



Tales Toolkit

Pilot Report

July 2025

Ceri Williams, Clare Huxley, Olivia Garner, Louisa Illidge,
Jade Talbot, Becci Newton



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 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

Executive summary **4**

Introduction **6**

Methods **15**

Findings **28**

Conclusion **68**

References **78**

About the evaluator

The project was independently evaluated by a team from the Institute for Employment Studies (the IES): Becci Newton, Ceri Williams, Clare Huxley, Olivia Garner, Louisa Illidge, and Jade Talbot worked across the research activities.

The lead evaluator was Becci Newton and Ceri Williams was the project manager.

Contact details

Ceri Williams
Institute for Employment Studies (IES),
City Gate,
185 Dyke Road,
Brighton, BN3 1TL

Tel: +44 (0) 1273 763 400

Email: ceri.williams@employment-studies.co.uk

Acknowledgements

The authors are indebted to all the staff at the early years settings who took part in the research for their time, energy, and commitment to the project. We would particularly like to thank all the staff who took part in the case-study visits and research surveys at an extremely difficult and busy time.

Our sincere thanks also go to the delivery team at Tales Toolkit and Goldsmiths, University of London: Kate Shelley, Dr Alice Jones Bartoli, Rhys Bevan, Georgie Matthews, and Kara Leatherbarrow for supporting us throughout the research activities. Special thanks to Gemma Tappenden, project manager at the Thames Valley Stronger Practice Hub for all her insights into the sector and support encouraging settings to take part in the research activities. Thanks also to all the project support officers who assisted us over the course of the project including Ehecatl Hunt-Duarte, Sara Butcher, and Zoe Gallagher. Finally, thank you to the Education Endowment Foundation (EEF) for funding and supporting this research. Our particular thanks to Aoife Duff, Daniela Alvarado, and Guillermo Romero at the EEF for their support and guidance.

Executive summary

The project

Tales Toolkit is a package of training and resources to support early language and social development through child-led oral storytelling activities. The language and narrative skills children develop are expected to support their writing development and their confidence, creative thinking, problem-solving, and self-regulation skills. The programme is designed for practitioners working with children from ages two to five, with children aged three to five (preschool or reception) being the focus of this pilot.

Settings receive access to five 50-minute online training videos and additional online resources and are expected to provide the training sessions to all staff working with children aged three to five over a three- to five-month period. Staff undertake the training together in groups, with accompanying group activities. Tales Toolkit sessions were expected to be delivered to all children at least once per week during term time or ongoing provision at settings that operated throughout the year. During sessions, practitioners guide children through the process of creating their own story, helping them to identify ideas for four key parts of their story—character, