach style affected how much attention was dedicated to some tasks. Academic versus consultant coaches approached and delivered the training differently. This affected the content and length of the training. Two of the three consultants added extra elements and delivered in line with their own way of delivering training and were more likely to fill the allocated time for delivery of the training. Coaches did not have the opportunity to observe other coaches, or to have their own training observed other than by the evaluation team; adding in some elements of observation and feedback for the coaches could assist with making sure that they cover the required content within their own coaching style and could help to support coaches with any areas they struggle with.

Practitioners and coaches thought it was important for coaches to visit the nursery setting to deliver training to understand the context and experience of practitioners. Coaches described how different settings had very different starting points in terms of their current practice and knowledge of dialogic reading and the TWiTCH techniques and that they needed to tailor the training they delivered to the starting point appropriate to the setting. In the first developer interview, the developer team members acknowledged that they had made certain assumptions about practitioners' ability to know how to use the provided TWiTCH materials, such as the word cards: 'Like even the sticks, the wooden sticks and things we didn't really talk to them as much about how they might use it and where, I think I just assumed that they'll know what to do.' This meant that sometimes coaches had to spend more time working with certain settings when the expected experience and knowledge was not there.

Support

When observing the first set of case study coach mentoring visits, the length of coach visits to nurseries varied between 35 minutes and three hours. The coach handbooks did not specify the length of these meetings, but the developer expectation was that they would be an hour. Most coach visits included an observation of one practitioner delivering TWiTCH to a small group of children of either a Week 1 or Week 2 activity. Some coaches were hands-on during the sessions, stepping in to model ways of delivery, whereas others only observed. Immediate feedback was given to the practitioner where possible by all coaches. The quality of the feedback observed was variable: while all practitioners were given positives about their delivery, in some cases they were not given feedback about areas to improve while in other cases they were given a long list of actions to take away. Interviews revealed that some practitioners did not feel they had been given things to work on. Some coaches tried to speak to all staff delivering TWiTCH and provided feedback to the champion. Practitioners reported being generally happy with the level of support that they received and felt that the coach role was important.

To summarise, training delivery was quite different between coaches in case study settings, particularly the length of a session, the content that was delivered, and how it was delivered when AV equipment was unavailable. There was a difference in approach to training between academic coaches and consultant coaches. Support sessions also varied in duration and the level of feedback given, though practitioners were happy with the level of support received overall. The success criteria—that 'TWiTCH training and support is delivered consistently across different settings and trainers' and 'the TWiTCH programme is sufficiently described to enable consistent delivery'—were not met and further development of the programme and coach training is needed to allow more consistent delivery within a trial while acknowledging the need to adapt the training to different settings and levels of practitioner knowledge.

Acceptability of programme

RQ4 Is the TWiTCH programme acceptable to settings and the individual practitioners working within them?

a. Is the training at the appropriate level and accessible to practitioners in each setting?

Data used to answer this question has been synthesised from the post-training survey and interviews of practitioners from case study settings and interview settings.

In the post-training survey, most practitioners agreed or completely agreed that the training was beneficial (40 of 47, 85%) and provided them with enough information to get started with TWiTCH (41 of 47, 87%), although a few disagreed that enough information was provided (3 of 37, 8%) (see Figure 2); 4% (2 of 47) did not think the training was beneficial for their personal development because they already felt knowledgeable and confident about the TWiTCH strategies (as highlighted in Figure 9). Some practitioners stated that they felt the training was too long, especially due to many of the strategies being already familiar.

When asked how accessibility of the training could have been improved, some practitioners would have valued seeing more real-life examples of TWiTCH with children with special educational needs (SEN) or English as an additional language (EAL) or more diverse circumstances. A small number of practitioners mentioned concerns over the suitability of the books chosen, some aspects of which, they felt, were no longer appropriate to use with children (for example, the evil stepmother in Hansel and Gretl). It was mentioned that the TWiTCH programme was perceived by some as being too complex for the children in their nursery implying that children in their setting did not have the required language skills. Two practitioners expressed concern at being asked to read aloud from the handbook during the session, which made them very anxious and unable to take in some of the content.

Around two thirds of practitioners felt that their understanding across most areas had increased as a consequence of the TWiTCH training (see Figure 3) with more reporting increased understanding for the 'RCRI strategies' and the 'choice-based language activities'. Many practitioners already had good understanding about 'using storybooks to help develop children's vocabulary' and 'deepening children's engagement with story through continuous provision'.

Figure 2: Post-training survey responses to questions about the suitability and accessibility of the TWiTCH training

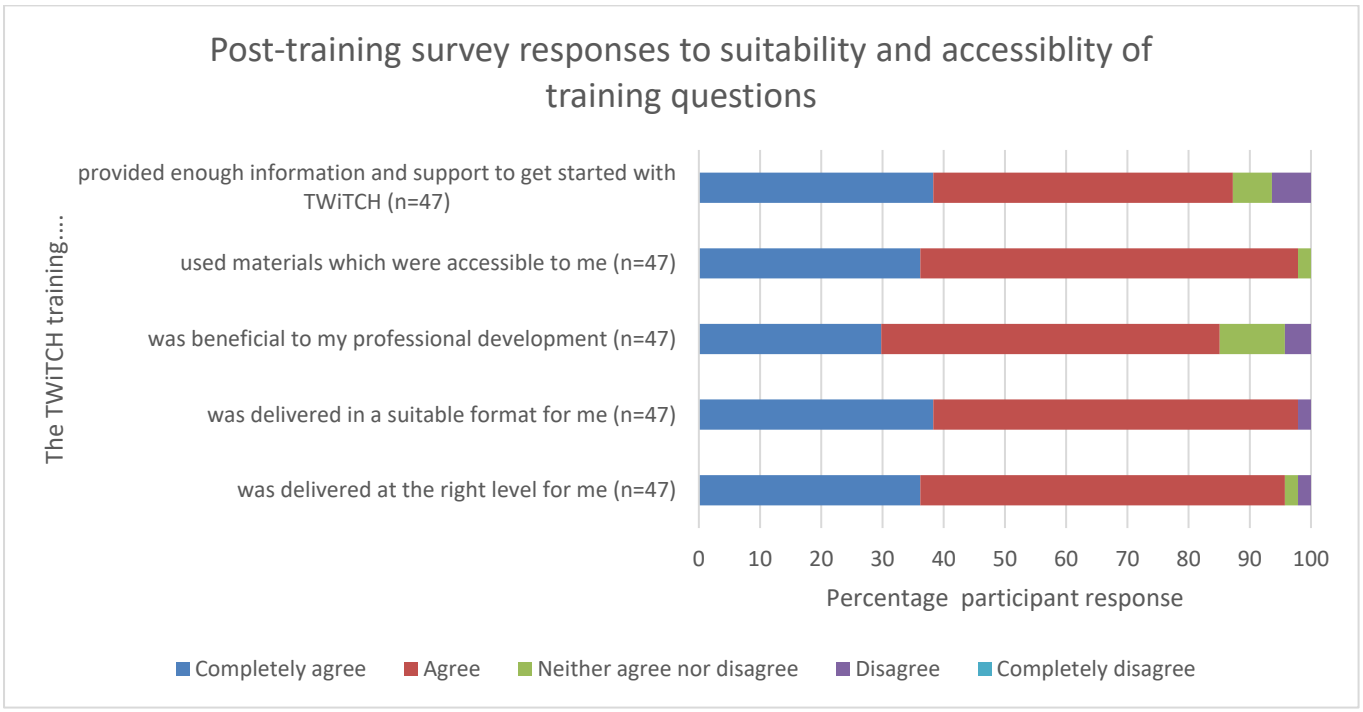
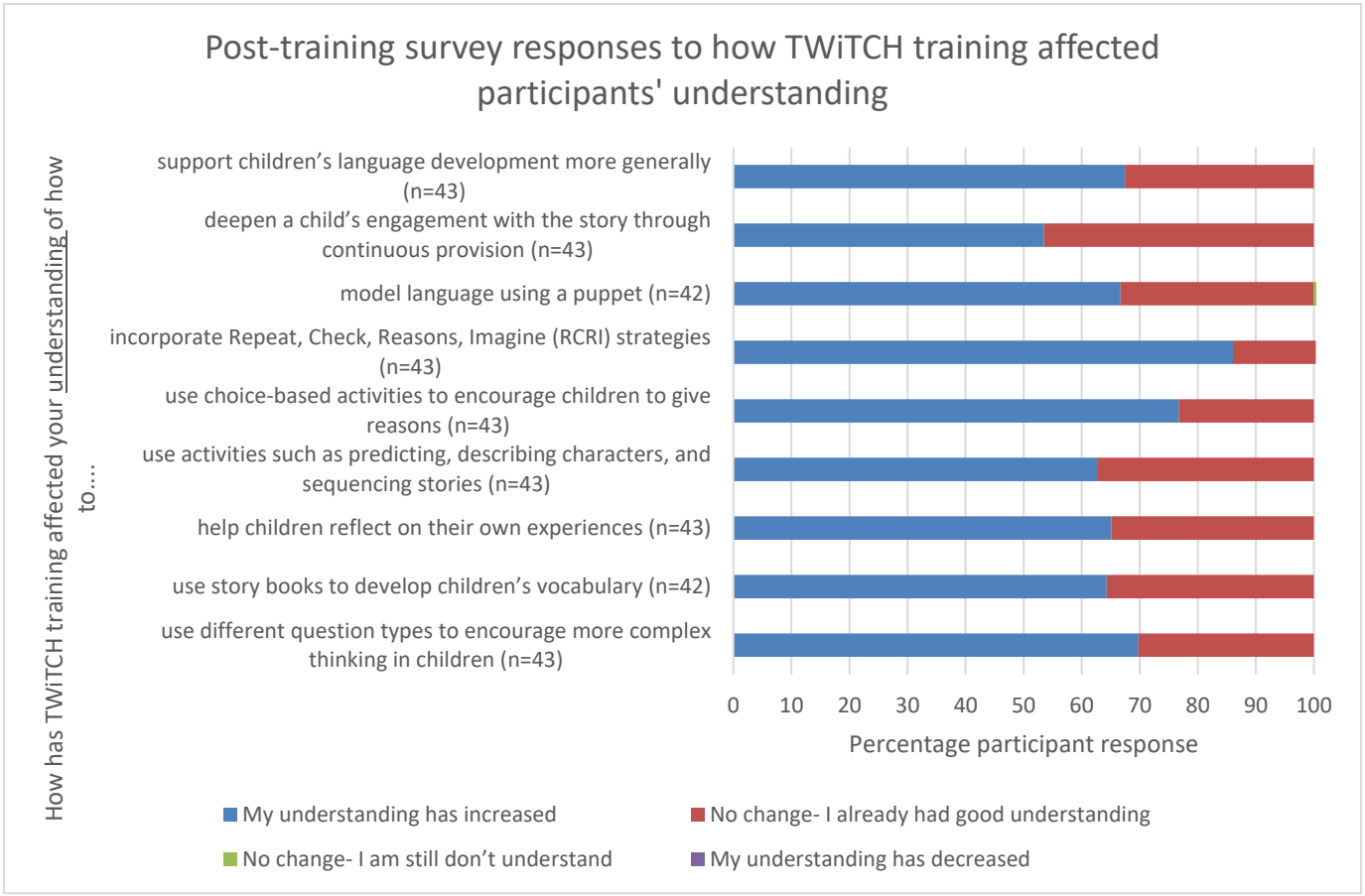


Figure 3: Post-training survey responses to how TWiTCH training affected understanding.



To summarise, the post-training survey showed that most practitioners found the training beneficial, with 85% agreeing it was helpful. However, a small number felt the information did not provide enough information to start delivering TWiTCH or that the programme was too complex for children in their setting; there were also a small number of critiques about the training's length, materials, and applicability to larger settings or younger or EAL children. Two-thirds of practitioners reported increased understanding of TWiTCH strategies and activities. The biggest increase was with RCRI strategies and choice-based activities to encourage children to give reasons. In terms of the success criteria, the training aspect of the TWiTCH programme was seen as accessible to the majority of practitioners.

RQ4b Is the level of support provided by the developer team sufficient and how are the different elements viewed (training, coaching visits, online support)?

Data used to answer this question has been synthesised from interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

Views of training

Practitioners in general were positive about the training and said they had a good understanding of TWiTCH after it. Individual suggestions for improvements were:

- no need for ice-breakers in a short session;
- no need to put people on the spot to act or read out in front of their colleagues; and
- ensure that the guidance for TWiTCH does not change after the training—that is, whether the children in the TWiTCH group should always be the same throughout the whole programme, or if anybody could take part at any given time.

This feedback was provided to the developer team during the project so that changes could be made for Phase 3 delivery where the team felt these were appropriate. The TWiTCH guidance had not stated that practitioners should read aloud during the training: this had been a decision made by the coach. There were more experienced practitioners who commented that they did not learn anything new or that one session would have been enough. On the other hand, these practitioners felt some of their colleagues, based on people's backgrounds, would have required more help, for example, personal guidance from their coach.

Views of coach support

In interview settings early in the programme, the majority of champions interviewed felt they had received useful advice from their coach. In case study settings, most staff valued coach support but this could vary between staff members even in the same setting. A coach's understanding of context and children in the setting was important in how the support was valued. Many settings in this study were in deprived areas and their intake had a high percentage of disadvantaged children with very low levels of language ability, often from families where English was not their first language. Where coaches took time to understand this context and make appropriate suggestions and adaptations to the programme they were viewed more favourably. However, in some cases practitioners felt that the coach did not understand the intake and children and so their support was viewed less favourably. Most practitioners interviewed felt that the level of support was right and they liked knowing they could contact their coach by email. It was felt that support could be improved by more practical modelling by coaches in a number of settings.

When interviewed again towards the end of the programme, champions were still largely satisfied with the level of support received from their allocated coaches. Eight of the nine champions from the interview settings, who commented on this point, were very positive about their coach. The champion that did not perceive the coach role to be important felt that peer learning would benefit their setting more than a designated coach and that it would have been more useful to get together with other practitioners delivering the programme. The champions that did value the coach supporting their setting felt that it had helped build confidence in staff to deliver TWiTCH (n = 2) and that it was important that the coach knew the setting and context (n = 2). Two champions particularly appreciated their coach demonstrating a TWiTCH activity in their setting with their children, with one of these going on to model what she had learnt to her colleagues. Two champions liked the opportunity to bounce ideas off a more experienced practitioner. For one champion, the coach was invaluable as they had just begun to deliver TWiTCH late in the programme, despite having attended the initial training, so they appreciated the opportunity of a refresher.

In one case study setting, a member of the SLT did not think an appointed coach was necessary for their setting, which was in stark contrast to a champion in the same setting who remarked:

> *'The support that we've had from the team has been amazing. We've been able to have the regular meetings. And then also, in between, if we've had any questions, we've been able to ask them and get support.'*

In this setting the champion was a recent graduate who was still building up their confidence in their practice. She particularly appreciated the observations and feedback of her practice and felt she had grown in confidence during the programme. The coach contact had been primarily with the TWiTCH champion in the setting, however, the coach had also met with the more experienced member of SLT staff and had observed her TWiTCH practice on one occasion. The member of SLT had felt that the coach had provided mainly positive feedback, which was not required, although did note also that the champion had grown in confidence during her time doing TWiTCH.

Views of online champion network meetings

Not all champions had chosen to attend the online network meetings and there were mixed feelings about this type of support. Where reasons were given for non-attendance, this was usually due to the champion not feeling like they had the time to attend or to not enough notice being given for them to be able to attend. Some champions thought it was good to hear experiences from other settings but others thought it was less useful, feeling that individualised support from the coach would have been more useful than sharing experiences with other practitioners.

One champion from a case study setting attended the online champion training and found the additional training helpful:

> *'It was a bit more in-depth than the practitioner training and just gave an insight as to what the champions were to do throughout the programme, and how to best support the other practitioners within the setting. So I found that very helpful.'*

This champion was given a follow-up one to one session with a member from the developer team as a second session in the afternoon.

Only two champions commented on the online support available in the second interview. One enjoyed sharing the experiences of the four settings she oversaw, particularly as they offered different perspectives of how TWiTCH was going. Another champion did not have a positive opinion of the online meetings but did not elaborate on this point and had missed some meetings due to not receiving a link. She was grateful for her coach who had helped with the information required for subsequent meetings.

To summarise, practitioners generally had positive views of the TWiTCH training, feeling well-prepared afterward. However, some suggested minor improvements to the way it was delivered. While experienced practitioners felt the training could have been shorter or more tailored, others needed more guidance. Regarding coach support, most practitioners valued it, especially when the coach understood the setting's context. Champions were mostly satisfied with their coach's support, with some appreciating demonstrations and opportunities for feedback. Opinions on online champion network meetings were mixed, with some finding them useful while others preferred more individualised support from coaches. In terms of the success criteria, the TWiTCH training and support was seen as acceptable and accessible to the majority of practitioners.

RQ4c How is the programme received by practitioners in different roles and different settings?

Data used to answer this question has been synthesised from the training and the post-intervention surveys, case study observations, and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

In the post-intervention survey, 24 of 34 practitioners (71%) said they would recommend TWiTCH to other early years settings; nine were not sure and one would not recommend it. This indicates that most found the TWiTCH programme a positive and worthwhile experience. This was similar across practitioners in PVI and maintained settings: in the former, 13 of 18 (72%) would recommend the programme while 5 of 18 (28%) were not sure. In maintained settings, of the 14 practitioners who responded, ten (71%) said they would recommend the programme, three (21%) were not sure, and one (7%) said they would not. Of those not sure or unlikely to recommend the programme, practitioners had a range of qualification backgrounds and were not consistently in one role in the project (champion, practitioner, or SLT). However, all but one of these practitioners had more than 13 years of experience working in the early years.

Most champions, when interviewed at the second timepoint, were very positive about the TWiTCH programme (n = 8). One with 20 years' experience commented that it was the first time she had spent three weeks on one book and was surprised at how much she was able to do with one book, such as focusing on language, maths, and retelling stories. Another practitioner was very positive but thought there were also challenges, particularly around how long the daily sessions took:

> *'So I do love it, I think it's fantastic, and it is a really good way for children to communicate, and build that language, but I do feel, for a preschool setting, I think it's a little bit lengthy, or it's a lot to do. I don't think it's as practical, in a preschool setting.'*

Another practitioner stressed the need for confidence to deliver TWiTCH to see the best outcome: 'And you've just got to be confident to go for it because when you see the impact you think, wow, yeah, this is really worth doing well.'

In one PVI setting, the uptake of the programme was 'hit and miss' according to the champion, which she attributed to a setting with a lot of experienced practitioners who did not feel they needed the additional training and were already confident to read stories. This champion oversaw a number of PVI settings and so was able to compare the uptake and outcome in these different settings. She noticed that the settings that were positive about the programme and with 'staff that really go for it' were the ones seeing an impact, whereas staff who said, 'Why do I have to do this?', were not seeing the impact.

One champion from a case study maintained setting had viewed TWiTCH as a language intervention programme for those with poor language skills and had done so from the beginning. When talking about the programme she commented:

> *'Those words are for children who are much further down the language journey, who are where they should be for being three and four. In which case, why would you be putting in an intervention? We just need a universal offer. If you talk about graduated response, which is what you're saying this is, this is an intervention at that next level, then it's not going to meet the needs for children with speech and language communication difficulties.'*

In summary, most champions were positive about the TWiTCH programme and would recommend it to other early years settings, noting its benefits, though some found the expected activities too long for preschool settings. Views of the programme did not seem to be particularly different for practitioners in different types of settings or with different qualification levels. Confidence in delivering the programme was mentioned as crucial for success. Uptake varied, with settings that embraced the programme seeing more impact. One champion viewed TWiTCH as more suitable for children further along in their language development.

RQ4d Is the intervention suitable for all children in the setting?

Data used to answer this question has been synthesised from interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

In first interviews with champions, most reported delivering TWiTCH to all children who were able to access it. Children not accessing it either had EAL, SEN, or speech and language (SAL) needs, or TWiTCH had been used as an intervention group specifically targeting children with low language ability rather than all the children of the appropriate age in the setting. Week 2 activities were considered too challenging at the beginning and some practitioners had reservations about their suitability for children. Additionally, the length of a TWITCH session was considered too long by some, especially at the start of the year. Champions mentioned that nurseries often chose not to offer TWiTCH to younger children, for example, two-year-olds in small settings with mixed age rooms or the SEN children. Also, some developmental delays in children resulted in them not being included in the TWiTCH group. On the other hand, EAL children or those who had been diagnosed with speech and language issues were given a priority to take part.

In the second interview, one champion commented that the level of TWiTCH may be more suitable for children entering the summer term since they found themselves heavily adapting the manual in autumn and spring because some of the questions were either too challenging or there were too many options for the children to choose from. By the summer, children had developed and were more able to access TWiTCH at a much deeper level. Another champion agreed that the level was pitched too high which resulted in simplification or making things age and stage appropriate.

In summary, the TWiTCH programme was considered suitable for most children in participating settings, with the exception of some children with EAL, SEN, SAL, or developmental delay. However, this was not the case in all settings and sometimes children with EAL or SAL were given priority. Due to the adaptations made by some practitioners, it was felt that the level of some activities would work better later in the year.

RQ4e How does the level of agency/decision-making given to the practitioner in the programme affect delivery and views of the programme?

Data used to answer this question has been synthesised from case study observations, coach logbooks, and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

Observations of case study settings and interviews with practitioners has illustrated two different approaches to the delivery of TWiTCH: those who follow the manual rigidly and do not adapt anything about the programme for their unique context and setting, and those who view the TWiTCH manual as a guide and have the confidence and experience (or are supported by a confident and experienced champion) to adapt the programme as they see fit.

In the first approach, these practitioners were not always aware that they had agency to make changes to TWiTCH, and they were the ones who were more likely to struggle to make the programme work or saw it in a more negative light. One practitioner commented on this in the post-training survey:

> *'I feel that it's quite prescriptive and gives little wiggle room for change. The majority of our children have English as a second language, and I don't feel it is aimed at those. I understand it's for practitioner learning but feel the children we are teaching has always been taken into consideration.'*

Another example is taken from an interview with a case study practitioner who felt that it needed to be clearer that it was possible to adapt TWiTCH:

> *'I think just re-looking at the [TWiTCH] textbook …There are a lot of different scary questions and lots of words. It makes it sound more complicated than it actually is and telling practitioners that it's okay to change some of the words if it doesn't suit … Some settings might look at that textbook and say, "Right. We have got to ask these questions", and they might not think. They might not pick up on that they're too advanced for the children they've got.'*

At the end of the interview, the practitioner was asked if she had any further comments and reiterated that permission to adapt could be more explicit: 'No. I just think just, from the start, making sure that practitioners are aware that they can adapt it, that it is okay to adapt it.'

In the second approach, settings that fully embraced the TWiTCH programme and extended its activities and question types into the full nursery context were seeing good outcomes. These settings also seemed to be the ones that had the confidence to fully own the programme for their unique context and children through adaptation and differentiation, rather than simply following it, as per the TWiTCH manual. For example, in the developer team second interview, they reported receiving feedback: 'TWiTCH has become embedded in settings beyond the life of the programme, a couple of settings have come back to us and said, "We're carrying on with it. We don't know in exactly what shape or form."'

On reflection, there was very little focus on decision-making or practitioner agency in the initial training, which might be why some settings struggled initially to deliver TWiTCH. Coach visits seem to have emphasised this flexibility to adapt through discussions with the practitioners directly. However, there could be a greater focus on practitioner agency from the outset, especially for those settings who had to wait a long time for their first coach visit. In many cases, coach visits were made in November and December after the training in the autumn but in some cases the first coach visits did not take place until January or February—or later for one case study setting. Early coach visits would also support any settings initial struggling to adapt the programme to their younger children.

In summary, observations and interviews revealed two approaches to delivering TWITCH: rigidly following the manual or adapting it to fit the unique context. Practitioners who strictly followed the manual often struggled, while those who adapted it saw better outcomes. Confidence and experience, or support from a champion, were key to successful adaptation. Settings that embraced and extended the programme into their full nursery context reported positive results. Initial training lacked emphasis on practitioner agency, which may have caused some struggles. Coach visits highlighted the importance of flexibility and adaptation, suggesting a need for greater focus on practitioner agency from the start.

RQ5 What costs and resources are required by settings implementing TWiTCH and is this feasible and acceptable to settings? What cost would settings be prepared to pay for the intervention?

Data used to answer this question has been synthesised from two semi-structured workshops with the developer team and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

Most TWiTCH champions (n = 9), when questioned in their second interview, did not report having to spend anything of significant value on delivering the TWiTCH programme. There were costs associated with buying resources for continuous provision, such as for baking (for example, ginger for gingerbread, or porridge for Goldilocks), but these were not considered to be above and beyond usual expenditure, 'just part of shopping'. One champion mentioned buying resources to make a TWiTCH display, including a cushion with a pouch for their rabbit puppet, and a floor book to evidence their reading with children, parents, and other professionals. Only one champion, out of the ten commenting on additional costs for TWiTCH, had to purchase extra books and puppets so that multiple practitioners could deliver TWiTCH simultaneously in their setting. Settings had been provided with two sets of materials where needed but had purchased additional books and puppets independently after it was not possible for the developer team to provide them with additional resources. Coaches had explored with the developer team whether it was possible for the setting to have further kits.

In one case study setting they had five groups running concurrently so had to photocopy and laminate four sets of materials and buy four sets of books. They also had to allocate one staff member to prepare the materials for the rest of the staff, which was general preparation time for the setting that was being completely overtaken by TWiTCH preparation.

When senior leaders in case study settings were asked about what they would be prepared to pay for TWiTCH, most felt that they would not be in a position to pay very much. Overall, the senior leaders valued the programme but had very little resource available for them to spend on CPD. One felt that if they were going to pay for the programme, they would like it to be free from errors and need fewer adaptations to make it suitable for delivery in their setting. Another would be prepared to pay up to £200 for the programme but did not feel they would need the coach visits. However, another senior leader felt that TWiTCH was invaluable to them and would be willing to pay for the training but could not put a price on it.

To summarise, in most settings, only minimal additional costs were incurred, such as items for baking, and this was deemed acceptable. A couple of settings had more considerable costs such as buying additional books, extensive photocopying, or allocating a practitioner to preparing materials for the whole group of practitioners. Although senior leaders in case study settings were asked directly, few commented on what they would be prepared to pay for TWiTCH, but felt it would be minimal due to very restricted budgets for CPD.

Programme fidelity

RQ6 Is the TWiTCH programme delivered with fidelity in settings?

a. To what extent do all practitioners attend the training? What factors affect whether practitioners can attend?

Data used to answer this question has been synthesised from the developers' administrative records (training attendance), the training survey, coach logbooks, and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

The number of practitioners attending training varied by setting: some settings sent all or most of their preschool staff to attend at least one session, however, the number ranged from one to eight; most sent three (mean, 3.54). The mean was similar across both maintained and PVI settings (3.58 and 3.50 respectively).

In two of the five case study settings, not all practitioners working with their three- to four-year-old children attended the training and for some attendance was only possible because the training was held in the nursery. This was a crucial factor that enabled more than one practitioner from a setting to attend. In one case study setting with ten members of staff, only three or four staff members were able to attend the original TWiTCH training sessions because not all staff could be released at the same time. The champion reflected that staff who were trained in TWiTCH were delivering it with more commitment and greater fidelity. Staff without training had observed and tried to replicate trained staff but did not have the same understanding or confidence to deliver the sessions as well.

In the first TWiTCH champion interviews, champions noted that not all practitioners would be able to travel from the setting for face to face training even if it was after 5 p.m.: issues such as not being able to drive and restrictive work patterns and other responsibilities could get in the way If the training was face to face, outside of normal working hours and demanded travelling, one champion mentioned restrictions on how many people could drive, and how work patterns and other responsibilities might prevent them from participating. A senior leader in a case study setting also noted that paying a large number of staff to work the additional hours needed for training was not feasible; another case study setting, however, had maximised staff attendance by making it part of a nursery inset day when children were not in attendance.

To summarise, not all practitioners attended the training, nor could they all attend. On average, 3.54 practitioners attended the training from a setting. Some of these practitioners were only able to attend because the training was held in the nursery. Reasons for not being able to attend training were that not all staff could be released at the same time when training was held during nursery hours, and that staff may not be able to attend outside of working hours. However, the success criteria that training sessions are attended by a minimum of three practitioners per setting for the majority of settings was met, 18 of 24 settings having achieved this.

RQ6b To what extent is the programme delivered in classrooms with fidelity? How does this vary between settings? How consistent is delivery across practitioners and settings? What are the reasons for variation in delivery?

Data used to answer this question has been synthesised from case study observations, interviews with the developer team, coach focus group, and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

To answer this research question, we summarise the extent to which the TWiTCH programme was delivered with fidelity in the five case study settings through the different elements of the programme and provide a comparative summary of the overall programme delivery in interview settings.

Case study settings

Organisation of TWiTCH

The case study settings organised TWiTCH in varied ways which depended on the size, timetable, space available, and nursery staffing. In four of the five settings, all the three- and four-year old children were able to access TWiTCH. Group sizes were also within the expected sizes in four out of five case study settings.

In one small setting there were nine preschool children in the nursery and both practitioners working with the group delivered the TWiTCH sessions together in a separate space, usually with a group of between seven and nine. However, due to one of these staff members working part-time, delivery of the sessions was across three days (later four days) rather than the five days expected in Week 1.

In a medium sized case study setting, TWiTCH was delivered by three practitioners to all children in consistent groups of seven to eight children which were run at different times during each day. This setting was newly opened and the TWiTCH programme had been embedded within the routines and plans of the setting from the start, facilitating delivery.

In the larger case study settings, delivering TWiTCH to all children was more difficult logistically. One ran four large TWiTCH groups (of around 13 children) simultaneously in the same large classroom space. It was unable to facilitate smaller groups due to staff ratios and, due to having a short nursery session time, it worked better not to have groups spread throughout the day. This set-up was challenging due to having to photocopy and purchase additional resources for four practitioners to deliver at the same time, as well as the noise from having four groups running concurrently.

In another of the larger settings, all children were able to opt in to TWiTCH sessions delivered by two TWiTCH practitioners. Due to the opt-in nature, the group size varied depending on the number of children who chose to take part on a given day. However, when more than ten children chose to take part staff would often repeat the session later in the day to keep the group size down. TWiTCH sessions would take part in a separate room when staff ratios allowed, which helped the children focus better on the sessions. One of the TWiTCH practitioners was unable to deliver TWiTCH five days a week for Week 1 activities and the activities were often compressed into three days in this setting.

In the final large case study setting, TWiTCH was offered as an intervention group to four groups of five to eight children (around a quarter of those registered) selected due to having lower language levels. Delivery was by three trained practitioners at times throughout the day. During the programme this setting changed the children who were in the TWiTCH groups as they found those with the very lowest language levels were not benefitting. During the year, this nursery started using the Week 1 activities across the nursery for their story-time sessions in key worker groups of ten to 13 children meaning that more children accessed these TWiTCH activities.

Fidelity to content

All of the case study settings reported following the TWiTCH handbook but made more adaptations to the activities of Week 2. Week 1 activities were followed closely, practitioners in all roles were positive about these activities, other than compressing the content from the five days of specified delivery into three days due to staff availability in two settings (described above); only minimal adaptations were made to the delivery of the Week 1 sessions. This included two case study settings using props instead of the pictures on the TWiTCH flash cards provided. Observation of Week 1 sessions across the case study settings also showed good adherence to the TWiTCH handbook using the specified strategies.

For the Week 2 content, three of the case study settings had adapted the content of the activities from the beginning of the programme due to feeling that the activities were too advanced for the level of the children. This adaptation had been planned in advanced by the TWiTCH champion and shared with the other practitioners. The adaptation involved simplifying the activities (for example, on an activity where children were presented with a choice of five items to pack for a trip and needed to give reasons for their choice, only two items were presented—and these one at a time—for children to respond to) and sometimes the language used was changed due to the concept likely being unfamiliar to some children (for example ‘ice skating’). In one of these settings, as practitioners got more confident with delivering TWiTCH, the champion stopped planning the adaptations and the practitioners were able to adapt the content as they delivered the sessions. The TWiTCH coaches had encouraged the settings to make these adaptations and felt that they were in line with the TWiTCH programme. Observation of Week 2 sessions in these settings showed good alignment to the programme and children engaging well with activities.

Two of the case study settings had tried to deliver the activities exactly as presented in the handbook, however, they found these sessions challenging to deliver and it difficult to get children to engage with the session. Practitioners in these settings did not always feel that the activities aligned well to the stories. In one of these settings the practitioners described skipping some of the Week 2 sessions early in the programme due to it being difficult to engage the children. However, feedback and modelling from the coach had supported them in delivering Week 2 sessions effectively and observation of a Week 2 session later in the project showed good alignment to the content and good engagement of the children. The other setting with a larger group size found it difficult to engage all the children in the activities within the time of the session and behaviour management of the group took up significant time. The coach in this setting encouraged the practitioner and champion to adapt the content to suit their groups and children but some of the practitioners lacked the confidence and expertise to do this successfully. Observation of Week 2 sessions and coach reflections for this setting indicated that the sessions were not being delivered as intended.

The rabbit puppet was being used as described across four of the five case study settings in Week 2. For the setting that did not use it the practitioners felt that it was confusing for the children. However, the language modelling that the rabbit puppet should have delivered was done instead by the two practitioners delivering the sessions together.

For the Week 3 content, all case study settings were delivering the continuous provision activities related to the story as described. However, in three of them, TWiTCH practitioners did not directly target ‘at need’ children with further TWiTCH activities. In one, this was because the Week 3 continuous provision activities were delivered by non-TWiTCH staff who had not been involved in delivering the Week 1 and Week 2 sessions. In the other two it was not felt necessary to target children with TWiTCH-specific activities as those who needed it were supported through normal practice.

In all case study observations of delivery there was good evidence of the use of the ‘repeat, check, reason, imagine’ (RCRI) strategy proposed by TWiTCH, even in those with weaker delivery.

Two case study settings had informed and encouraged parent involvement with the TWiTCH programme, going beyond what was specified in the TWiTCH handbook. This included parent noticeboards, sending information and activities home, and getting parents to record readings of the book in the child’s first language.

TWiTCH champion role in supporting other practitioners

There was evidence of the TWiTCH champion actively supporting other staff to develop their practice in using TWiTCH across three of the case study settings. This included the champion preparing TWiTCH sessions, materials and adaptations to share with other staff in the setting in two of these settings. In two settings the champion modelled TWiTCH for other practitioners, observed practice, and provided feedback to support staff with delivery.

In summary, in the five case study settings, fidelity to the organisation of TWiTCH with delivery to all children in small groups was met in three of the settings with the two other meeting either the group size or full cohort requirements. In terms of fidelity to the content of the TWiTCH cycles, all the settings described sticking closely to the content in Week 1 with minor adaptations whereas Week 2 content required simplification to be delivered effectively, which was done by three of the settings. The Week 3 continuous provision activities were delivered across all case study settings, but three of the five did not target specific children in Week 3 with additional TWiTCH activities. The TWiTCH RCRI techniques were being used by all settings in their practice. In three of the settings there was evidence that the TWiTCH champion was actively supporting other staff to develop their practice.

Interview settings

In interviews with champions, we asked about the general delivery of TWiTCH and how this was being applied in their contexts. In follow-up interviews in May and June 2024, we asked champions whether the delivery had changed since we last spoke to them, or we asked for further information about how TWiTCH was being delivered if this was not captured in the first interviews.

The group size in many nurseries was five to eight children. Some, however, had more than ten in their TWiTCH groups and some were experimenting with bigger groups, and one-on-one. Some nurseries managed to keep the group small by running separate morning and afternoon sessions, or parallel sessions if they had enough staff. Some did not have staff/child ratios to run smaller groups.

Week 1 seems to have been delivered with fidelity with the occasional combining of two sessions into one (that is, not doing TWiTCH each day of the week as per the manual). Even experienced practitioners were positively surprised by the depth of discussion they could draw from the text during TWiTCH. Practitioners reported that the Week 1 activities were not too dissimilar to usual practice, however, the texts were repeated more than they would usually do and the questioning was more structured than usual practice when reading a book.

Week 2 was more varied in delivery: many practitioners reported needing to adapt the Week 2 activities, some reported doing Week 2 activities less frequently than specified or skipping them altogether, and some did activities every day. Practitioners often felt the Week 2 activities were too advanced for the children in their settings, especially early in the programme. Most changes to the Week 2 activities were about simplifying the language and tasks to make it more suitable for the language ability and age of the children. Sometimes experiences referred to in the manual were unfamiliar to children, especially in most deprived settings.

Week 3 had a mixed delivery. All champions interviewed discussed using continuous provision. Some settings set out continuous provision in Week 1 to run through all three weeks. There was limited evidence of working with individuals or small groups of children who needed additional support (not all settings did this). All children who needed it were supported as part of normal practice rather than it being TWiTCH-specific. Where the story fitted into what the whole setting was doing, continuous provision activities were easier to implement and nurseries were more likely to have resources.

Use of resources differed per setting. The rabbit puppet was widely used but not in all settings. Many practitioners were very positive about using rabbit to engage children in the session (often using it to greet children at the beginning of the session), but there was some reluctance to use it as it was ‘distracting’ or not seen as appropriate (for example, the practitioner wanted children to see their mouth and facial expressions while talking). The carrot was used less frequently as some settings already had systems for turn-taking. The squeaker in the carrot was considered a distraction by some practitioners. Resource cards were often being used, but not all resources. Practitioners expressed a preference for real objects over pictures, and real photos over drawn images.

Discussion of fidelity in the developer interview

In two interviews with the developer team we explored the topic of fidelity, including discussions around the core components of TWiTCH. In the second interview we discussed in more detail what settings would need to be doing specifically to be following TWiTCH with high, medium, and low fidelity.

The developer team observed that high fidelity to the TWiTCH programme would be completion of the nine TWiTCH cycles using the supplied books, embedding the RCRI strategies, modelling language to the children, introducing and explaining new vocabulary, and implementing the full three-week cycle of activities for at least six of the TWiTCH cycles. It would also mean delivering the TWiTCH activities to groups of approximately eight to ten children and them being accessible to all preschool children in the setting.

For medium fidelity, they would like to see the completion of six cycles, the embedding of RCRI strategies, modelling language to the children, introducing and explaining new vocabulary. and implementing the Week 1, Week 2, and Week 3 activities for at least four cycles. Delivery of the activities may be to a larger group of children (more than ten) or only to a select group.

For low fidelity, they would see this as completing less than six TWiTCH cycles, attempting but not embedding RCRI strategies, and not completing activities across the three-week cycles (that is, excluding the activities of a certain week).

As the developer team was not able to discuss this until the end of the programme, a fidelity measure would require further development before it can be implemented as a reliable measure. The team found it difficult to comment on how consistently they felt that TWiTCH was being delivered due to not having quality assurance of delivery built into their delivery model. For a future trial, tracking of delivery across settings could be built into the delivery model through the coach support.

The developer team discussed the potential lack of consistency or fidelity of TWiTCH delivery in settings as not wholly a negative thing. The idea of agency to adapt TWiTCH could be considered a positive:

> *'So you could say lack of consistency or fidelity, or you could say it's responsive because of the diversity of the sector and the setting … potentially that responsiveness is a strength, not a weakness.'*

When looking at the variability of how settings target children who need additional support in Week 3, the second developer interview opened a discussion about inequalities and whether settings not targeting 'at need' children could be widening the attainment gap between disadvantaged and non-disadvantaged children. However, team countered by observing:

> *'I think … the flipside is the fact that we're seeing the repetition having the impact on those that are disadvantaged, means that maybe we didn't—maybe we were making assumptions about expectations of children that they actually are able to access it from day one because of the repetition of it. And therefore there wasn't as much targeted support needed in Week 3.'*

The team went on to acknowledge that we cannot be sure about this hypothesis as children's outcomes were not the focus of the evaluation.

To summarise, adaptation of the TWiTCH programme occurred in case study and interview setting contexts to suit the unique contexts and children's needs, with common challenges such as managing group sizes, ensuring consistent participation, and addressing language barriers. In case study settings, Week 1 was generally delivered with fidelity, while Weeks 2 and Week 3 had varied delivery, often involving simplified language and tasks. There was varied use of resources like the rabbit puppet and carrot and limited evidence of targeted support for individual children. Interview settings saw differences in delivery frequency (from daily sessions to three times a week), group sizes (from small, flexible groups to larger, structured groups), and practitioner involvement (multiple practitioners versus fewer staff). In regard to the success criteria that the 'programme is delivered in settings of different types with medium to high fidelity', it is seems that for the majority of settings this has been met, however, clearer records of programme delivery—and variability of delivery—across settings would help to provide more data here, particularly in respect of who the programme is delivered to: notably, the mixed delivery of the Week 2 and some of the Week 3 activities leaves some questions about whether 'high fidelity' was met.

RQ6c What role does the TWiTCH champion play in supporting the intervention delivery?

Data used to answer this question has been synthesised from case study observations, interviews with the developer team, the coach focus group, and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

In the first interviews, most TWiTCH champions reported that they were involved in supporting TWiTCH delivery in their setting. Some saw the role as a coordination role, with the coach and other staff, to support practitioners to do TWiTCH consistently. Champions reported often doing additional planning that was cascaded to other practitioners and TWiTCH was often discussed at staff meetings and informally to share experiences and trouble-shoot challenges. Overall, champions appreciated having the role, which gave them the opportunity to talk to other practitioners in the nursery.

Similarly, the role was appreciated in case study settings. The champion from one of these enjoyed taking on the role of TWiTCH champion for the opportunity given to her to share what she had learnt with other practitioners in the setting. This included talking to non-TWiTCH practitioners about how to use TWITCH activities in the nursery's daily activities, sitting in on story-reading sessions and giving advice, as well as giving and receiving feedback.

In the first developer team interview (December 2023) when all settings had received TWiTCH training and most were delivering their second or third TWiTCH cycle, the role of the champion was highlighted as being key to TWiTCH. The developer team felt that without someone to lead on the delivery, then TWiTCH would need to be a shared commitment and more easily forgotten, likely not having someone driving its implementation in the nursery setting. When questioned about who should take on the role of champion in a second interview, the developer team thought that someone with seniority would be the ideal candidate. This was to ensure that the champion would have the ability to make decisions and, therefore, make TWiTCH happen in a setting. Alternatively, if not a single person with seniority, an enthusiastic practitioner looking for a development opportunity could take on the role of champion, but they would need to have the buy-in from the senior leadership team so that there was the authority to make things happen. In one case study setting, the champion did not have the time to commit to the role as she would have liked due to other responsibilities in the nursery.

In the second interview, champions in interview settings were asked about any training they had delivered to new staff members in their setting. Three out of ten champions had trained other staff members, and an additional setting expected to do some training in September. One of the champions conducted the training so that a colleague could cover her while she was on leave, while the other settings trained new staff who had joined the nursery after the TWiTCH training had already been delivered. The champion's role was described as having a role in planning, often collaboratively with other practitioners, sharing ideas, and preparing topic boxes.

In the second developer interview (September 2024) after the settings had delivered the programme, the developer team indicated that a success of TWiTCH had been the formation of a community via champion networks that they hoped would continue on the website. Champions had been given links to the website to access materials and content.

In summary, the champion supported other TWiTCH practitioners by coordinating with other staff and the coach, and by preparing activities and materials for TWiTCH sessions. Some champions appreciated the opportunity to share good practice with colleagues. Training other staff members in TWiTCH had already been done by some champions.

RQ6(d) How is the programme adapted by those delivering it?

Data used to answer this question has been synthesised from case study observations, interviews with the developer team, the coach focus group, and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

In the first round of interviews, the TWiTCH champions reported following the manual but one said that they created their own shorter A4 version of instructions, which they copied and laminated for other practitioners to keep TWiTCH delivery consistent. Most changes to the programme involved simplifying the language or making the games simpler to suit the children. Mostly this was done at the point of delivery, rather than during planning, by using different wording when children did not seem to understand. For example, when talking about a 'cellar' most children did not have a cellar, so the practitioners talked about how there are different names to the same rooms in different homes. When sharing TWiTCH between multiple practitioners, some wrote the word definitions down to encourage consistency. Most of the

champions had adapted the TWiTCH cycles to either better suit their existing topics that follow the changing of seasons or ordered the books according to their complexity and length.

When interviewed a second time, champions in many settings described that they had continued to deliver TWiTCH as reported during their first interviews. In terms of changes made, one champion commented that they had had to heavily adapt the questions and tasks for children in the autumn and spring terms because the level was too challenging, or because there were too many options for children to consider. The champion felt that at the time of questioning—during the summer, after delivery of around six or seven TWiTCH cycles—the children were more able to access TWiTCH as described in the manual with fewer adaptations and that 'they are getting better at accessing it on a deeper level', engaging with less adult scaffolding with the reasoning prompts. Another champion also felt that the level was too high and so they had simplified the materials and adapted them to 'the age and stage' of the children.

Champions were specifically asked about their use of the TWiTCH materials (puppet, carrot, word and picture cards, spoons etc.) in the second interview. There was explicit mention of use of the rabbit puppet (n = 5), the turn-taking carrot (n = 5), and cards used directly used with children (n = 5). Two champions had decided to swap the cards for real objects, and two champions described using the cards as prompts for staff rather than for use with children. The spoons were being used (n = 2), but in one setting they were described as 'short-lived' and children quickly lost interest in them compared to other small word-play items.

The TWiTCH cycle process was followed to differing degrees. In one setting they had decided to deliver TWiTCH five days a week, even in Week 2 and Week 3 of the programme, as they believed the consistency was beneficial for the children. In one setting they sometimes did TWiTCH with the full group, especially if they only had one copy of the book they were reading and so were unable to divide the group. Another setting had reduced the days from five to three in Week 1 and were continuing this pattern in Week 2, with no Week 3 at all. Reducing Week 1 to three days a week instead of five enabled the setting to incorporate non-fiction books, which link in with the story, for the other days. The champion in this setting described that they had chosen to introduce Week 3 continuous provision activities from Week 1 and remove Week 3 from their TWiTCH cycle altogether:

> *'We're still doing it with the depth but three weeks for books for children that are so young—we find it's quite a long time and they start to get bored quite quickly. So, by bringing those weekly activities in straightaway to hook them in and get them interested we can eliminate that week, that final third week … there have been times when they've said, "Oh not again!", which is why we changed it and tweaked to the two weeks and not every day. I think everyday was a little bit much for them.'*

This was another motivation to introduce the non-fiction books, to add 'another dynamic and keep the children engaged.'

Champions in four settings explicitly mentioned running the same groups of children for TWiTCH as before, and the champion in one setting had introduced a new group of children to TWiTCH. One champion explained how before TWiTCH they had included 'vocabulary around text such as introducing the idea of the author, the illustrator, the blurb, the page, all those types of things,' which TWiTCH did not have, but they felt it was important to continue this established practice of theirs.

Week 3 continuous provision activities were put out from Week 1 in some settings (n = 2), with others sticking to the three-week cycle. In the second interview, champions were asked if they targeted specific children in Week 3. In response, some champions (n = 4) said they had done so but the selection process of who to select differed per setting. One champion described simply focusing on the children who had struggled with one of the TWiTCH activities in Week 2 and gave them further opportunities to explore the task by using different books in Week 3; others described focusing on children who lacked confidence (n = 2), and who they thought might benefit from one to one support, or who are 'low ability children' needing more language. In the latter setting, they repeated the storytelling activities from Week 1 in the third week and focused on repeating language or questioning. One champion described targeting two children with EAL, but they did not engage with TWiTCH. Other champions in interview settings (n = 4) did not target specific children in Week 3 and either continued to work with their usual group/s (n = 2) or did this as part of their general practice: 'We're always picking up our bottom 20%, it's like our bread and butter.' The champion who reported this was in the setting which had condensed the TWiTCH cycle into two weeks. In one of the settings where they were not targeting specific children, the champion described that they had originally done so at the start of TWiTCH but now felt that the child they had worked with, who had communication needs, was now able to access the programme with the rest of the group.

To summarise, champions generally followed the manual but made on-the-fly adaptations during delivery of sessions like simplifying language and games as well as adaptations to simplify language and tasks made in planning sessions before delivery in some settings. By the second round of interviews, many settings continued with these adaptations, finding the original materials too challenging. Use of materials varied, with a preference for real objects over cards. The delivery of the TWiTCH cycle varied: some settings maintained a five-day schedule while others reduced it to three days or eliminated Week 3 altogether, integrating continuous provision activities from the start. Week 3 activities were sometimes targeted at specific children needing extra support, but this varied by setting. In some settings they focused on children who struggled with Week 2 activities or lacked confidence, while others did not target specific children, integrating support into general practice.

RQ7 What challenges/barriers are faced by settings in delivering the intervention?

Data used to answer this question has been synthesised from case study observations, coach logbooks, the coach focus group, and interviews of practitioners from case study settings (n = 5) and interview settings (n = 12).

In the first round of interviews the TWiTCH champions commented that one of the biggest challenges for the settings had been part-time children (reported by three private and one school nursery). For the study, in some settings they were only targeting TWiTCH at children who were consistently attending otherwise they would have had to adapt the programme for children who only attend part time by condensing the programme into fewer sessions. Another challenge was staffing (n = 4): how to always maintain a sufficient child-adult ratio for all groups of children. A third challenge was finding time and a quiet space for the session (n = 4). The staffing issues and lack of quiet space meant that TWiTCH sessions had to be very carefully planned. In one nursery, where three practitioners were each doing TWiTCH, they took two thirds of the children (with the two-year-olds who did not take part in TWiTCH) with two adults out, while one practitioner with eight children stayed in to do TWiTCH. They swapped after 15 to 20 minutes and again after 30 to 40 minutes to be able to run TWiTCH for three small groups in a quiet space when the rest of the nursery were playing outdoors. The fourth challenge most mentioned by champions (n = 4), was the duration of TWiTCH. In one nursery, children were bored to be doing the same story for three weeks. In other settings, the stories took longer than the advised ten to 15 minutes mentioned in the manual, and children's attention span and willingness to sit down could not be sustained for 20 to 25 minute sessions.

Other challenges were mentioned only once or twice by champions. One said they could not find any use for the flash cards. Another said the props did not always match the story. For example, they had been given dog, cat, and pig masks, but the story had a duck and a rat in it. Another comment was that some of the language or morals of the story were challenging in the modern world: the Princess and Frog book involved teaching children to keep secrets and Jack and the Beanstalk used, 'I smell the Englishman and I'd like to grind his bones to make his bread.' The moral dilemmas of these stories are a key feature of the intervention and especially in the Week 2 activities where choice based language games are used to discuss reasoning for and against decisions related to the stories and some level of dilemma or challenge is required for these discussions, however, some of the issues in the stories were not seen as suitable by some practitioners who were having to lead discussions about uncomfortable topics. Finally, it was thought that the programme needed tweaking for EAL, SEN and two-year-old children in some settings to simplify the language and tasks or provide adaptations that could be tried for these different groups.

In the second interview, champions (n = 3) again spoke about the difficulty of accommodating part-time children in the TWiTCH programme. In one setting the champion noted that the programme had only worked with some of the morning and afternoon children, but not the whole class due to attendance patterns. When interviewed for the second time, a champion from a case study setting explained how they had children who only attended one or two days that did not fall on a TWiTCH day and which may fall on a different day the next week. This made it challenging to ensure that these children had had enough exposure to the book to be able to do the TWiTCH activities. The champion did observe that parental involvement in the programme gave them some reassurance that they may have further opportunities to explore the book at home. The developer team also acknowledged part-time attendance as a potential challenge for settings when questioned in a second interview. They expanded on this point to differentiate between school and PVI settings, noting that the former would find it easier to implement TWiTCH due to more consistent attendance patterns in a school setting.

Other main challenges for settings that were mentioned in second interviews by champions were staffing and space. In terms of staffing, staff leaving (n = 1), staff absences (n = 1), and having enough staff to split the children into smaller groups (n = 2) were noted as challenges. Being able to access a suitable space for the small group work was also a

challenge (n = 2), as was generally fitting TWiTCH into the day and routine (n = 2). In one setting, splitting up SEND children and managing the sleepers was causing challenge.

Week 2 of the TWiTCH programme was mentioned by four champions as a continued challenge. One champion attributed this to staff needing time to understand and process what to do for this part of TWiTCH. Another champion thought that the skills required of practitioners was more advanced for this stage of the TWiTCH cycle. This champion believed that when elements of delivery were scripted, practitioners had a better chance of succeeding, but when practitioners had 'to go off on a tangent to go with the children's interests', this demanded a higher level of confidence, patience, and skill to pitch the language and questioning at the right level, which the less-experienced practitioners found challenging. She commented that all practitioners were still having a go and showing improvement. One champion commented that the level of the activities was too high for the children and so adapting the content of Week 2 has been a challenge for the setting and has taken time for practitioners to 'get their head around'.

The developer team, when questioned in the first interview, acknowledged that there may need to be greater focus on practitioner autonomy and judgement. They reflected that practitioners find it challenging to stick rigidly to the TWiTCH manual and perhaps do not use their autonomy to adapt the programme for their setting: 'I think some of the stresses that some of the practitioners have had with TWiTCH is that they approached it like … It says that in there, I've got to do this.' They discussed that for future iterations of TWiTCH there may need to be 'the permission to actually experiment and use your judgment of the children' while acknowledging this may not be easy for all practitioners due to confidence and their expectations from previous training: 'It's hard to unlearn that, though, for a lot of practitioners, that model of intervention.' Thus, practitioners who view TWiTCH as an intervention or compare it to other training they have received may hesitate to adapt the materials and activities for their children because they believe its efficacy is tied up in following the manual as closely as possible. The developer team discussed an assumption they had about practice and how permission to adapt might need to be more explicit:

> *'Yeah. It took me back about the thing you said about the word cards and things, so when they said, "Oh, our children it's really hard for them," so I think in my head my assumption was that actually they'll look at that and think, "Oh, actually let's see if we can make up some cards that's going to be right for our kids." But because it's there in the pack they assume that I've got to use this even if the children don't have hardly any language to start.'*

The developer team's assessment—from discussions with others they encountered in the field through their training delivery and other sector events—was that the sector generally seemed to be experiencing a lack of confidence. According to the developer team, the sector has lost a lot of experienced people and new people coming into the sector have low levels of confidence, particularly in how to plan activities and in relation to knowledge of child development to support planning. TWiTCH offers these practitioners the structure that people are asking for because 'they've got something to hang on to'. Similarly, as inexperienced staff are coming into the sector, this puts pressure on the experienced staff to care for the children and meet their needs, leaving little to no time to coach and support those newcomers.

In summary, part-time children, staff ratios to facilitate small group work, access to quiet space for small group work, the length of TWiTCH sessions when running over ten or 15 minutes, and Week 2 activities for both children and staff were the main challenges experienced by practitioners, particularly in relation to being able to adapt the materials so that they were suitable for their children. The issue of being able to facilitate small group work across the full cohort of children was an issue for a many settings as was the level of the Week 2 activities for children early in the programme. Due to these issues we feel that the success criteria 'absence of any major barriers to delivery for the majority of settings' was not fully met.

RQ8 What factors support the delivery of the programme?

Data used to answer this question has been synthesised from case study observations, interviews with the developer team, coach logbooks, the coach focus group, and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

Many practitioners valued the idea of small group work and may have tried this in the past with other activities:

> *'So, working in small groups, the huge one … obviously we've always tried to do activities in small groups, but this has forced us to make it a priority again. And it's really, really beneficial for those quiet children who*

don't have the confidence to speak in a larger group or can't make themselves heard over certain other children.'

But not all settings could facilitate small group work due to space and staffing. The one setting which withdrew from the programme explained that one of the reasons was due to space issues.

Practitioners who had the confidence and skillset to adapt TWiTCH materials and activities for their own unique context and group of children were the ones who embraced TWiTCH and had a positive view of the programme. However, not all practitioners seemed to understand that they had the agency to adapt the materials as this was not focused on during the training and coach visits were often not until after settings had implemented a few cycles of TWiTCH.

Nursery type and context might be a factor that affects the delivery of the programme. In discussion with the developer team in the second interview, the team highlighted the differences between school and PVI settings—that the former are able to retain staff, offer staff training, and offer better pay: 'We tend to lose people from PVI into the school sector because of the pay conditions … They tend to recruit the most talented.' Shorter working days in school settings also means that CPD can be offered outside of working hours, which is harder to accommodate in PVI settings. The issue of inconsistent attendance of children in PVI settings was flagged again. On reflection, the developer team was very impressed with how PVI settings embraced the TWiTCH programme, despite all these additional challenges facing their particular context. For PVI settings, higher levels of attendance is neither expected nor enforceable according to the developer team, 'The rhetoric is all about, this is just about getting parents into work, so you need maximum flexibility to fit in with the working patterns of parents, rather than the outcomes for children.'

Planning and sharing TWiTCH collaboratively with staff seemed to benefit consistent delivery and gave practitioners confidence to deliver. A case study setting noted that frequent meetings to share resources and ideas across its four sites had benefited champions and fostered a collaborative—but also competitive—element in practitioners, helping to make TWiTCH as good and as exciting as can be for the children.

Embedding the programme beyond storytelling was a feature observed in settings with the most successful delivery, especially when, for example, TWiTCH was being embraced by non-TWiTCH practitioners in their own practice, key vocabulary and question types were being used by all staff even outside of storytelling, parents were taking TWiTCH questions and texts into the home learning environment, and children themselves were using reasoning language and questioning in free play. In the second developer team interview, the team also felt that TWiTCH was particularly successful when it was viewed and embedded beyond a standalone storytelling session and involved 'integration into their planned practice, so it stops just being an add-on, that they start to see it as part of their planning, part of their curriculum'—which then also fosters an idea of shared delivery and community in the setting. The developer team had initially hoped that the champion network meetings would fulfil the anticipated need for a community of practice outside of the setting, but seeing it emerge within some settings was considered even better for the sustainability of TWiTCH. Embedding TWiTCH in this way is how practitioners can change their behaviour and interactions with children beyond storytelling. According to the developer team:

'It's actually hard to know whether—when we're looking at the impact of TWiTCH—is it TWiTCH itself or is TWiTCH almost like a catalyst to make other things that are happening even more effective?... It's not really the magic bullet intervention in that way. It's about behaviour change of the adults, isn't it?'

In summary, a number of factors have emerged as supporting the delivery of the TWiTCH programme including a setting having the physical space and staffing to support small group work, having practitioners who have the confidence to adapt the programme to suit their children, the unique nursery context, planning and sharing TWiTCH collaboratively with staff, and embedding TWITCH beyond story-time.

Evidence of promise

RQ9 How different is the TWITCH programme to (a) other CPD programmes the settings have participated in, (b) existing knowledge, and (c) usual setting practice?

Data used to answer this question has been synthesised from baseline and post-intervention surveys, the developers' administrative records (for example, the MoU sign-up form), and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

In a series of questions about normal nursery practice in the baseline survey, the average storytelling duration in the full sample of respondents (n = 31) was 13.83 minutes (SD: 4.78; range: 5–25 minutes).[8] This was shorter in PVI settings (n = 17 from 12 settings; M: 12.79; SD: 3.74; range: 5–20) but longer in maintained settings (n = 14 from 11 settings; M: 14.82; SD: 5.76; range: 7.5–25 minutes), compared to the overall average. These storytelling durations are not largely dissimilar to the expected duration of TWiTCH sessions.

Most practitioners in both PVI and maintained settings reported having story-time in their setting once or more than once a day (87.5%, see Table 10), which is comparable practice to the TWiTCH programme.

Table 10: Frequency of storytelling in the three- to four-year-old group, from baseline survey responses (n = 32)

	Total number of responses (%)	Practitioners in PVI settings (%)	Practitioners in Maintained settings (%)
1–2 times a week	1 (3.1 %)	1 (5.6%)	0 (0%)
3–4 times a week	3 (9.4%)	1 (5.6%)	2 (14.3%)
Once a day	13 (40.6 %)	6 (33.3%)	7 (50%)
More than once a day	15 (46.9%)	10 (55.6%)	5 (35.7%)

Half of practitioners reported personally delivering the storytelling session less than every day (see Table 11). This finding is different to the TWiTCH programme for Week 1, which should involve a practitioner delivering story-time every day. However, an additional 1.3% of practitioners reported doing story-time three to four times a week, which compares to the frequency of storytelling activity in Week 2 and Week 3.

Table 11. Frequency of practitioner doing story-time, from baseline survey responses (n = 32)

	Total number of responses (%)	Practitioners in PVI settings (%)	Practitioners in Maintained settings (%)
1–2 times a week	6 (18.8%)	4 (22.2%)	2 (14.3%)
3–4 times a week	10 (31.3%)	4 (22.2%)	6 (42.9%)
Once a day	16 (50 %)	10 (55.6%)	6 (42.9%)

Most practitioners reported that their group size for story-time was more than ten children (62.5%, Table 12), and just over a quarter (28.1%) had story-time with 7 to 9 children, which is closer to the specified group size for the programme. Those in maintained settings were more likely to do storytelling with a group of more than ten, with only 14.3% doing storytelling with smaller groups, compared to 44.5% in PVI settings doing storytelling with smaller groups. Where practitioners reported group sizes of more than ten, we do not know whether this was whole-class practice or still split into smaller groups: this was a limitation in the design of the survey responses. Anecdotal evidence from nurseries indicated that story-time was often done as a full group activity with one practitioner while others set up or cleaned up other activities.

Table 12. Size of group for storytelling, from baseline survey responses (n = 32)

	Total number of responses (%)	Practitioners in PVI settings	Practitioners in Maintained settings
4–6 children	1 (3.1%)	1 (5.6%)	0 (0%)
7–9 children	9 (28.1%)	7 (38.9%)	2 (14.3%)

[8] One participant who completed the survey was missing data for this question.

More than 10 children	20 (62.5%)	8 (44.4%)	12 (85.7%)
Depends on the day	2 (6.3%)	2 (11.1%)	0 (0%)

When practitioners were asked in the post-training survey about how TWiTCH training had affected their understanding of children's language development and dialogic storytelling practice, around a third felt that they already had good understanding of most of these practices before TWiTCH. Those in maintained settings were more likely to report this than those in PVI settings (36% compared to 28% respectively). Almost half of practitioners felt they already had good understanding of using continuous provision to deepen a child's engagement with the story, again, slightly higher for those in maintained settings—48% compared to 44%. Practitioners were less familiar with the RCRI strategies, choice-based activities, and using different question types and similar levels of knowledge were reported on these across both types of setting although maintained settings' knowledge was still a little higher. The majority of practitioners completing the post-training survey reported increased understanding on all elements of TWiTCH due to the training provided (see Figure 3).

According to the developer team, when questioned in the first interview about early years language programmes that nurseries would be accessing, WellComm training and Elklan training had been heavily promoted in the last five or six years of going through things like the Early Outcomes Projects. Other alternatives included Helicopter Stories, developed by a speech and language therapist in Sheffield and very popular, but it was not an ongoing programme of support: 'You go in and somebody's doing it but then the book's put away and then it doesn't carry on in the same way.' And then there were more local programmes like Stories for Talking. The developer team thought there were a number of options available to settings, with NELI being a key one with current funding from the DfE and widely available. However, according to the developer team, TWiTCH differs to these programmes in its approach: 'What we're trying to do is improve the quality of adult/child interactions.'

In the second developer team interview, the TWiTCH programme's distinctness compared to other programmes was explored:

> *'So actually there's quite a lot of research about the value of repeated reading. So that's, kind of, reinforced our thought, that's actually what the "TWiTCHness" of that TWiTCH: it has to have that repeated reading for it to be a TWiTCH session.'*

Repetition was considered beneficial for the children, particularly those needing extra support:

> *'Particularly around children with SEND or with English as an additional language, having the concrete book alongside that repetition, so they've got a hook to hang, you know, the language onto, really helps it to become embedded, and them to feel levels of confidence.'*

And also practitioners:

> *'But it's also for the practitioners, the repetition; it's actually giving them more time to—more space to start to do more about the story, to think more about the story, to feel confident themselves about reading a story. Because the more they're reading, one or two of them have said to us that actually it's improved practitioner confidence in reading stories.'*

When questioned about practitioner experience of story reading prior to TWiTCH in the second interview, the developer team thought this was an area new practitioners lacked confidence in: 'It's often one of the things they find the scariest … Because there's the performance element to it, and the visibility, and other people hearing them do it.' Through their experience of training apprentices and new practitioners, the developer team have found that storytelling is something practitioners 'shy away from' due to a lack of confidence and that TWiTCH is 'building a skill that should be central to early years'.

In interviews with the champion from a case study setting, the champion, who had many years' experience of trying different language and early years programmes and interventions in her setting, discussed how she had initially perceived TWiTCH to be a language intervention programme to help children with low language ability, and consequently they had initially stopped doing another language intervention programme, TalkBoost, to be able to commit to TWiTCH. After doing TWiTCH and not finding it to support children with the lowest language levels as well as

TalkBoost had previously, the lowest language learners were removed from the TWiTCH programme group so they could restart TalkBoost. TWiTCH was continued with a low to mid language level group. The champion reflected that they felt: '[TWiTCH] needs marketing more as a storytelling tool and developing storytelling skills, rather than as something for language development, really.' However, the champion felt that the TWiTCH storytelling methods were very valuable and had adapted storytelling practice throughout the setting to use the TWiTCH techniques.

To summarise, TWiTCH does differ to usual setting practice, particularly in relation to story-time group sizes and the frequency with which any given practitioner might tell the story. Similarities between TWiTCH and usual setting practice include the duration of a storytelling session and the frequency that the three- to four-year-old group have story-time. Some practitioners had existing knowledge of language and storytelling practices used in the TWITCH programme, but many were less familiar with RCRI, choice-based activities and using different question types. The developer team felt that the repeated reading, promotion of quality adult-child interactions, and building the skill of storytelling is what differentiates TWiTCH from other programmes. In a case study setting, TWiTCH had initially replaced the usual practice intervention used for children with low language levels, however, it was felt by the champion and senior leader in this setting that this was inappropriate after trying TWiTCH, and that TWiTCH, instead, added value to their usual story time practice across all children in the nursery.

RQ10 Does practitioners' practice change (and if so, how) as a result of taking part in the TWiTCH programme (measured through self-report and audio observations)?

Data used to answer this question has been synthesised from baseline and post-intervention surveys, structured and coded pre and post observations of practitioner interaction practice during story-time, and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

According to self-reports, practitioners' practice does change after following the TWiTCH programme. Matched data from baseline and post-intervention survey results show an increase in the self-reported frequency of the use of language support strategies and of storytelling practices. Paired samples t-tests (see Table 13) indicated that there were significant differences in responses to these scales describing practice from the baseline to the post-intervention with higher scores post-intervention on both scales (analysis was run in SPSS and syntax is included in Appendix D.1). This indicates that practitioners increased their use of language supportive practices both in their practice generally and while doing story-time during the project.

Table 13. Descriptives and paired t-test results for frequency of use of language support strategies and storytelling practices scales from pre-post survey (matched group)

	Baseline mean (SD)	Post-intervention mean (SD)	95% CI mean difference	df	T	Effect size estimate Cohen's d (95% CI)
Use of language support strategies (n = 31)	22.45 (3.47)	25.45 (2.90)	1.6–4.41	30	4.36***	0.78 (0.38; 1.18)
Use of storytelling practices (n = 32)	24.56 (6.61)	29.19 (5.55)	2.13–7.12	31	3.78***	0.67 (0.28; 1.05)

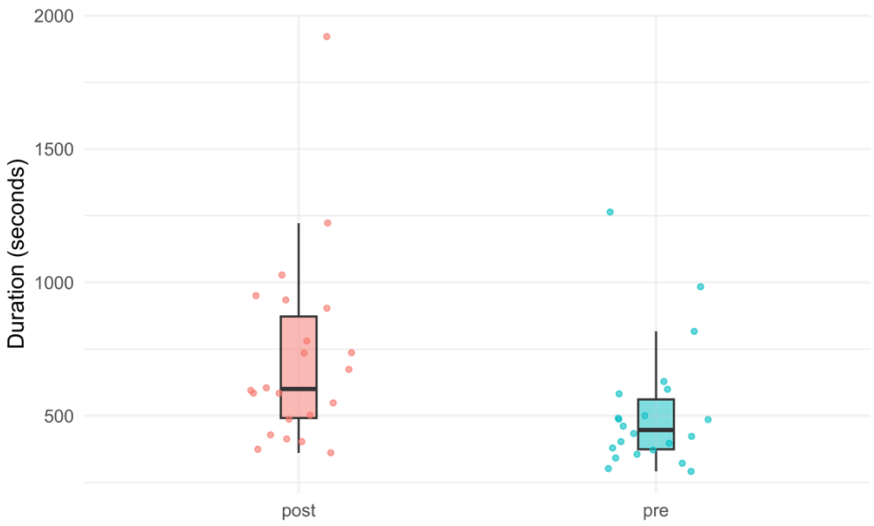
*** $p < 0.001$.

The findings from the survey are supported by interview data. Practitioners reported that even outside of TWiTCH, they now give children more time to respond and they use better questions to engage children in the book, especially more open-ended questions. Practitioners reported trying to encourage a range of answers and that they repeat what a child has said to validate their thoughts to a greater degree, following TWiTCH. Practitioners have also noticed the impact of TWiTCH beyond the sessions and in the wider nursery. For example, some non-TWiTCH practitioners have observed their TWiTCH colleagues delivering TWiTCH and have incorporated some of the techniques into their own practice. In some settings, Week 1 methods have been incorporated throughout the nursery and built into plans during the year. Another wider impact has been the use of TWiTCH to engage parents through sharing stories and activities, involving parents in translation, and sending out newsletters about TWiTCH books. The success of TWiTCH is evidenced by the fact that many nurseries plan to use it again next year in some way—either through running the full programme again and repeating all the same texts, choosing some of the books to incorporate into planning, or using Week 1 activities and incorporating some reasoning questions across practice. In one setting being inspected by Ofsted, a TWiTCH session was delivered and this was positively commented on.

Audio observations

In addition to self-report measures, audio observations of practice indicate a change in storytelling practice as a result of taking part in the programme. According to analysis of the audio recordings of storytelling at baseline (n = 22) and post-intervention (n = 22), the total duration of audio recordings increased from baseline (M: 515s; SD: 234s) to post-intervention (M: 717s; SD: 355s). The effect size, as measured by Cohen's d, was d = 0.67 [0.06, 1.28], indicating a medium effect. This indicates that practitioners, on average, were spending longer on reading and engaging children during storytelling (Figure 4) after the implementation of the TWiTCH programme.

Figure 4. Duration of audio recordings of storytelling in baseline and post-intervention recordings

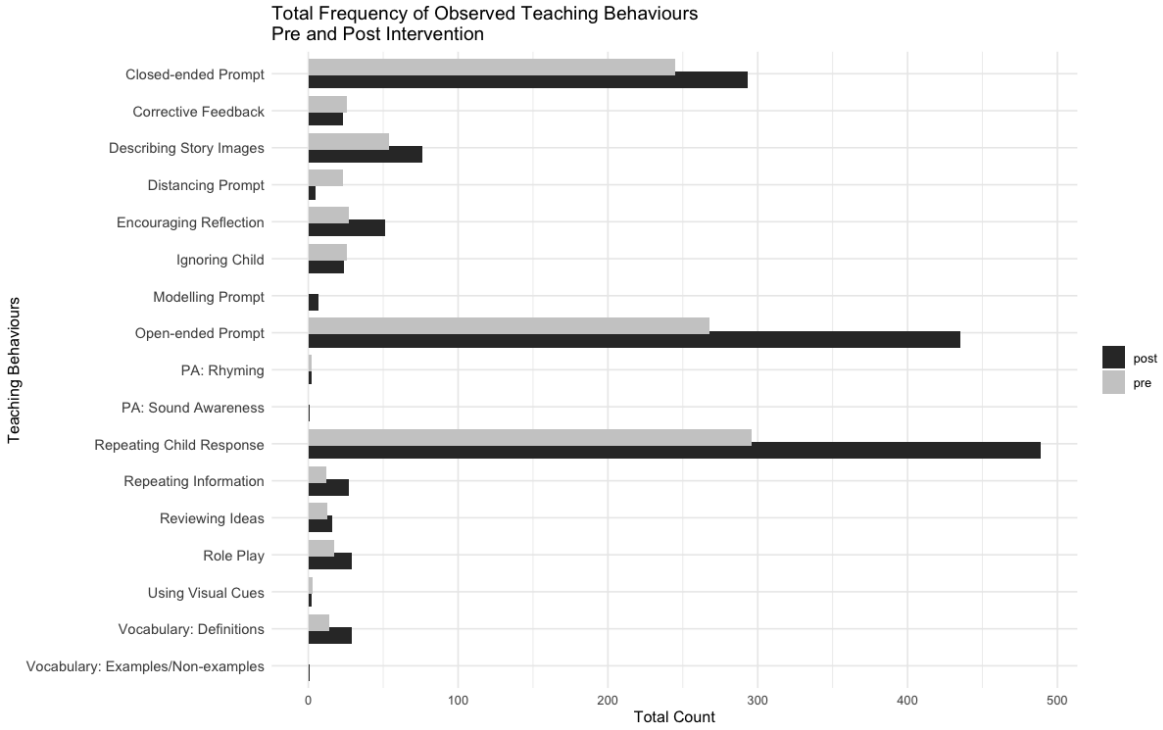


Practitioner talking time during storytelling was higher in post-intervention recordings (M: 334s; SD: 193s) compared to baseline (M: 314s; SD: 157s), with an effect size of d = 0.11 [-0.48, 0.7], which indicates a very small effect. However, to account for variations in session length and to provide a more nuanced interpretation, we examined the proportion of time that practitioners were speaking during the storytelling episode. This proportion *decreased* from baseline to post-intervention from 82.36% (SD = 9.79), on average, to 71.21% (SD = 18.87). Thus, although the absolute time increased slightly, practitioners were proportionally talking *less* during storytelling after the intervention, suggesting a shift toward greater child verbal engagement.

In recordings, practitioner talking time around storytelling (before or after the story was read) was higher post-intervention (M: 145s; SD: 92s) compared to baseline (M: 80s; SD: 58s), with an effect size of d = 0.84 [0.22, 1.45], which is considered a large effect size. However, the percentage of time practitioners spoke during these pre- and post-story segments decreased slightly from 62.79% (SD: 20.76) at pretest to 57.90% (SD: 8.37) post-intervention. This might indicate that practitioners were spending more time engaging children in discussion before and after storytelling after following the TWiTCH programme compared to their normal practice before.

In all 44 recordings—22 pre and 22 post—the number of occurrences of different question types and dialogic reading practices was averaged. The most noticeable changes between baseline and post-intervention recordings include an increase in the average frequency of open-ended prompts (pre: 12.18; post: 19.77), closed or completion prompts (pre: 11.14; post: 13.32), and repetition of what the child said (pre: 13.45; post: 22.23). There were also increases in other TWiTCH-specific pedagogical practices such as role playing, encouraging reflection, and describing story images (see Figure 5). There were few or no occurrences of practices relating to print concepts and phonological awareness in baseline or post-intervention recordings, but these are not features of the TWiTCH programme and so this is not a surprising finding.

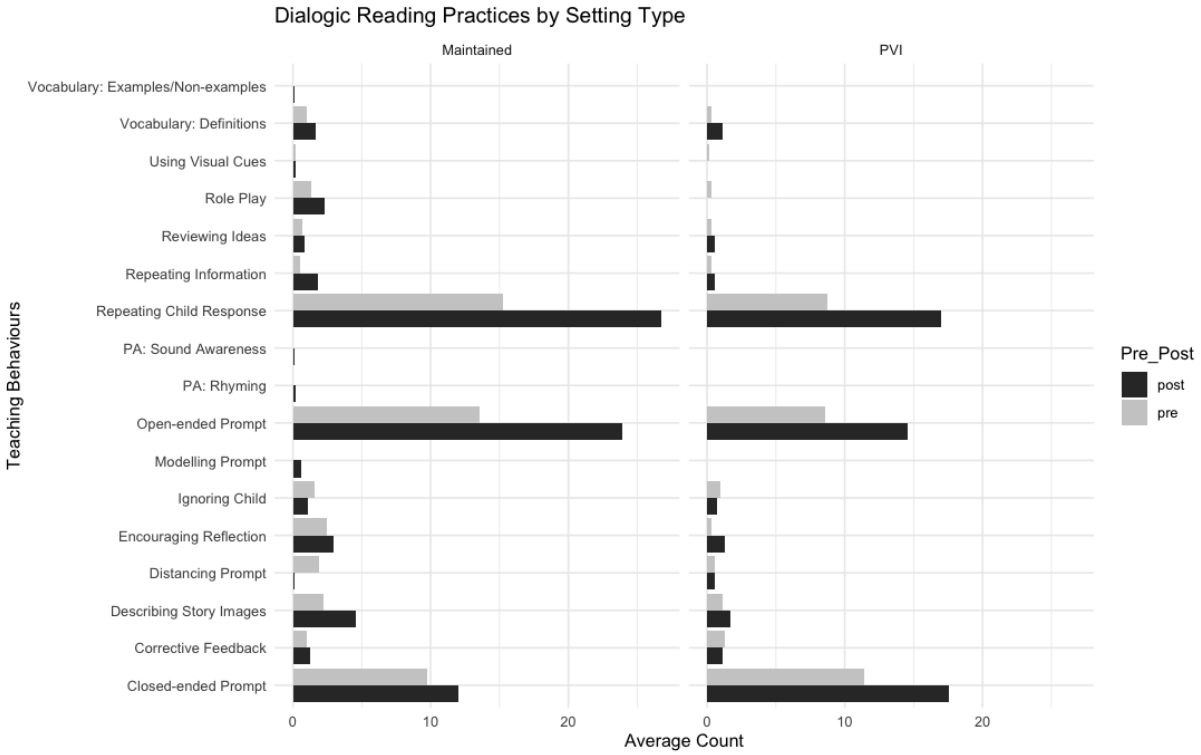
Figure 5: Frequency of dialogic reading practices observed during and around storybook reading



PA = phonological awareness.

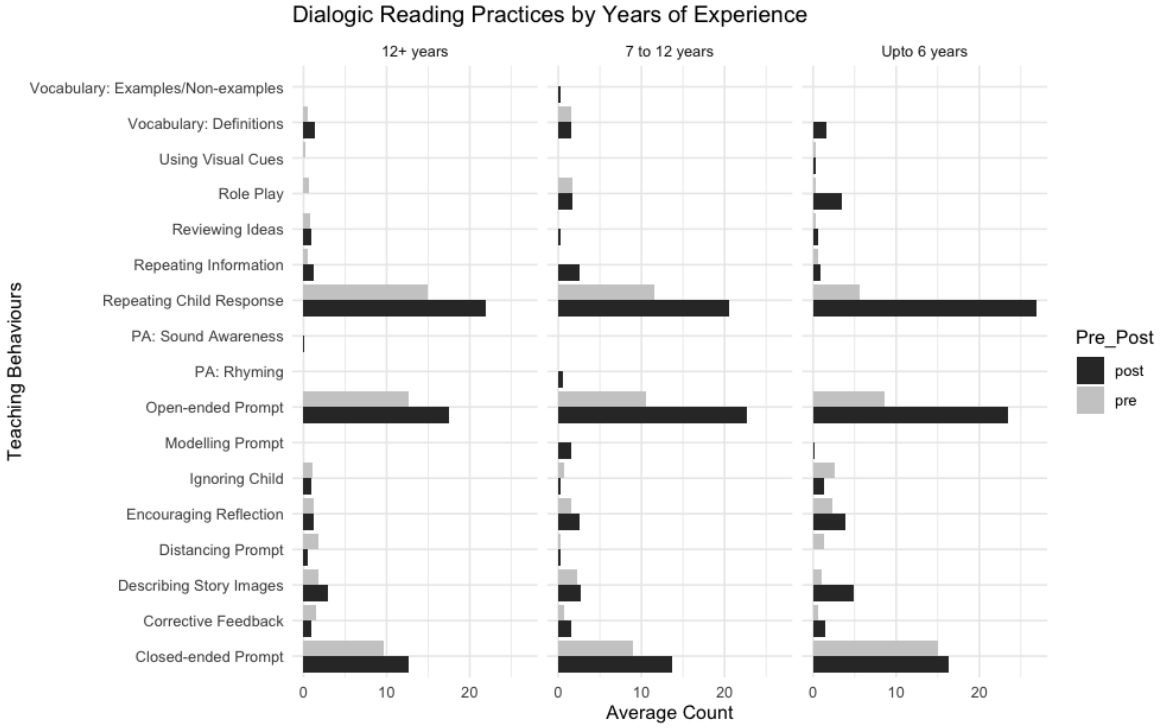
When comparing dialogic reading practices in maintained and PVI settings, we observed the same pattern of a noticeable increase in open-ended and closed prompts, and in repeating a child's response from baseline to post-intervention recordings (see Figure 6), with larger gains in maintained settings.

Figure 6: Frequency of dialogic reading practices observed during and around storybook reading by setting type



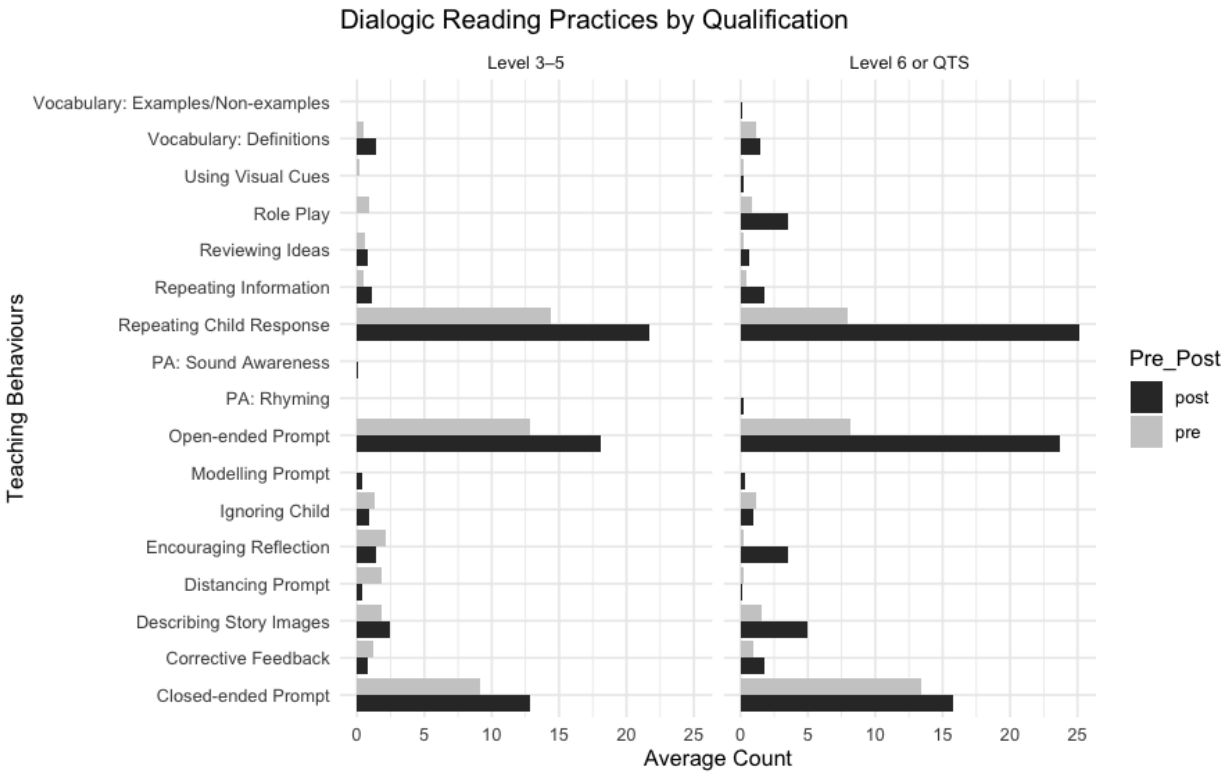
Comparing practitioners' use of dialogic reading practices by years of experience (Figure 7), this indicates the greatest gains in the frequency of use of open and closed questions and of repeating a child's response were in less experienced practitioners (less than 12 years).

Figure 7: Frequency of dialogic reading practices observed during and around storybook reading, by practitioners' years of experience in early years



Qualification level, on the other hand, indicates that greater gains in dialogic reading practices were observed in more qualified practitioners (Figure 8).

Figure 8: Frequency of dialogic reading practices observed during and around storybook reading, by practitioner qualification



In summary, evidence from survey, interview, and audio recording data indicates a change in practice following the TWiTCH programme. Self-report data evidences a significant increase in the use of language support strategies and storytelling practices, with the impact going beyond practitioners trained in TWiTCH strategies to the wider nursery and, in some cases, non-TWiTCH practitioners. In audio observations, noticeable differences were found in the amount of time practitioners spent on reading a story, and also the amount of time they devoted to discussing the story before and after the session, as a consequence of the programme. Open-ended, closed, or completion prompts and repeating what a child says were strategies which saw an increase in frequency following the TWiTCH programme. The success criterion that there are 'changes in practitioner practice in the majority of settings' has been met.

RQ11 How does practitioner confidence and knowledge for supporting children's language development change (i) immediately following training and (ii) throughout delivering the programme as a result of the TWiTCH programme?

a. Does this vary across settings, practitioners, and roles?

Data used to answer this question has been synthesised from baseline and post-intervention surveys, the training survey, and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

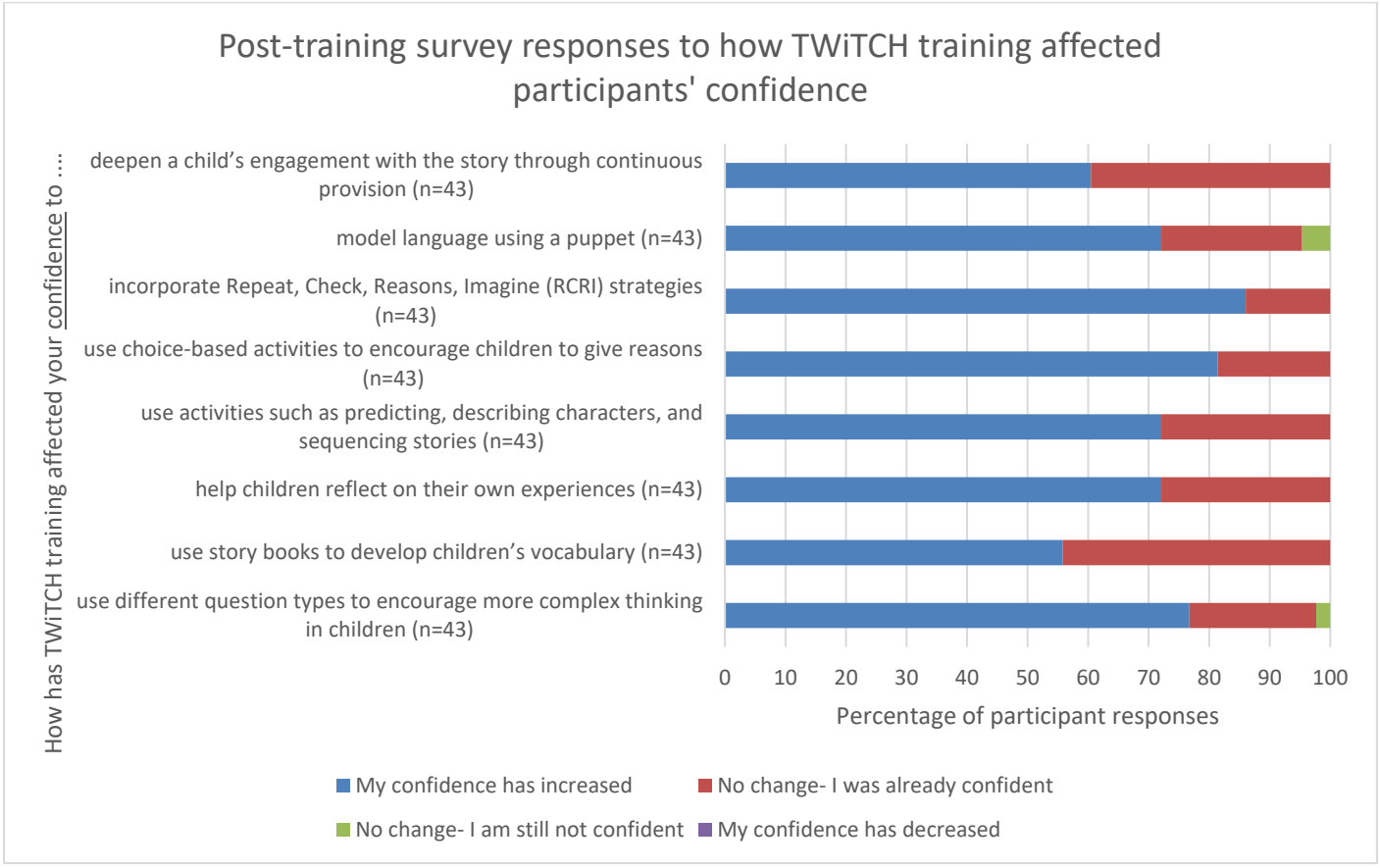
Confidence immediately following training

Practitioner confidence in supporting children's language development increased following training. In the post-training survey,[9] the majority reported 'my confidence has increased' for all elements of the TWiTCH programme (Figure 9), however, some—13% to 44% depending on the aspect of TWiTCH—reported no change as they were already confident prior to the programme. With regard to using continuous provision to deepen children's engagement and using storybooks to develop children's vocabulary, around 40% of practitioners reported no change due to already feeling confident in these areas. This indicates a perceived lack of impact for a group of practitioners. Very few reported still

[9] At the time of survey completion, four respondents had completed less than one cycle, 11 had completed one or two cycles, and 29 respondents had completed three or more.

lacking confidence and nobody reported a decrease. Increases were particularly large in RCRI strategies, choice-based activities, and using different question types. The TWiTCH programme may need to be specifically targeted at practitioners who do not already feel confident to use dialogic approaches with storytelling.

Figure 9: Change in confidence for different elements of the TWiTCH programme following the training (n = 43)



Knowledge immediately following training

Practitioner knowledge for supporting children's language development also increased immediately following training. The majority reported that the training had increased their understanding of all elements of the programme (see Figure 3). Increases were especially large in RCRI, choice-based language games, and using questions.

Confidence throughout delivering the programme

Practitioner confidence for supporting children's language development increased during the programme delivery period: it increased significantly from the baseline survey (M: 32.53; SD: 7.08) to the post-intervention survey (M: 38.72; SD: 5.30), with a significant effect size—*d* = 1.13 (see Table 14).

Table 14: Comparison of confidence scores for matched group of participants from baseline to post-test on confidence scale

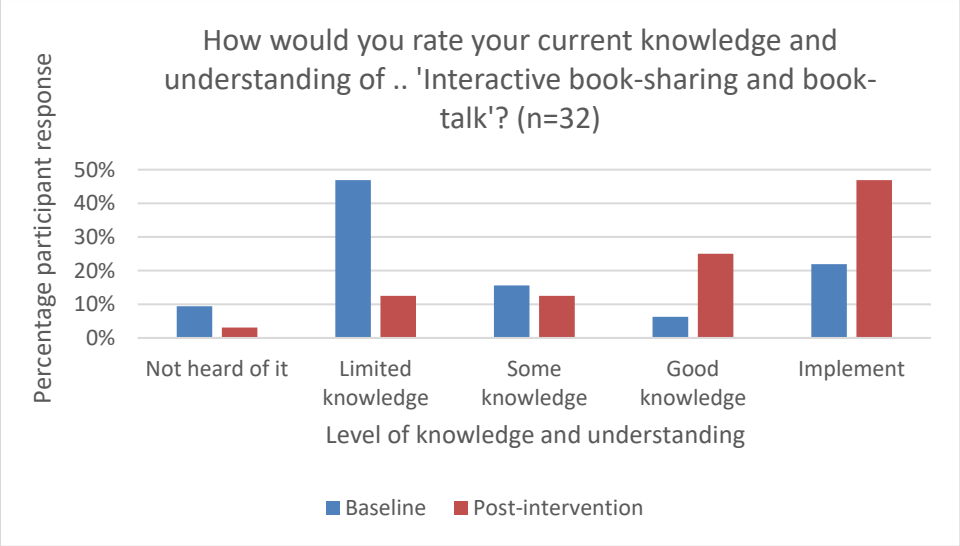
	Baseline mean (SD)	Post-intervention mean (SD)	95% CI for mean difference	df	T	Effect size Estimate Cohen's d (95% CI)
Confidence scale (n=32)	32.53 (7.08)	38.72 (5.3)	4.21–8.16	31	6.39 *	1.13 (0.68; 1.57)

* Significance at the p < 0.001 level.

Knowledge throughout delivering the programme

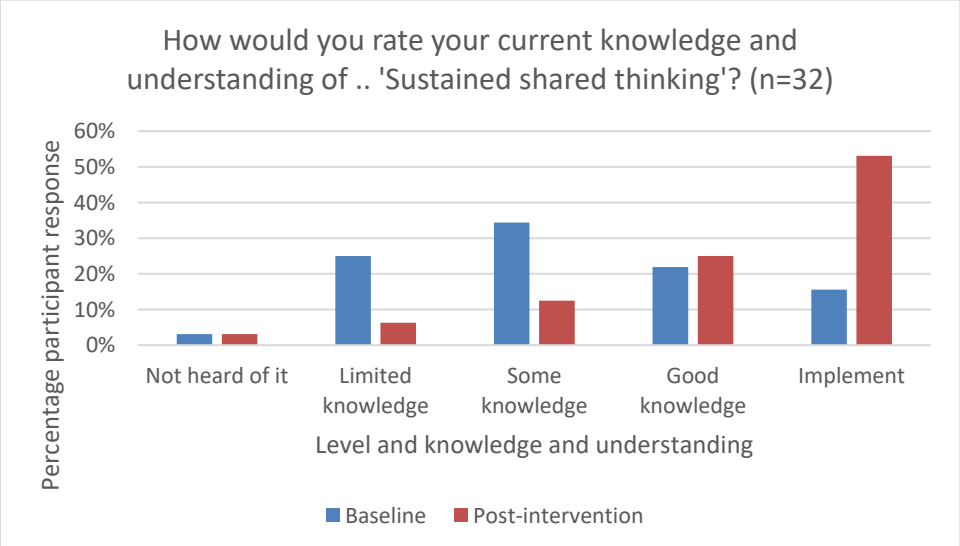
Practitioner knowledge for supporting children's language development increased during the programme delivery period. While at baseline only 28% of practitioners felt they had good knowledge or were implementing interactive book sharing, this was 72% after doing the intervention (see Figure 10). Similarly, with sustained shared thinking (SST) 37% felt they had good knowledge or were implementing SST but this increased to 78.1% by the end of the programme (see Figure 11). For the choice-based language activities, only 31% felt they had good knowledge or were implementing these in their practice, but this increased to 81% by the end of the programme (see Figure 12).

Figure 10: Change in interactive book sharing knowledge from beginning to end of delivering the programme



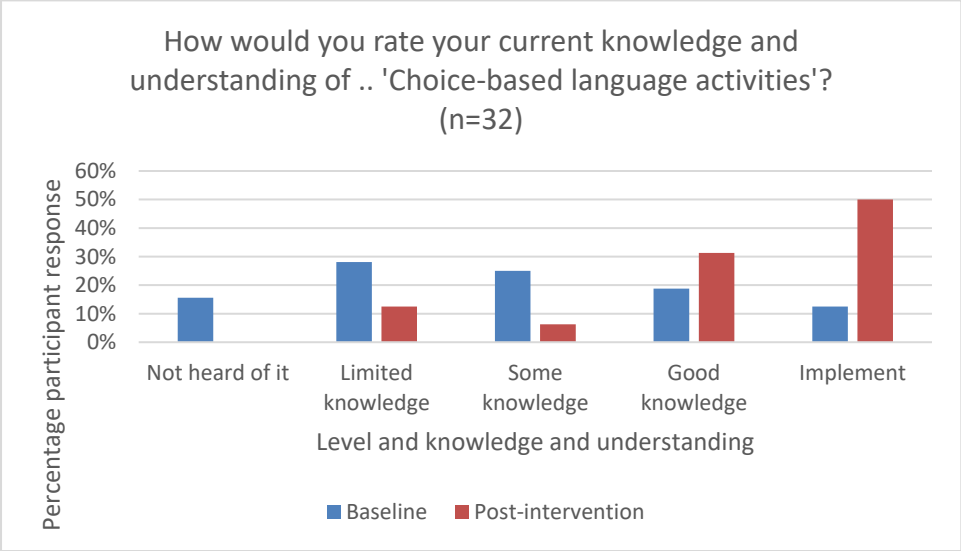
n = 32, matched sample.

Figure 11: Change in sustained shared thinking knowledge from beginning to end of delivering the programme



n = 32, matched sample.

Figure 12: Change in choice-based language activities knowledge from beginning to end of delivering the programme



n = 32, matched sample.

Practitioner comments in free text boxes in the post-intervention survey provide further evidence for increased confidence and knowledge following the programme. Practitioners were asked to describe the impact of the programme with the most common response being that it had improved their confidence (13 responses) particularly in reading stories and in using questioning. Many also described that TWiTCH has supported them with their knowledge and practice of using better questioning both while reading stories and sometimes in their other practice. Some practitioners described that they felt TWiTCH allowed them to have more in-depth discussion with children about stories and encouraged children to contribute more to the discussion. A few practitioners related that they now used the TWiTCH-specific techniques (for example, repeating and checking answers and giving children more time to respond), which they had not been doing before and that they had a better understanding storytelling practice (for example, the value of repeated readings, small group work, and specifically introducing new vocabulary). A small number felt that TWiTCH had increased the consistency of what children receive across the nursery.

> *'It has given me more confidence when sharing stories to be animated and the importance of repeating phrases and acting out the stories' (practitioner survey).*
>
> *'I am more confident in reading a story and asking questions' (practitioner survey).*

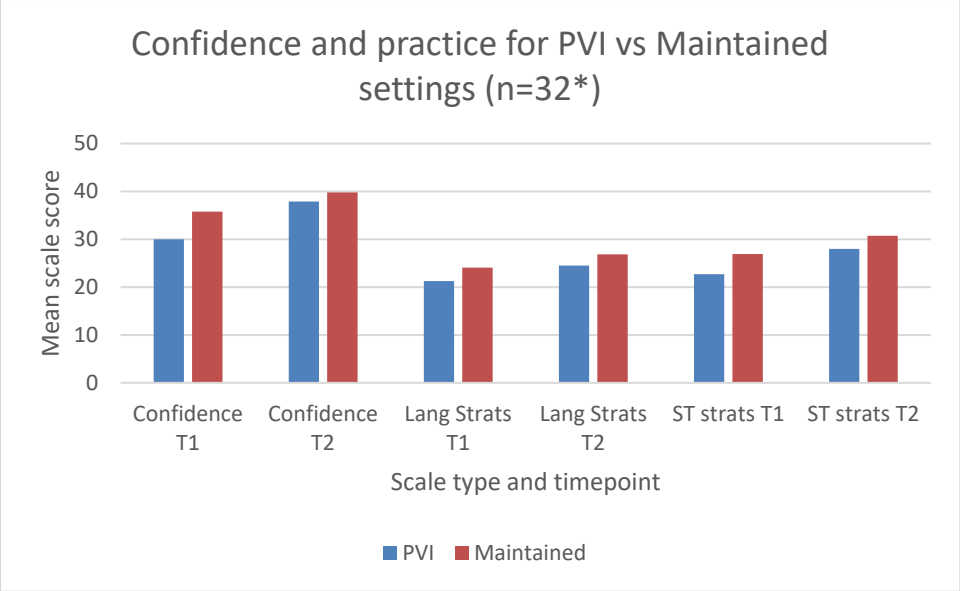
Data from the first champion interviews supports the survey findings: six reported that their confidence in reading stories improved, which in some cases was also mentioned by the senior leadership team about their staff (three responses). Confidence to use a puppet improved and some staff reported becoming more confident to adapt the programme for their setting as time went on. In terms of changed practitioner knowledge and understanding, some champions said their expectations of children's reasoning skills had been raised and that they could see the value in repeated reading of texts.

In the second interview, champions were asked about any impact the TWiTCH programme was having on staff. Six thought that an increase in confidence was the biggest impact on themselves or other staff delivering TWiTCH. This was predominantly in relation to telling stories, asking questions, or allowing children to respond and lead the direction of the discussion. One champion mentioned that staff in the setting were 'more confident to tailor it to the children now rather than just following the handbook as it's written'.

Variation in relation to confidence across settings, practitioners, and roles

Practitioner confidence and practice varied by type of setting. Higher levels of confidence and frequency of use of language-supporting and storytelling strategies were found for practitioners in maintained settings compared to PVI settings at both survey timepoints (Table 14): maintained settings expressed higher confidence at baseline compared to PVI settings, however, the latter appeared to improve in confidence to a greater degree than maintained settings, narrowing the gap in scores post-intervention.

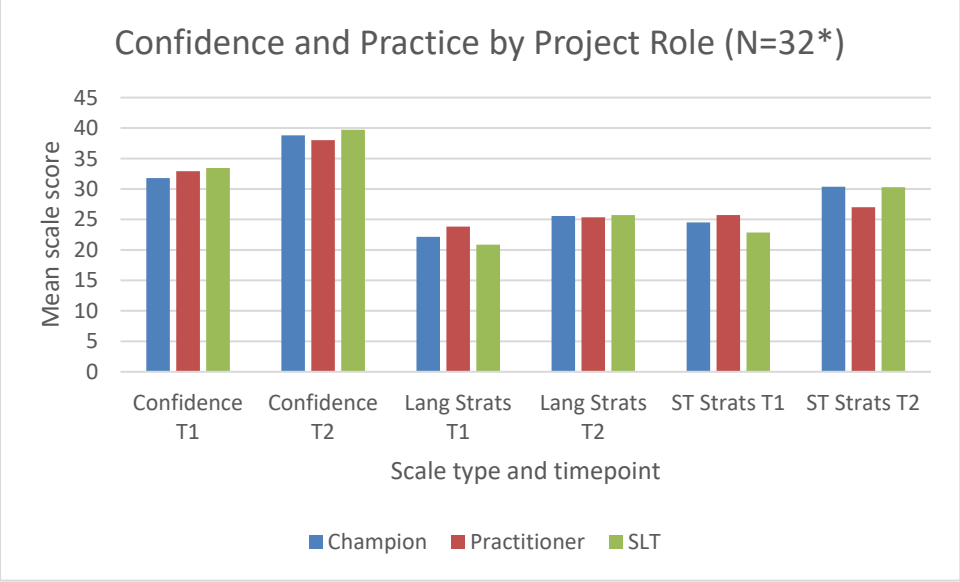
Figure 13: Comparison of practitioner confidence and practice scores between baseline (T1) and post intervention (T2) in PVI and maintained settings



PVI: n = 18; maintained: n = 14.
* One practitioner missing data at T1 on language strategies scale.

When comparing practitioner confidence and practice by role (Figure 14), SLT members showed higher levels of confidence compared to champions or practitioners but this was not reflected in more frequent use of language-supporting or storytelling practices at baseline. This may be because senior leaders are less frequently in the classroom. Champions and senior leaders saw greater gains in the use of strategies than practictioners during the project. All roles gain in confidence during the project.

Figure 14: Comparison of practitioner confidence and practice scores between baseline (T1) and post-intervention (T2), by practitioner role

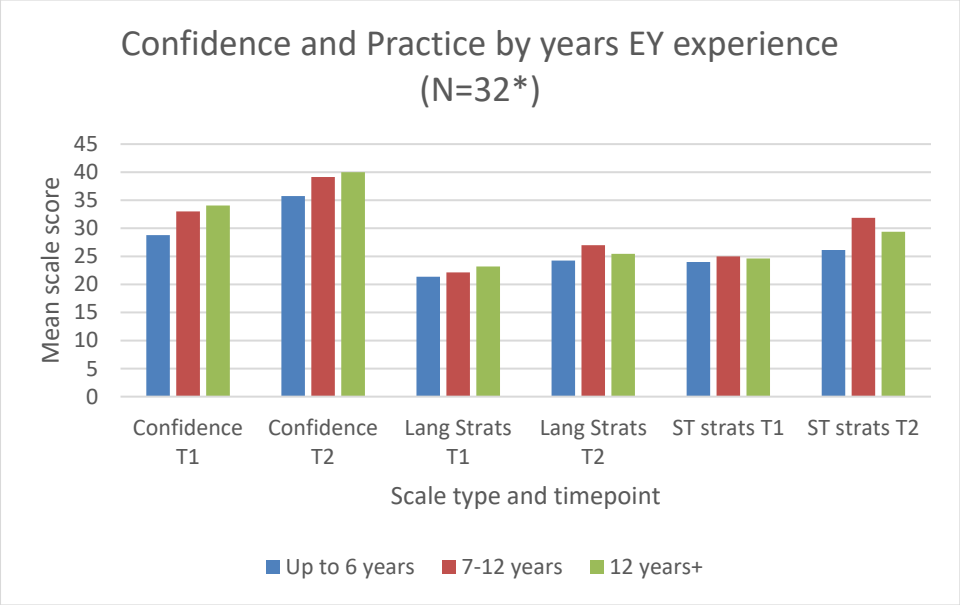


Champions: N = 14; practitioners: N = 11; senior leaders: N = 7.
* One practitioner missing data at T1 on language strategies scale.

Considering length of early years' experience in practitioners (Figure 15), confidence was higher in those with more experience at both baseline and post-intervention, with similar gains in confidence after the project regardless of experience. At baseline, use of language-supporting strategies was fairly comparable in practitioners, but greater gains were seen in those with seven to 12 years' experience. Practice using storytelling strategies did not vary much by

experience at baseline, but practitioners with the most experience (12 or more years) seemed to make the greatest gains.

Figure 15: Comparison of practitioner confidence and practice scale score between baseline (T1) and post-intervention (T2,) by years' EY experience



Up to six years' experience: N = 8; seven to 12 years: N = 8; more than 12 years: N = 16.

Confidence gain

Multiple regression analysis was run to predict gains in the practitioner confidence and practice scales from type of setting (PVI vs maintained), length of experience, or level of qualification. It was expected that due to fewer opportunities for CPD, practitioner confidence and practice may improve to a greater extent for PVI settings than for maintained settings. It was also expected that greater gains may be observed for those with fewer years of experience in early years and those with a lower level of qualification.

For the gains in the confidence scale, Step 1 of the model including setting type, indicated that setting type predicted change in confidence with a higher level of confidence gain for PVI settings, however, this only explained a small amount of the variance ($R^2 = 0.13$, $p < 0.05$ for Step 1). On addition of the other predictors at Step 2 the model was not significant ($R^2 = 0.16$, $p = 0.17$) and added little additional prediction of the change to scores (Appendix D.4). It should be noted again, however, that the sample size was likely too small for the use of multiple regression and therefore this analysis is underpowered.

Gains in practice for use of general language-supporting strategies and storytelling strategies

Again, multiple regression analysis using the change in practitioner use of general language strategies and the change in use of storytelling strategies from baseline to post-intervention as dependent variables was used. This followed the same analysis structure as for confidence gain. None of the predictors explained a significant amount of the variance and the models were not significant (see Appendix D.4). Gains in the use of strategies for storytelling or in general practice did not seem to be predicted by the type of setting, the qualification level of the practitioner, or their early years' experience.

In summary, practitioner confidence and knowledge for supporting children's language development did change following training and through delivering the programme. The most noticeable increases in confidence were seen in RCRI strategies, choice-based activities, and using different question types. There was a statistically significant increase in self-reported practitioner confidence compared to the baseline survey. Free text survey responses and interview data support this finding, with many practitioners reporting improved confidence as the biggest impact of the programme. Confidence was higher for practitioners in maintained settings than for those in PVI settings and for the latter, confidence seemed to improve more through the programme. However, when aggregating the data by setting type, practitioner role,

and years' experience, the findings were not statistically significant in the underpowered multiple regression analyses undertaken. In terms of the success criteria, the data supports the view that there has been 'increased practitioner confidence and knowledge in supporting children's language development across the majority of practitioners' and that this criterion has been met.

RQ12 What impact do practitioners feel the programme is having on children's language development? Are there any unintended outcomes?

Data used to answer this question has been synthesised from baseline and post-intervention surveys and interviews with practitioners from case study settings (n = 5) and interview settings (n = 12).

Impact on children's language development

Practitioners described the impact they thought being part of TWiTCH had had on children in the nursery in the post-intervention survey. The most common responses were that TWiTCH had led to:

- improvements in children's communication, language, and listening skills;
- children having a deeper understanding of the story and improved ability to retell stories;
- improved confidence in children to talk as part of a group; and
- greater use of reasoning in language, for example, giving more reasons or asking why.

Practitioners reported that children enjoyed the TWiTCH sessions. Some described how TWiTCH engaged children who were not normally engaged, that it supported quieter children to be able to contribute more, and that the repeated reading was particularly useful for EAL children or those with language delays. A few had also seen children using TWiTCH vocabulary and stories across other activities. A small number of practitioners described seeing improved concentration in children, in a wider range of books. One felt that TWiTCH had not been as impactful as a previous intervention in terms of language gains.

Similar reports of such language development gains were shared by TWiTCH champions in their first interview. Champions said they had seen benefits for the children especially in their comprehension and engagement with books. They also noticed that children were more confident in expressing themselves and could concentrate better for the book reading sessions. Children were reported as engaging well with TWiTCH and looking forward to sessions and that an increased enjoyment of reading and stories overall had been observed. Champions noted that children's understanding of TWiTCH stories has grown through repeated reading and that they were better able to retell stories and talk about characters. Children were talking about stories at home and to other staff. Improvements in language and vocabulary were reported by some champions and in children's ability to give reasons using 'because'. In Case Study Setting 1, the champion felt that children were talking more.

> *'They often come round and say, "I'm using my 'because' today,", and they're giving their reasonings to pretty much anything that they like to say.'*

In the second interview, champions from four settings explicitly mentioned language development as an observed outcome. For example, one described how vocabulary had improved:

> *'So definitely for us, we're seeing language acquisition improving for the children. So children are using a variety of vocabulary that … we wouldn't have necessarily heard.'*

The same champion gave more detail about how this extended beyond the TWiTCH sessions:

> *'And generally in the children's free play, they're using the language, but they're using sentence structure as well. So, things like the "I wonder" questions and "What happens if…?" and "How do you feel about…?"—they're using some of those questions when they're in free play, which is something that we haven't really seen before, but because it's modelled and used really regularly, it's becoming part of their natural way of speaking ... because of the repetition and the familiarity of the story, I feel they're learning more of those phrases, and then they're bringing them into the play, which is only enriching the play for the children.'*

The champion in another setting also emphasised how children's language had improved and that they were seeing examples of it used in free play. Another commented on language improvement in the 'more able children' specifically

in terms of reasoning and asking questions themselves, as well as noticing improvements in the 'lower' children. One setting had noticed an improvement in children's ability and confidence to recall a story, and this had been evidenced through recent literacy assessments. In the same setting, 'reluctant talkers' were benefiting from the repetition and small-group work to talk in front of others.

When asked about the impact of TWiTCH on children, in addition to the above-mentioned language improvements, champions mentioned greater confidence (n = 3), an increased interest in reading (n = 2)—particularly in boys in one setting, better quality imaginative play (n = 1), and improved concentration to listen to a story for a longer period (n = 2), even in two-year-olds, which the champion believes is in part due to the rabbit and the children's engagement with it. There was also more interest and excitement in reading and books (n = 2) for children.

In one case study setting, teachers in the children's future primary school had noticed children's high levels of vocabulary when they were on their settling-in days during the summer term. These teachers commented that the way children were building their sentences and using reasoning was advanced. The case study attributes this not only to TWiTCH sessions but also to using this new language beyond the TWITCH sessions through play and with their peers:

> *'You often hear the children squabbling, I would say. And they would say to each other, "Oh, you're not using your 'because you're not telling me why'." And we've noticed that has come about since starting the TWiTCH programme, which has been nice.'*

The developer team reported that practitioners in some settings had raised their expectations of children, which they saw as a positive of the programme. For example, children with SEN have been able to take a more active part in TWITCH than practitioners thought they would, particularly due to the repetition of the storytelling, which has allowed these children the time to access the story and then engage with it.

Other unintended outcomes

Parental engagement and home learning was an unintended outcome of the TWiTCH programme. A champion in a case study setting commented in the first interview that they had noticed 'the engagement or the want for engagement from the parents'. This had emerged from 'the children going home and talking about the stories that they've read and then also the new words that they've learnt as well from the stories too'. By the second case study visit and interview, the setting had responded to parents' engagement and had contacted them using their app to explain TWiTCH. Additionally, they then created a corner in the setting for parents to be able to see the resources used for TWiTCH, the handbook, and then ask any questions that they might have. Parents have reported that they have been reading stories at home with their children, implementing TWiTCH strategies as best they can. A champion in another setting described how they had also shared information with parents on their nursery communication app :

> *'And then we often get—in the mornings or when they get dropped off or picked up—the mums or dads will mention retelling the stories … the story's going home. They're talking about it. And then sometimes they'll ask, like, who's twitch? They talk about Twitch a lot.'*

The developer team also picked up on parental engagement as an unintended outcome during their second interview. One outcome had been that some children were taking the story into the home and introducing their siblings to TWiTCH. The developer team thought this, in addition to the above, meant that 'parents are getting nudged in lots of ways', which was a positive outcome of TWiTCH.

Enjoyment of the TWiTCH programme was an unexpected outcome for the developer team (second interview):

> *'I don't think we envisaged practitioner enjoyment … But given the, you know, retention and recruitment crisis, something that actually gives pleasure is actually quite important.'*

The enjoyment of the children was anticipated, but the effect that it had on practitioners was not expected:

> *'Yes, a lot of, like, displays … created in the storybook corner, or the continuous play. I mean, the best one, I think, was Jack and the Beanstalk, where they'd made this massive tree and, you know, they'd done all these bits and pieces and … they were enjoying it themselves, so they were putting all this extra work that they didn't have to do, obviously… almost, like, feedback that the children are enjoying it and are engaged, and the professionals are feeling that they're doing a good job, because they can see the levels of engagement and enjoyment that the children are showing.'*

In interviews, some champions were asked about any negative impact experienced through doing TWiTCH. While in some settings (n = 3) champions specifically commented that they had not experienced any negative consequences, a small number did mention some negative impacts (n = 3). The champion in one setting commented that they felt TWiTCH directed attention away from other things that they should be doing in the nursery. Another champion felt the same, specifically that they would have spent more time on phonics had they not had to give their full attention to the TWITCH programme. A third champion thought that the other 'focus groups' their setting was delivering 'had to be maybe sacrificed the wrong way'. Given the time and space demands on TWiTCH, any future scaled-up evaluation should investigate the potential loss of other activities in nursery to accommodate the TWiTCH programme.

In summary, survey and interview data indicates that many practitioners felt that TWiTCH did have an impact on children's language development. This was particularly noticeable in children's ability to retell a story, their use of reasoning language and vocabulary, and confidence in expressing themselves. These findings indicate that the success criterion, 'practitioners in the majority of settings perceive TWiTCH to have a positive impact on children's language development', has been met. Unintended outcomes of TWiTCH included parental engagement and transfer of learning to the home environment, and practitioners' enjoyment of the programme. The only negative impact mentioned in a few settings was having less time to focus on practice in other areas of development that they had been able to focus on before TWiTCH.

Success indicators

Table 15 shows the extent to which the success criteria have been met by the results above. In the final column, green indicates that the criteria have been met and orange indicates this has not been fully met.

Table 15: Indicators specified at the start of the evaluation to measure the success of the TWiTCH programme

Pilot criteria	Success Indicators	How to assess this?	Research question	Criteria met?
Evidence of promise	Changes in practitioner practice in the majority of settings.	Audio recordings, survey, case studies, interviews, coach feedback.	10	Fully met
	Increased practitioner confidence and knowledge in supporting children's language development across the majority of the practitioners.	Surveys, interviews.	11	Fully met
	Practitioners in the majority of settings perceive TWiTCH to have a positive impact on children's language development.	Interviews, case studies.	12	Fully met
Feasibility of implementation	Programme is delivered in settings of different types with medium to high fidelity as judged by the developer team.	Developer team interview, coach feedback, interviews with TWiTCH champion and practitioners.	6	Mostly met
	Initial training sessions are attended by minimum of three practitioners per setting for majority of settings and are accessible to practitioners.	Training attendance records, post-training surveys.	6	Fully met
	At least six TWiTCH cycles completed by majority of settings.	Coach feedback, surveys, interviews.	2	Fully met
	Absence of any major barriers to delivery for majority of settings.	Interviews, case studies.	7	Some moderate barriers raised but setting managed to deliver

	TWiTCH programme seen as acceptable and accessible to majority of practitioners.	Surveys, interviews, case studies.	4	Fully met
Readiness for trial	TWiTCH training and support is delivered consistently across different settings and trainers.	Observations.	3	Not fully met
	TWiTCH programme sufficiently described to enable consistent delivery.	Interviews, coach feedback.	3	Not fully met

The success criteria were deemed to be fully met for all areas excluding four. In three of these four areas there were aspects identified that had not been fully met while in one area it was approaching being fully met.

For the 'programme being delivered in settings of different types with medium to high fidelity' criterion, this was deemed to be 'mostly met' as most settings (PVI and maintained) were delivering the programme in line with the TWiTCH guidance, however, some settings were not specifically targeting children in Week 3 for additional work and in some of settings (especially larger ones) group sizes were larger than specified or TWiTCH was being done with only some of the nursery children.

For 'absence of any major barriers to delivery for the majority of settings' we highlight the issue again that not all children were receiving the intervention due to settings having logistical issues with space and staffing to deliver the intervention in small groups to all children. In Readiness for Trial, these two items have been flagged as not fully met due to the variability seen in delivery of the TWiTCH training by coaches and in the delivery of Week 2 sessions during the programme. Both these issues have been highlighted elsewhere in the report.

Cost evaluation

Based on information shared by the developer team in two cost workshops and using EEF costing guidance, we calculated the cost of delivering the TWiTCH programme during this pilot. We then used assumptions about an ongoing model for TWiTCH for continued delivery within a setting for a further two years to calculate the total cost of delivery for the 25 settings in the pilot. This allowed us to calculate a cost per setting for the programme and a cost per child. Since this pilot was subsidised by the EEF, settings did not pay anything towards the programme and were provided with additional funds for cover, which are included in the cost calculations here.

Two further calculations were made for the cost implication for nursery settings: (a) based on modifications made for Phase 3 and as an ongoing package and (b) as a three-year programme. The two other calculations forecast the cost implications for nursery settings to purchase the programme.

Table 16 details a breakdown of the programme costs of the TWiTCH programme as it was delivered during this pilot project for the Year 1 delivery. The delivery period these assumptions are based on are June 2023 to July 2024.

Table 16: Costs of TWiTCH programme delivery for this project

	Type of cost	Item	£
YEAR 1 DELIVERY	**Developer personnel for preparation and delivery**		
	Set-up cost: staffing	Developer and lead team (1 x senior lecturer and 1 x professor) Prepare and deliver training for coaches Prepare and deliver coach meetings Development of handbooks and guidance in line with feedback Provide ongoing support to nursery settings Prepare and deliver online champion meetings	£51,900
	Set-up cost: staffing	Project manager Support with recruitment of nurseries Liaise and support nursery settings throughout the programme Liaise and support coaches throughout the programme Create and prepare materials for TWiTCH packs Organise delivery of TWiTCH packs to nursery settings	£39,875
	Set-up cost staffing	Project officer Administrative support throughout the project	£1,500
	Set-up cost staffing	Five TWiTCH coaches (lecturer/early years professional) 22 days per coach, at £350 per day, 5 settings per coach Attend coach training and meetings Prepare and deliver in-person practitioner training Visit settings for observations and coaching, and offer ongoing support online	£38,500
			Travel costs
	Set-up cost: travel	Travel for TWiTCH coaches to in-person meetings	£6,981
	Resources for preparation and delivery		
	Set-up cost resources	2 x resource packs for each of the 25 settings (at £510 per resource pack)	£25,506
		1 x resource pack for each of 5 coaches	£2,550
		Delivery of resource packs to nursery settings via courier, or mileage	£400
		Additional costs (photocopying, additional packs etc.)	£8010
	Setting costs		
	Set-up cost: cover to attend training	Cover costs for practitioners to attend TWITCH training, plan for delivery, and for champions to attend four online meetings (£175 per setting) *	£4,375
		Additional costs, for example, other books, printing, additional manuals (cost across all settings)	£250
		Total cost Year 1	£171,917
YEAR 2 DELIVERY	**Developer personnel for delivery**		
	Ongoing: coach support	Five TWiTCH coaches; three days per coach at £350 per day.	£5,250
		Total cost Year 2	£5,250
YEAR 3 DELIVERY	**Developer personnel for delivery**		

	Ongoing: coach support	Five TWiTCH coaches; three days per coach, at £350 per day.	£5,250
		Total cost Year 3	£5,250
TOTAL COSTS			
	Total cost for delivery over 3 years		£182,417
	Cost per setting for 3 years delivery		£7,296.68
	Cost per setting per year of delivery		£2,432.23
	Cost per child per year of delivery—assuming mean number of children per setting in pilot (32.8) per setting per year		£74.15
	Cost per child per year of delivery—assuming maximum number of children in a setting in pilot (100) per setting per year		£24.32
	Cost per child per year of delivery—assuming minimum number of children in a setting in pilot (11) per setting per year		£221.11

* Cover costs here are based on the £175 provided by the developer team to settings to enable them to attend training. This was offered to all settings and most settings invoiced for this.

The total cost for the first year of programme delivery would be £171,917. While the cost evaluation did not include costs for undertaking the evaluation activities, the TWiTCH programme was being delivered as part of a research project, which meant delivery was still affected by these elements which included additional time spent on recruitment of particular settings, administration time spent on reimbursing cover costs, and making changes to programme delivery during the programme in response to feedback. Including staff time for these elements means that this figure is likely to be a slight overestimate of the costs of delivery outside of a research project.

Discussions with the developer team highlighted that they would not expect ongoing delivery of TWiTCH for a second and third year to require any additional materials but would continue the ongoing relationship with each TWiTCH coach through three, one-hour online meetings per year per setting. Assuming that coaches would continue to support five settings each this would mean three days of time for each coach (15 days in total) at £350 per day—a total of £5,250 for delivery in each of the second and third years.[11] This cost assumes a continuing relationship with the same coach and that further training for coaches is not required: if further training is required, there would be additional costs for the delivery of the programme should new coaches be required and trained but we have assumed these costs would be only incurred every three years and that they are covered within the costs for the first year of delivery. The total cost for delivery across the three years would therefore be £182,417.

With 25 settings participating in the programme, this model of TWITCH delivery would cost a nursery setting £7,296.68 for three years of delivery. The cost per year per setting when assuming a three-year delivery period would be £2,432.23.

Over three years of delivery, three cohorts of children would be receiving the TWiTCH programme. Using the mean number of children in the evaluation settings (M: 32.8; range: 11–100) we assumed that 98 children per setting would receive the programme over the three years. This would mean a cost per child of £74.15 per year for a nursery with similar numbers of children to the average. However, as the TWiTCH model has been designed to be offered as a one-off setting cost, which would be the same for all nurseries rather than a per-child cost, this would mean that smaller settings would have a much higher rate per child compared to larger settings. This would mean that for the smallest cohort in our data (n = 11 three- and four-year-olds) the cost per child per year would be £221.11 and for the largest cohort (n = 100 three- and four-year-old children) the cost per child would be £24.32.

Estimated costs for an ongoing delivery model of the TWiTCH programme to nursery settings, as informed by the delivery model for Phase 3 and as discussed with the developer team, are detailed in Table 17. These costs are based on programme delivery from June 2024 to July 2025—after this trial (13 months). Personnel costs have been calculated

[11] There may also be some ongoing administration costs for the developer team, but these are likely to be fairly low for subsequent years of delivery and therefore have not been included in the calculation.

using the job role and grade for these roles (as supplied by the developer team) and the corresponding starting salaries for these grades as per academic role salary scales.[12]

Table 17: Cost of TWiTCH programme delivery as an ongoing model beyond the pilot

	Type of cost	Cost	Cost per item	Number of items	Total
YEAR 1	**Developer personnel for preparation and delivery**				
	Set-up cost	1x TWiTCH developer and lead (senior lecturer, Grade 8) Prepare and deliver training for coaches to deliver practitioner training Prepare and deliver coach meetings Provide ongoing support to nursery settings Prepare and deliver online champion meetings	£45,585	0.4 FTE	£18,234
	Set-up cost	1 x TWiTCH project manager and administrator (senior administrator, Grade 6) Liaise and support nursery settings throughout the programme Liaise and support coaches throughout the programme Create materials for TWiTCH packs and order resources Organise delivery of TWiTCH packs to nursery settings	£31,396	0.6 FTE	£18,838
	Set-up cost	1 x TWiTCH financial support (senior administrator, Grade 5) Oversee financial aspects of the programme, such as receiving payments and issuing receipts	£25,742	0.1 FTE	£2,574
	Set-up cost	5 x TWITCH coaches (lecturer/early years professional) Attend coach training and meetings Prepare and deliver in-person practitioner training Visit settings for observations and coaching, and offer ongoing support online	£400 per day per coach	24 days per coach	£48,000
	Resources for preparation and delivery				
	Set-up cost	1x TWiTCH resource pack for each of the 5 coaches	£500	5	£2500
	Set-up cost	2x TWiTCH resource packs for each of the 25 settings	£500	50	£25,000
	Set-up cost	Delivery of resource packs to nursery settings via courier or postage	£400	1	£400 *
	Set-up cost	Additional printing costs/ extra copies of manuals	£10	25	£250
	Travel costs				
	Set-up cost	Travel costs for TWiTCH coaches to in-person meetings	£6,981	1	£6,981 **

[12] We used the Durham University guidelines for these calculations.

	Set-up cost	Travel for 1 TWiTCH developer to attend coach training/meetings at 3 timepoints	£200	2	£400 ***
		Total Year 1 cost to deliver the TWiTCH programme to all settings			£123,177
YEARS 2 AND 3		**Ongoing support for each nursery setting**			
	Ongoing cost	TWiTCH coach support of 3 x online meetings per year (1 hour per meeting and 1 hour for preparation time), per setting per year	4 hours	2	£12,000
		Total Year 2 and 3 costs to deliver the TWiTCH programme to all settings			£135,177

* This number is subject to review as it would depend on the location of settings, which may be further afield and therefore more expensive.
** This cost is subject to change and is dependent on the location of coaches and the location of settings in any future TWiTCH programme.
*** This cost is subject to change and is dependent on the location of settings in any future TWiTCH programme, and whether or not the developer would have to stay overnight.

The total cost of TWiTCH using the Phase 3 delivery model is estimated to be £123,177 in the first year of this revised calculation and £135,177 for delivery across three years, less than costs for the pilot programme. With 25 participating settings, this amounts to a setting cost of £5,407.08 for three years of delivery and a cost of £1,802.36 per setting.

The time commitment for practitioners to deliver the programme over the first and subsequent two years is outlined in Table 18. Estimates for the time practitioners need to initially plan to deliver TWiTCH and ongoing preparation times have been informed by discussions with practitioners during interviews and through feedback from the developer team. For the delivery over a second and third year, the time presented is for each year but is expected to be the same for both years due to delivery with different cohorts of children.

Table 18: Total time required of nursery personnel for training, preparation, and delivery

Type of cost	Staff member	Activity	Activity time (hours)	Number of staff	Overall time (hours)
YEAR 1					
Set-up	Senior leadership	Meet with TWiTCH coach	0.5	1	0.5
Set-up	Nursery staff delivering TWiTCH	Attend TWiTCH training 2 x 3 hours	6	3	18
Set-up	Nursery staff delivering TWiTCH	Plan TWiTCH implementation	1	3	3
Ongoing	Nursery staff delivering TWiTCH	Weekly preparation for delivery (each week for 27 weeks)	0.3	3	24
Ongoing	Nursery staff delivering TWiTCH	Delivery of TWiTCH sessions (Week 1: 5 x 20 mins; Week 2: 3 x 20 mins; Week 3: 3 x 20 mins) over 9 cycles assuming delivery of full TWiTCH programme by 3 practitioners	33	3	99

Set-up	TWiTCH champion (practitioner or SLT)	Attend champion training session	1	1	1
Ongoing	TWiTCH champion (practitioner or SLT)	Attend champion network meetings (3 x 30 mins)	1.5	1	1.5
Ongoing	TWiTCH champion (practitioner or SLT)	Attend online meetings with coach (3 x 1 hour)	3	1	3
Ongoing	TWiTCH champion (practitioner or SLT)	Liaise with coach throughout programme	1	1	1
Ongoing	TWiTCH champion (practitioner or SLT)	Observation and feedback from coach (3 x 2 hours)	6	1	6
Ongoing	TWiTCH champion (practitioner or SLT)	Support other practitioners in setting	1.5	1	1.5
YEAR 2 and YEAR 3—time requirements per year					
Ongoing	Nursery staff delivering TWiTCH	Weekly preparation for delivery (20 minutes each week for 27 weeks)	0.3	3	24
Ongoing	Nursery staff delivering TWiTCH	Delivery of TWiTCH sessions (Week 1: 5 x 20 mins; Week 2: 3 x 20 mins; Week 3: 3 x 20 mins) over 9 cycles assuming delivery of full TWiTCH programme by 3 practitioners	33	3	99
Ongoing	TWiTCH champion (practitioner or SLT)	Attend champion network meetings (3 x 30 mins)	1.5	1	1.5
Ongoing	TWiTCH champion (practitioner or SLT)	Attend online meetings with coach (3 x 1 hour)	3	1	3
Ongoing	TWiTCH champion (practitioner or SLT)	Support other practitioners in setting	1.5	1	1.5

Contribution analysis results (how the theory of change was supported)

The contribution analysis framework has four steps which allow us to assess whether the TWiTCH programme has contributed to achieving its intended impacts. The first step is that the programme is based on a reasoned theory of change, which is described in the Introduction section for TWiTCH and included in Appendix C. This was put in place before the programme commenced and consideration was given to the assumptions and causal pathways across the three levels of coaches, practitioners, and children through a series of workshops with the developer team and the EEF. These workshops also considered the strength of existing evidence supporting the theory of change and the theory of change was revised in line with the evidence and discussion. The framework of data collection for the evaluation was then developed based on this reasoned theory of change to collect data to evidence described causal pathways with priority given to the practitioner pathway where most data collection was focused. In Step 2 we looked at whether the activities and inputs of the programme were implemented as planned and in Step 3 we looked at whether the expected chain of results from the theory of change occurred. We have taken a systematic approach of evaluating each step of the theory of change to look at whether each aspect (across the inputs, outputs, outcomes, and impact) has been achieved, the evidence we have to support our conclusion, and whether we believe it is possible to attribute this to the TWiTCH programme. In Step 4 we looked at whether there were any contextual factors that may have influenced the results found and the extent to which changes can be attributed to the programme and this is also presented below. The results of this analysis are presented in appendices C.5 to C.8 with colour coding indicating whether the aspect was evident and whether there is evidence that this aspect contributed to outcomes. A summary of what this analysis found is presented below.

Step 2: 'Were the activities within the programme implemented as planned?'

In summary, for Step 2, the evidence indicates that most of the programme activities were implemented as stated in the theory of change model.

At the coach level, coaches were recruited and sufficient training (although not consistent for the five coaches) and resources were provided for these coaches to provide CPD and support the settings taking part.

At the practitioner level, practitioners taking part in TWiTCH attended the training provided by coaches in the settings, although in many cases this was not attended by all practitioners working with the three- to four-year-old age group. Practitioners received and used the resource packs provided. Practitioners also received ongoing support from their coaches through in-person mentoring visits and, in most cases, through remote contact in between visits. TWiTCH champion training took place and was attended by the majority of champions; however this training did not follow the plan set out. Online network meetings for champions also took place throughout the year, with all settings being involved in at least one, but attendance and engagement in these sessions varied.

At the child level, in all settings three- to four-year-old children were receiving the TWiTCH storytelling and activities delivered by their practitioners. However, this was not all children of this age in every setting; different settings took different approaches to which children were included. Most delivered eight or nine cycles of TWiTCH using the provided TWiTCH books, however, few settings had progressed to using TWiTCH techniques with their own books for a tenth cycle. Children were generally accessing TWiTCH as part of a small group of up to eight children, but there were a few settings when delivery was done with a larger group. There was variation and adaptation of how the TWiTCH activities were delivered with Week 1 activities and Week 3 continuous provision having greatest fidelity; Week 2 activities were significantly adapted for delivery and the Week 3 specific targeting of 'at need' children was not evident.

Step 3: Did the expected chain of results occur?

In summary, for Step 3, the evidence indicates that the causal pathways from programme inputs through to outputs and outcomes has been fairly well supported and that this means the expected chain of results did occur.

For practitioners, the findings show that the training led to improved knowledge and confidence for delivering the programme and that as a result of taking part and delivering the activities, practitioners improved their confidence and knowledge for supporting children's language development and changed their practice in the way that they interact with children both in normal practice and particularly during story-time sessions. Changes to practices in storytelling were specific to the aspects that the TWiTCH programme was aiming to improve, while no improvements were seen on other aspects of storybook reading (phonological awareness, print instruction) that were not related to the TWiTCH programme. Practitioners appreciated the support of their coaches for delivery and those in champion roles also improved their confidence in supporting other staff in the setting from beginning to the end of the programme.

For children, practitioners reported improvements in children's language and oral communication, which they believed were due to taking part in the programme. Children were reported as having better comprehension of the stories, of using story and TWiTCH vocabulary outside of TWiTCH sessions and improving in their ability to retell the story and talk about the characters. Children were also reported to have increased confidence to express their ideas in a group and to have increased their use of reasoning language—even outside of TWiTCH sessions. Children's language will improve significantly through natural development, and it is difficult to be certain that improvements in children's language were due to TWiTCH, however, the reported outcomes were expected by the theory of change and most are fairly specific to the TWiTCH activities adding greater confidence that they occurred due to TWiTCH. Data on children's outcomes is not collected directly from children but from practitioner report, meaning that we are we are less secure in this finding.

For children, investigating the longer-term impact was not a part of the evaluation with most of the impacts occurring beyond the scope of the project, which mostly focused on the practitioner level of the theory of change.

Step 4: Could any other contextual factors have influenced the programme and made a significant contribution?

We considered whether the TWiTCH programme was different to what practitioners would be doing normally to assess whether these things would be already happening. While some practitioners felt that they already had knowledge of the TWiTCH strategies and confidence in delivering them after attending training, the evaluation still found that practice and

confidence changed from beginning to the end of the project and TWiTCH seems to be the most likely cause for this. In a case study setting where practitioners were already very knowledgeable, experienced, and effective in supporting children's language development, they felt that the TWiTCH Week 1 activities around storytelling enhanced their current practice, and during the project they changed to use the TWiTCH strategies with all core texts in the nurseries. In the Week 2 activities, practitioners in the same setting were surprised at the levels of reasoning that the children in their setting could achieve through the sessions and planned to incorporate some aspects of these into their future nursery practice.

We also investigated whether practitioners had any other training through the year which could explain the observed changes in knowledge, confidence, and practice. In the post-intervention survey, 22 of 31 practitioners reported that they had not had any other training focused on supporting children's language development during the course of the TWiTCH project. Of the those that had received other training, seven described involvement in training with a communication focus (Elklan or other communication programmes), two practitioners had taken part in NELI training, which supports children's early literacy, and EAL training and WellComm training were mentioned by one practitioner each. As the majority had not engaged in other training it seems likely that TWiTCH would be the primary attribution to the change in practitioner outcomes. However, other communication training may likely have also contributed for those involved.

Looking at children's outcomes, it is important to note that the natural development of children's language will be improving all the time during the year of the programme. This is likely to contribute significantly to gains in children's language and communication during this time. It is difficult in the design of this evaluation to separate gains due to TWiTCH from those due to natural development. Some practitioners in the case studies commented on this and that it was difficult measure whether TWiTCH had impacted on children's development.

Contribution analysis summary

It seems very likely that the TWiTCH activities delivered as part of the programme contributed to and directly led to changes in the outcomes, especially at the level of coaches and practitioners. The evidence is less strong for directly attributing changes in children's outcomes to the TWiTCH programme due to the design of the study as a single group pilot, the data collection of impact on children's outcome being through practitioner report, and the natural development of children over the course of the project. However, the provision of specific examples of where children have directly been using techniques and vocabulary specifically from the TWiTCH session in other practice gives some strength to support the view that the TWiTCH programme is directly contributing to children's language outcomes.

Conclusion

Table 19: Summary of pilot findings

Area of research	Key finding
Acceptability	The TWiTCH programme was generally seen as acceptable to settings. Although coach training and support to settings was delivered inconsistently by different coaches, training and coach support was seen by practitioners as adequate and effective at improving practitioner knowledge and confidence for delivering TWiTCH for the majority of practitioners. Settings were positive about the programme, however, Week 2 activities were often considered to be pitched at too high a level for children, especially at the beginning of the programme.
Fidelity	Almost all settings (92%) were able to implement at least six cycles of the programme during the year, with 38% completing all nine and 29% completing eight—some with time remaining to complete a ninth cycle at the time of data collection. The programme was delivered with a varying degree of fidelity for different elements across settings. Practitioners attended training and engaged with coaching. The group of children who received TWiTCH varied across settings with some settings not delivering TWiTCH with their full cohort of preschool children due to logistical challenges related to space and insufficient staffing numbers to facilitate smaller group work for all children. Children attending nursery part-time were also excluded, sometimes due to not being in nursery for all TWiTCH sessions. In the delivery of TWiTCH, Week 1 sessions of shared book reading were delivered with high fidelity to the programme, however, the Week 2 choice-based language activities were delivered with more variability due to perceived difficulty of the activities and subsequent adaptations made by practitioners. Targeting specific 'at need' children during Week 3 of the cycles was often not done. Practitioners did not always see the value of targeting specific children with TWiTCH activities over their usual practice of targeting 'at need' children more generally.
Evidence of promise	The evaluation showed evidence of promise, with practitioners' quantitative and qualitative data suggesting that the TWiTCH programme contributed to their increased knowledge and confidence in supporting children's language development, and to changes in practice both during storytelling and general practice. Practitioners reported observing improvements in children's communication, language, vocabulary, use of reasoning language, and confidence to contribute in a group setting, which they attributed to the programme. Practitioners felt that taking part in TWiTCH led children to have a better understanding of—and ability to retell—stories.
Readiness for trial	The TWiTCH programme is close to being ready for a larger-scale trial. Readiness for trial is conditional on updating training and materials for Week 2 to ensure practitioners can deliver the content to the target age and ability of children, tailoring training and materials to make sure delivery is suitable in larger group sizes, and refining training for coaches with the aim to achieve greater consistency in delivery of training and coaching across settings. The contribution analysis found that the TWiTCH theory of change was mostly supported and concluded that the expected TWiTCH outcomes occurred and could be attributed to the programme. For a future trial, consideration needs to be made of the target group for the programme given the inconsistency of delivery to the full year group in the pilot.

Interpretation

The evaluation set out to answer research questions investigating the acceptability of the TWiTCH programme, the extent to which the programme was delivered with fidelity, the evidence that it was achieving its intended outcomes, and the feasibility of a future trial. We discuss the results in these four areas below followed by the extent to which the TWiTCH intervention seems ready for trial.

Acceptability of the TWiTCH programme

The TWiTCH programme was seen as acceptable by 24 of the 25 nursery settings recruited to take part in the trial. In all 24 settings, practitioners had delivered at least six cycles of TWiTCH with many completing all nine specified cycles and some moving on to delivery with their own choice of book. These settings had participated in the TWiTCH training, and feedback from surveys and interviews provided evidence that practitioners felt that the training and support was appropriate to enable them to deliver the programme. While the majority of practitioners gained knowledge and confidence as a result of the training, for some aspects of TWiTCH—particularly those relating to dialogic story reading and continuous provision related to stories—some practitioners reported not gaining additional knowledge or confidence due to already being knowledgeable and confident before the TWiTCH programme, and this was higher for practitioners in maintained settings than those in PVI settings.

It was seen as important that training was onsite so that multiple staff were able to attend and the training could be adapted to the context and needs of the setting. In many cases it was not felt to be feasible for the setting to have all practitioners working with three- and four-year olds attend the training due to the logistics and additional cost. Ability to attend the training did not seem to differ between PVI and maintained settings with similar numbers of staff able to attend the training from both.

Almost all staff interviewed were positive about the programme, and it was felt that the programmed aligned well with what staff already did and knowledge they already had; TWiTCH was seen as a reminder of what good practice is and provided a structure to help implementing it. The Week 2 choice-based language activities were less familiar to practitioners and these were seen by staff as being too difficult for children at the beginning of the programme, especially in the way described in the handbook. Practitioners felt that they needed to simplify these activities to deliver them with the younger children. Some also lacked confidence to use the puppet as described to model language in these activities. As children developed, and practitioners delivered more of the programme, the Week 2 activities were seen as more appropriate. The depth of discussion and interaction facilitated by the Week 2 activities sometimes surprised practitioners with what the children were able to do. Staff felt that TWiTCH was suitable for most children in their settings, however occasionally there were children with English as an additional language, children with extremely low language skills, or who had special educational needs where it was seen as less appropriate. In many settings, however, children with SEN or EAL were successfully included in TWiTCH and practitioners reported benefits of the programme specifically for these children. These views were from staff across all types of settings and there did not seem to be different patterns of feedback from PVI and maintained settings.

Programme fidelity

Different aspects of the TWiTCH programme were delivered with varying degrees of fidelity. In terms of the training, while some of the settings were able to allow all practitioners working with three- to -four-year-olds to attend, this was not the case for all settings where fewer practitioners attended due to logistical and cost challenges for the setting. These challenges included practitioners being unable to attend training after hours due to other commitments (such as childcare), settings not being able to release all practitioners during working hours due to needing to maintain ratios and being unable to pay practitioners to work beyond their standard hours due to budget restrictions. In almost all settings, however, the practitioners who delivered TWiTCH during the programme had attended the training sessions and engaged with the mentoring visits and contact throughout the programme. Consideration could be given to whether the programme needs all practitioners working with the age group to attend or whether the champion role could support practitioners unable to attend.

The organisation of TWiTCH in nurseries was not always as the developer team had envisioned in terms of *all* nursery children accessing the TWiTCH sessions in smaller groups (of eight to ten children). Some settings chose to run TWiTCH only with selected groups of children, such as poorer language learners, or to allow children to opt into sessions. Some settings also prioritised the delivery of TWiTCH to those children attending nursery on more days (sometimes excluding part-time children): this was often done for logistical reasons with it being difficult to find space, time, and resources to run multiple small groups for all children to participate. There were settings that implemented TWiTCH with all the children, but this was easier in smaller settings: it posed significant logistical challenges for larger settings especially in terms of using small groups for TWiTCH work. There were a number of settings across the pilot that used larger groups of more than ten children: it was often more difficult to deliver these sessions with fidelity. The organisation of TWiTCH did not seem to vary depending on the type of nursery with a range of practices across both PVI and maintained settings.

The Week 1 sessions, which focused on the reading of the story and exploration of the story using dialogic pedagogy, were delivered with high fidelity across most settings. These were mostly delivered as five separate sessions across the week focusing on the specified areas and were delivered with little adaptation. The structure and content of the Week 1 sessions was the same across all book cycles, which likely facilitated high fidelity. This was also using practice that practitioners felt was familiar and so they were confident to deliver these sessions. Week 2 sessions, which focused on the choice-based language games and use of high-level concepts to facilitate reasoning, were more varied in their delivery: the activities were perceived by practitioners as too difficult for the children in their setting and the resulting less confident delivery meant that there were lots of adaptations made. Not all the recommendations for Week 2 were followed: some sessions were done less or more frequently than specified and the activities were often simplified or sometimes completely changed. In Week 3, the continuous provision activities were done with high fidelity across settings. These activities felt very familiar to practitioners and aligned with their usual practice. However, the targeting of 'at need' children during Week 3 with TWiTCH-focused activities was only implemented in a small number of settings. Often practitioners felt that usual nursery practice in targeting these children was adequate and that the TWiTCH-specific focus was unnecessary.

The TWiTCH champion role was valued and followed by all settings: the champions coordinated TWiTCH delivery with other staff in the setting, liaised with the coach, and often supported other staff in the nursery to develop their practice through provision of training or peer reflection experiences. The champion role seems to be important for the delivery of this programme and has provided some staff the opportunity to develop themselves as they develop other practitioners' skills.

Some nurseries did experience challenges—particularly at the beginning of the programme—with the logistics of delivering the programme to children in small groups. There were often challenges to access quiet space to facilitate the small group work, and the ratios of staff to children in the nursery did not always allow for groups smaller than ten children without having extra staff available. Practitioners also discussed the challenge of delivering TWiTCH with children attending the nursery part-time where these children missed the same sessions each week in the storytelling. This feels a particularly salient issue for the TWiTCH programme to address as disadvantaged children are more likely to be part-time due to their entitlement to only to 15 hours of nursery provision compared to the 30 hours provided to working families. Although expected that maintained nurseries may be better prepared to implement consistent small group work with children attending more regularly, both PVI and maintained nurseries found the logistics of implementation challenging, particularly those with larger numbers of children.

Where practitioners were able to adapt the programme to suit the needs of their children and their nursery context TWiTCH worked particularly well. However, this required confidence and skill from the practitioner and not all were aware that they could adapt the materials or make skilful adaptations to change their practice to work for their setting. Planning TWiTCH sessions in advance and sharing planning collaboratively between staff before session was often able to facilitate this and it may be that further guidance in the handbook and at training could support staff with how to successfully adapt the activities for different children and circumstances.

Evidence of promise

The evaluation showed significant evidence of promise, particularly for practitioners, indicating that the TWiTCH programme contributed to increasing practitioners' knowledge and confidence in how to support children's language development as well as leading to changes in practice both during storytelling and in general practice. Data from multiple sources triangulated to strengthen these findings, including the self-report of practitioners in responding to survey scale items, interviews, and coding of practitioners' audio recordings of storytelling practice. In the audio recordings, baseline to post-intervention changes in practice were specific to TWiTCH techniques adding weight to the evidence that it was the programme that led to these changes.

While practitioners in PVI settings reported lower confidence ratings before TWiTCH than those in maintained settings, they had greater confidence gains afterwards meaning that TWiTCH may have supported narrowing the confidence gap between those in PVI and maintained settings, Practitioners in both types of setting made similar gains in practice after taking part in the programme according to the self-report data. Audio recording data indicated that practitioners in maintained settings made larger gains in their use of dialogic storytelling strategies after doing TWiTCH. This indicates that practitioners in both maintained and PVI settings can benefit from the programme.

Practitioners across all setting types also reported that the programme had a positive impact on children's outcomes. This included improvements in children's communication and language attributable to TWiTCH with some children

improving their vocabulary, improving their ability to use reasoning language, and using TWiTCH -specific language in other areas of nursery. Better understanding of the stories and the ability to recall and retell the stories was also thought to have improved in children. Practitioners also reported that TWiTCH had led to an increase in confidence for children to express themselves as part of a group and there were many specific cases given where it was felt a child who had been unwilling to engage with talking as part of a group at the beginning of the programme was facilitated through TWiTCH to be able to share their ideas in a group setting. However, as this data is not collected directly from the children participating, and it is expected that children's language and other development would improve significantly through the delivery period even without TWiTCH, some caution should be taken in interpreting whether it was TWiTCH that has directly led to the changes.

Practitioners reported that some TWiTCH techniques and activities were not dissimilar to their usual practice. Before TWiTCH, most settings were already delivering daily storytelling sessions of a similar length to TWiTCH, however, this was usually in larger group sizes with more than ten children,[13] particularly in maintained settings where practitioners rarely reported using groups sizes smaller than ten. The audio recordings indicated that practitioners were using fewer open-ended prompts before TWiTCH and there was often minimal discussion of the books alongside the reading of the story. In practitioner interviews and the developer interview, the frequency of *repeated* book reading in TWiTCH was described as higher than in their usual practice even with core texts. This text repetition was seen as particularly valuable for less able children and those with English as an additional language. The familiarity of the programme and its techniques is likely one of the key facilitators of its delivery, increasing its acceptability and making it easier for practitioners to embed across the setting. However, this may also mean that for some practitioners who are already confident and knowledgeable to support children's language development, there may be limited scope for new learning.

Feasibility of trial

This pilot evaluation indicated that the TWiTCH programme and evaluation was appealing for nurseries to take part in and that it was feasible to recruit both maintained and PVI settings from disadvantaged areas to take part in the programme. Both types of settings were able to access and deliver the programme.

The timing for the delivery was feasible for most of the settings indicating that in a trial situation it would be possible to deliver the full programme within the nursery year. However, there was variation in when settings undertook their initial training sessions and for a few settings this was delivered late in the first term leading to less time for delivery of the programme and not leaving enough time for the delivery of nine cycles of TWiTCH during the year. If recruitment of children and baseline data collection was needed as part of a trial, further consideration would need to given to the timeline for the programme and the number of cycles that settings may be able to implement.

The programme used a 'train the trainer' model to allow delivery to a larger number of settings and to facilitate a future scale-up, which would be necessary if the programme were to move to trial. In the pilot, not all coaches received the same TWiTCH training, depending on when they joined the project. Coaches that joined later experienced a better developed and more standardised form of the training, which would be replicable for delivery to more coaches. There was inconsistency in the ways that different coaches delivered the training to settings, particularly in regard to the time they spent doing the training and around the provision of opportunities for practitioners to practise the techniques and to plan for delivery with some of these changes seeming to be linked to whether coaches were employed by the university or as consultants: consultants tended to spend more time training and supporting settings. Having all coaches experiencing the same training may lead to a more consistent approach in the future with consideration and adaptations made to delivery of training based on coach background and experience. We also recommend that the developer team plans for a quality assurance process for delivery of training across coaches that can involve feedback to coaches to support them in delivering the training as described (see Formative Findings below).

Readiness for trial

The success indicators included in the study plan were mostly achieved, indicating fairly good readiness for trial. The contribution analysis concluded that, in most cases, the TWiTCH activities were delivered and led to improvements in the expected areas. The key areas of weakness in terms of readiness for trial related to:

[13] Storytelling may have been done in whole class groups but our survey was not able to distinguish between small groups of more than 10 people and whole class delivery.

- consistency of training delivered by coaches for settings;
- delivery of the programme not being to *all* children in some settings and, at times, to groups of more than eight;
- variation in the delivery of Week 2 activities across settings; and
- delivery of targeted activities for 'at need' children during Week 3 not always being delivered

The developer team has already tried to address most of these areas for the Phase 3 delivery.

Regarding the consistency of training, the developer team has already made changes to the way that coaches are trained for delivery in Phase 3 including training coaches together, updating the training script with greater emphasis on consistency, and providing more opportunities for coaches to practise and experience TWiTCH. The team has also changed the way that training is delivered for settings so that practitioners attend additional online training delivered by the developer before coach training, which will improve the consistency of some areas of the training. This short additional training will be able to be watched later if practitioners are unable to attend live and is intended to give practitioners an overview of the TWiTCH programme covering some of the background information that coaches had consistently omitted in their training. While this is an additional time commitment for settings, the developer team felt that it was important to the programme that settings are provided with this information. The training script for coaches will also now specify the amount of time to be allocated to different parts of the training; in addition, a quality assurance plan for the training is being put into place for Phase 3 delivery. All of these changes are expected to lead to much greater consistency in the delivery of the training and support provided by coaches. The train-the-trainer model seemed to work reasonably effectively during the pilot and these changes will support greater replicability of the programme for delivery at scale in a trial.

In terms of the inconsistency in delivery to different groups of children in the nursery rather than the whole setting, settings need to be provided with greater clarity around what the intervention is and who it is for on signing up to the project—and in their training. Continued variability, as seen in this trial, would make it difficult to look at outcomes for children when different children are being included in different settings. TWiTCH is designed to be an inclusive whole-class intervention. For Phase 3 delivery, the developer team has increased the specified maximum group size to ten. In the pilot programme the group size specification started out at 'five to eight children', which some settings stuck to throughout, however, this was increased during the project to 'up to ten' due to logistical difficulties in some settings with working with smaller groups. This change in group size from the outset is expected to make it easier for settings to deliver this to all their children rather than just a group. However, increased group size may change delivery and make it more difficult to have dialogic interactions with all children. Future research should look at the difference that group size makes to delivery and the effectiveness of the programme.

To increase fidelity across Week 2 activities and deal with practitioners' concerns that the activities were too difficult for younger children, for Phase 3 delivery the number of questions has been reduced, the activity format has been simplified, and additional questions that can be used flexibly by practitioners have been provided in an appendix. Practitioners are also being provided with additional support during training to enable them to be flexible to the needs and interests of the children in their group. The manual has also been adapted to provide practitioners with examples rather than a script encouraging them to use the programme flexibly. This should help support with greater fidelity with the Week 2 delivery and increase readiness for trial.

To encourage greater targeting of children who need additional support during Week 3, the training for settings has been adapted to give greater prominence of this element during discussion of the Week 3 activities, presenting it before the continuous provision activities. With practitioners implementing a TWiTCH cycle in between the two training days, the second training day also now specifically encourages discussion and reflection on how settings targeted children in Week 3 during the implementation of the cycle. Further changes could still be made to the practitioner handbook to emphasise this element of the programme and to provide further guidance on which children to target and examples of activities to do.

For consideration of scale-up beyond funded delivery as part of an evaluation, there is a potential issue with the costing model and how nurseries would access this programme. In interview, senior leaders reported having very minimal training budgets and a willingness only to pay up to a few hundred pounds for the training, support, and resources in the programme. As the costs are significantly higher than this, alternative funding options should be considered to make it accessible to nurseries.

Methodological considerations for a trial

This pilot evaluation found it difficult to collect accurate programme implementation information to monitor fidelity. There is potential in using coach logbooks to monitor this as coaches are visiting each of the nursery settings, discussing with practitioners how the programme is being implemented, and observing practice and completing a logbook entry after each mentoring contact. In the pilot project these logbooks were completed in different ways by each coach and often did not contain specific information on delivery, for example, which children were receiving TWiTCH, size of group, variations in the structure and delivery of sessions. For a trial, coach logbook templates could be adapted and training specifically provided about completing the logbooks, so that practitioners record a greater level of detail about delivery and do this more consistently.

This pilot evaluation found that the audio recording measure, developed as part of this project, captured practitioner story reading practice well and demonstrated change from baseline to post-intervention. A future trial could use this measure to look at change in practitioner practice again. We found that the coding of these audio recordings was particularly resource intensive, taking up a lot of researcher time, so for trial it may be worth using this with a subgroup of practitioners or using only at the endpoint testing, comparing practice between intervention and control groups. If researchers wanted to also code children's interactions with the practitioner during storytelling we would recommend video recording sessions rather than audio only due to the difficulty in distinguishing different children's voices.

This pilot evaluation indicated that, for many settings, the time and space demands of TWiTCH meant changes to usual practice and, in some cases, the loss of other activities. It would be important to systematically consider the changes that were being made to usual practice to implement TWiTCH and the value of the activities that were stopped as part of a future trial.

A small change to the usual practice survey should be made for future trials to allow the distinction between whole-class story reading and larger groups (but not the whole class) story reading; currently, the top category is story reading with 'ten or more children'. Adding in a full-class option in addition to the existing options would allow better understanding of what the usual story reading practice in nursery settings is before TWiTCH and how big a change TWiTCH is to usual practice.

In this study there was a substantial difference between the number of practitioners that completed the baseline survey and those that completed the post-intervention survey. At the beginning of the evaluation, all practitioners who worked with three- and four-year-olds were asked to complete the baseline survey before any TWiTCH training occurred. This resulted in a large number of practitioners completing the survey who did not go on to attend the training or implement TWiTCH. Consideration should be made for a future trial of the timing and design of the baseline survey to minimise this attrition and only target practitioners who would participate in TWiTCH. Completing the survey at the beginning of the first training session would help to minimise this attrition.

One limitation of this pilot evaluation is that it did not investigate children's outcomes. There is indicative evidence, based on practitioners' perceptions, that children's communication improved as part of the programme and that children increased their use of more complex sentence structures and reasoning language. Children were also reported to improve their understanding of stories and their narrative ability to retell stories. For trial, children's outcome measures should include a measure of oral language output, which would ideally capture reasoning language. A specific measure which focuses on children's narrative skills—something expected to improve as a result of TWiTCH—could be effective. Children's outcomes could also include a standardised measure of vocabulary, which would be expected to improve as a result of increased exposure to a wider range of vocabulary through books and specific focus on vocabulary. For the greatest likelihood of seeing change in this area, an assessment focused specifically on the vocabulary contained in the TWiTCH books would give the best chance to see an effect of the programme. As one of the expected impacts is that TWiTCH improves children's EYFS communication scores, we also suggest follow-up data collection of this information at the end of reception year.

Formative findings

The evaluation's findings highlighted some areas where some improvements could be made to the way the intervention is delivered. These areas were discussed with the developer team at two feedback workshops during the delivery period. Some of these recommendations have already been implemented by the developer team for delivery to Phase 3 settings.

To support increased consistency of the training delivered by coaches to nursery settings we have a few suggestions. Training all coaches together—so that they can discuss and develop a shared understanding of the programme together—may help to increase consistency. Coaches also suggested that the opportunity to plan training sessions with other coaches would help them to deliver the training and may also lead to improved consistency. Implementing a quality assurance process, where coaches are observed delivering and provided with the opportunity to discuss feedback, may help also support with consistency and ensure that coaches are delivering the programme as guided. Further guidance on which areas of the training can be adapted to the setting and which need to be delivered as described may also provide further help to coaches when they adapt the training to ensure all the core elements are delivered.[14]

For some settings, delivery was not deemed feasible with smaller groups of children: additional guidance to help such settings deliver to larger groups would be helpful.. It may be useful to provide additional guidance for settings in how to deliver TWiTCH in larger groups to facilitate delivery in these settings. This would be especially useful as the developer team plan to change the guidance for the group work to increase group size maximum to ten children.

TWiTCH worked best when practitioners were able to take the activities but then adapt them to the interests, needs, and contexts of the children in their setting. Explicit permission needs to be given to practitioners in both the training and the handbook so that they feel they are 'allowed to' and are able to make these changes. Further guidance would be helpful on ways that practitioner could adapt the sessions to make the activities easier or more challenging for their group.[15] Practitioners also highlighted that they would appreciate seeing the TWiTCH techniques modelled with different groups of children with greater diversity of ability and background. Videos showing these could be used, or a coach demonstrating TWiTCH at training with children in their setting could enable practitioners to see how TWiTCH can work with children in their context.

Additional guidance should also be provided for practitioners about ways they can include and support part-time children with doing TWiTCH. To support better uptake of the work targeting 'at need' children in Week 3 of the cycle, more support and guidance could be provided to help practitioners identify which children they should work with. Part-time children who missed out on earlier TWiTCH sessions could be specifically mentioned here as children to focus on during Week 3 sessions.

Limitations of the evaluation

This pilot evaluation had a number of limitations. The evaluation focused on the coach and practitioner level of the theory of change, looking at whether the expected outcomes were achieved on these levels. However, the evaluation collected very limited data on the child level, and the data on whether children's outcomes improved as part of TWiTCH is only the self-report of practitioners. Future trials would need to look at whether data collected directly from children aligns with practitioner reports.

As a pilot evaluation, we were working with a small sample size of settings and practitioners. Estimating sample size at the beginning of the project assumed that all practitioners in the nurseries who were working with three- and four-year-old children would participate in TWiTCH and complete both baseline and post-intervention measures. However, in practice, a smaller group of practitioners (around three per setting) completed the training and in many cases, not all of those practitioners were involved in delivering TWiTCH in the setting. This resulted in a much lower response rate for the post-intervention survey. This meant that while the evaluation aimed to look at differences between groups of practitioners in terms of outcome measures, the sample size was too small to run the multiple regression models that would explore this.

The scale measures developed to look at practitioners' confidence, practice, and outcomes were specifically developed to align closely with the TWiTCH programme activities and techniques and were reliant on practitioner self-report. This limits the generalisability of the findings to say how well TWiTCH improved practitioners' overall practice and may be biased by practitioners' views of themselves. Use of a more objective observational measure of setting practice (for

[14] The developer team has now adapted the training programme for delivery in Phase 3 to include multiple plenary training sessions for coaches and has also put together a plan for quality-assuring coach delivery.
[15] TWiTCH training and handbooks have now been updated to support practitioners to adapt the programme to the context of their settings.

example, the Sustained Shared Thinking and Emotional Wellbeing, 'SSTEW', measure), could investigate how much delivering TWiTCH transfers to practitioners' usual practice and would be free from the potential bias of self-report.

The evaluation struggled to accurately track implementation fidelity due to coach logbooks not containing information about delivery and the developer team not having a quality assurance process built in. The developer team also found it difficult to articulate the aspects of TWiTCH that should be there to describe 'high fidelity' implementation versus 'low fidelity' in the early stages of the project, however, this was achieved by the end of the project. A more structured template for coaches to gather delivery and implementation fidelity details in the logbooks during their visit would facilitate the understanding of delivery in each setting for a future trial.

In the coding of the audio recording data, ethical and methodological restrictions meant that we were unable to analyse the child input in the observations of the shared reading sessions. This means that we were unable to capture the types of responses and input the children had in the session, potentially excluding important aspects of the interactions between the practitioner and the children. This issue was due to using opt-out consent for children to be included in the audio recordings and audio recording focused on the practitioner making it difficult to interpret children's voices. Using opt-in consent and video recording may allow the observation and coding of children's contributions as well a practitioner contribution.

Future research and publications

Due to the whole-class, inclusive nature of the TWiTCH programme and the fact that changes to nursery practice due to TWiTCH have been seen beyond the group delivering TWiTCH, we suggest that a future evaluation should use a randomised controlled trial design with randomisation at setting level (excluding multiple settings from a nursery chain or academy). Children's outcomes should be the focus of this trial as these were not captured in this pilot, which only captured perceptions of children's outcomes.

Based on the significant variations seen in delivery during the pilot, we feel that investigating the impact of the group size used for TWiTCH, looking at how practice changes for larger versus smaller groups, and whether this impacts on the effectiveness of the programme, would be important. Investigating whether outcomes change depending on dosage or exposure to the programme would also be interesting, especially considering whether part-time children are benefitting as much as children attending full-time. Specifically looking at the impact of TWiTCH on children with lower levels of language proficiency and those with English as an additional language would also be recommended given the greater potential impact on these group and the benefits noted by some practitioners of TWiTCH for these subgroups. It may be necessary to further adapt the TWiTCH materials to make sure that those with lower language levels are able to access the programme from the start.

Further publications

We expect to write an academic journal paper using the audio recording data collected at baseline in this trial to look at current practice in book-sharing in nurseries over the next year and plan to write a joint journal paper with the developer team including the findings from the pilot evaluation.

References

Ab Wahab, M. K., Zulkifli, H. and Abdul Razak, K. (2022) 'Impact of Philosophy for Children and Its Challenges: A Systematic Review', *Children (Basel)*, 9 (11*)*:1671. DOI: 10.3390/children9111671

Alexander, R. J. (2017) *Towards Dialogic Teaching: Rethinking Classroom Talk (5th edn)*, York: Dialogos.

Barone, C., Chambuleyron, E., Vonnak, R. and Assirelli, G. (2019) 'Home-Based Shared Book Reading Interventions and Children's Language Skills: A Meta-Analysis of Randomised Controlled Trials', *Educational Research and Evaluation*, 25 (5–6), pp. 270–298. DOI: 10.1080/13803611.2020.1814820

Bingham, A. J. (2023) 'From Data Management to Actionable Findings: A Five-Phase Process of Qualitative Data Analysis', *International Journal of Qualitative Methods*, 22. https://doi.org/10.1177/16094069231183620

Brodie, K. (2014) *Sustained Shared Thinking in the Early Years: Linking Theory to Practice*, London: Routledge.

Burgoyne, K., Hargreaves, S., Akhter, N., Cramman, H., Eerola, P., Einbeck, J. and Menzies, V. (2024) 'Lack of Effect of a Parent-Delivered Early Language Intervention: Evidence from a Randomised Controlled Trial Completed During COVID-19', *JCCP Advances*, e12279. https://doi.org/10.1002/jcv2.12279

Cooper, J. O., Heron, T. E. and Heward, W. L. (2014) *Applied Behavior Analysis (2nd edn)*, Pearson Education.

Demissie, F. (2019) '*TWiTCH (Thinking With Tales for Children) Project Report*, (unpublished).

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L. and Cooper, P. J. (2020) 'Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis. *Child Development*, 91, pp. e383–e399. https://doi.org/10.1111/cdev.13225

Eadie, P., Bavin, E., Bretherton, L., Cook, F., Gold, L. and Reilly, S. (2021) 'Predictors in Infancy for Language and Academic Outcomes at 11 Years', *Pediatrics*, 147 (2), e20201712. DOI: 10.1542/peds.2020-1712

EEF (2018) 'Preparing for Literacy: Improving Communication, Language and Literacy in the Early Years', London: Education Endowment Foundation.

Flack, Z., Field, A. and Horst, J. (2018) 'The Effects of Shared Storybook Reading on Word Learning: a Meta-Analysis', *Developmental Psychology*, 54, pp.1334–346. DOI: 10.1037/dev0000512

García-Carrión, R. and Villardón-Gallego, L. (2016) 'Dialogue and Interaction in Early Childhood Education: A Systematic Review', *REMIE – Multidisciplinary Journal of Educational Research*, 6 (1), pp. 51–76. DOI: 10.17583/remie.2016.191

Garcia-Moriyon, F., Rebollo, I. and Colom, R. (2005) 'Evaluating Philosophy for Children: A Meta-Analysis', *Thinking: The journal of Philosophy for Children*, 17 (4), pp. 14–22.

Gorard, S., Siddiqui, N. and See, B. H. (2015) 'Philosophy for Children: Evaluation Report and Executive Summary', London: Education Endowment Foundation.

Hardman, J. (2019) 'Developing and Supporting Implementation of a Dialogic Pedagogy in Primary Schools in England', *Teaching and Teacher Education*, 102908. https://doi.org/10.1016/j.tate.2019.102908

Harrell, F. E. Jr. (2015) *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis (2nd edn)*, Nashville, USA: Springer.

IBM Corp. (2023) *IBM SPSS Statistics for Windows, Version 29,* Armonk, NY: IBM Corp.

Jay, T., Willis, B., Thomas, P., Taylor, R., Moore, N., Burnett, C., Merchant, G. and Stevens, A. (2017) 'Dialogic Teaching: Evaluation Report and Executive Summary', London: Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Dialogic_Teaching_Evaluation_Report.pdf?v=1734707307.

Justice, L., Jiang, H. and Strasser, K. (2018) 'Linguistic Environment of Preschool Classrooms: What Dimensions Support Children's Language Growth? *Early Childhood Research Quarterly*, 42, pp. 79–92. https://doi.org/10.1016/j.ecresq.2017.09.003

Laird-Gentle, A., Larkin, H., Kanasa, H. and Grootenboer, P. (2023) 'Systematic Quantitative Literature Review of the Dialogic Pedagogy Literature', *Australian Journal of Language and Literacy*, 46, pp. 29–51.

Lindorff, A., Sylva, K., Ereky-Stevens, K. and Joseph, A. (2022) 'Coaching Early Conversation, Interaction and Language (CECIL): Impact Evaluation*',* London: The Sutton Trust. https://www.suttontrust.com/wp-content/uploads/2022/03/CECIL-Evaluation-Oxford-14.02.2022.pdf

Lipman M. and Sharp, A. M. (1978) *Growing Up with Philosophy*, Philadelphia, PA: Temple University Press.

Lord, P., Dirie, A., Kettlewell, K. and Styles, B. (2021) 'Evaluation of Philosophy for Children—An Effectiveness Trial: Evaluation Report', London: Education Endowment Foundation.

Lumivero (2024) *NVivo* (Version 15) www.lumivero.com

Mayne, J. (2008) 'Contribution Analysis: An Approach to Exploring Cause and Effect. *ILAC brief*.

Mayne, J. (2012) 'Contribution Analysis: Coming of Age?', *Evaluation*, 18 (3), pp. 270–280.

McKeown, M. G. and Beck, I. L. (2006) 'Encouraging Young Children's Language Interactions with Talk with Tales for Children (TWiTCH) Programme Stories', in D. K. Dickinson and S. B. Neuman (eds.), *Handbook of Early Literacy Research (2nd edn)*, New York: Guilford Press (pp. 281–294).

Meade, A., Williamson, J., Stuart, M., Smorti, S., Robinson, L. and Carroll-Lind, J. (2013) 'Adult-Child Sustained Shared Thinking: Who, How and When?', *Early Education,* 53, pp. 7–12.

Merrell, C., Little, J. and Coe, R. (2014) 'Is the Attainment Gap among Primary Aged Children Decreasing?', in C. Wood and R. Scott (eds), *Harnessing What Works in Eliminating Educational Disadvantage: A Tale of Two Classrooms*, London: Demos (pp. 45–52). https://dera.ioe.ac.uk/23193/1/Two_classrooms_-_web.pdf

Ministry of Housing, Communities and Local Government (2019) 'English Indices of Deprivation 2019'. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

Mol, S. E., Bus, A. G. and de Jong, M. T. (2009) 'Interactive Book Reading in Early Education: A Tool to Stimulate Print Knowledge As Well As Oral Language', *Review of Educational Research*, 79 (2), pp. 979–1007.

OUP (2018) *Why Closing the Word Gap Matters: Oxford Language Report*, Oxford: Oxford University Press.

Piasta, S., Justice, L., Cabell, S., Wiggins, A. and Turnbull, K. (2012) 'Impact of Professional Development on Preschool Teachers' Conversational Responsivity and Children's Linguistic Productivity and Complexity', *Early Childhood Research Quarterly*, 27, pp. 387–400. https://doi.org/10.1016/j.ecresq.2012.01.001

Sammons, P., Elliot, K., Sylva, K., Melhuish, E., Siraj-Blatchford, I. and Taggart, B. (2004) 'The Impact of Pre-School on Young Children's Cognitive Attainments at Entry to Reception', *British Educational Research Journal*, 30, pp. 691–712. https://doi.org/10.1080/0141192042000234656

Sammons, P., Gu, Q., Day, C. and Ko, J. (2011) 'Exploring the Impact of School Leadership on Pupil Outcomes: Results from a Study of Academically Improved and Effective Schools in England', *International Journal of Educational Management*, 25, pp. 83–101.

SAPERE (n.d) 'What is P4C'. sapere.org.uk/what-is-p4c/

Siraj-Blatchford, I. and Manni, L. (2009) '"Would you like to tidy up now?": An Analysis of Adult Questioning in the English Foundation Stage', *Early Years*, 28, pp. 5–22.

Siraj, I., Kingston, D. and Melhuish, E. (2015) *Assessing Quality in Early Childhood Education and Care: Sustained Shared Thinking and Emotional Well-Being (SSTEW) Scale for 2-5 Year Olds Provision*, London: Trentham.

Sénéchal, M. and LeFevre, J-A. (2002) 'Parental Involvement in the Development of Children's Reading Skill: A Five-Year Longitudinal Study', *Child Development*, 73 (2), pp. 445–460.

Sénéchal, M., Ouellette, G. and Rodney, D. (2006) 'The Misunderstood Giant: On the Predictive Role of Vocabulary to Reading', in S. B. Neuman and D. Dickinson (eds), *Handbook of Early Literacy (2nd edn)*, New York: Guilford Press (pp. 173–182).

Snow, C. E., Burns, M. S. and Griffin, P. (1998) *Preventing Reading Difficulties in Young Children*, Washington, DC: National Academy Press.

Stanley, S. (2012) *Why think?: Philosophical Play From 3–11*, Bloomsbury Publishing.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (2004) 'The Effective Provision of Pre-School Education (EPPE) Project Technical Paper 12: The Final Report-Effective Pre-School Education', London: IOE.

Sylva, K., Melhuish, E., Sammons, P., Siraj, I., Taggert, B., Smees, R. and Hollingworth, K. (2014) 'Students' Educational and Developmental Outcomes at Age 16: Effective Pre-School, Primary and Secondary Education (EPPSE 3-16) Project' (research brief, September 2014), London: Department for Education.

Taggart, B., Sammons, P., Smees, R., Sylva, K., Melhuish, E., Siraj-Blatchford, I. and Lunt, I. (2006) 'Early Identification of Special Educational Needs and the Definition of "At Risk": The Early Years Transition and Special Educational Needs (EYTSEN) Project', *British Journal of Special Education*, 33 (1), pp. 40–45.

TASO (2022) 'Impact Evaluation with Small Cohorts: Methodology Guidance', London: TASO, London. https://cdn.taso.org.uk/wp-content/uploads/2022-06-17_Impact-evaluation-with-small-cohorts-methodology-guidance_TASO.pdf

Tower Hamlets (2009) 'Progression in Language Structures', Tower Hamlets EMA team in collaboration with Tower Hamlets teachers. https://www.themeadtrust.org/wp-content/uploads/2019/09/Tower-Hamlets-Language-Structures-2.pdf

Tymms, P., Merrell, C., Hawker, D. and Nicholson, F. (2014) 'Performance Indicators in Primary Schools: A Comparison of Performance on Entry to School and the Progress Made in the First Year in England and Four Other Jurisdictions', London: Department for Education. https://www.gov.uk/government/publications/performance-indicators-in-primary-schools

van der Wilt, F., Bouwer, R. and van der Veen, C. (2022) 'Dialogic Classroom Talk in Early Childhood Education: The Effect on Language Skills and Social Competence', *Learning and Instruction*, 77, 101522. https://doi.org/10.1016/j.learninstruc.2021.101522

Weadman, T., Serry, T. and Snow, P. C. (2022) 'The Development and Psychometric Properties of a Shared Book Reading Observational Tool: The Emergent Literacy and Language Early Childhood Checklist for Teachers (ELLECCT)', *First Language*, advance online publication: https://doi.org/10.1177/01427237211056735

Further appendices

Please find the further appendices in an accompanying document on the project page.

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

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

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn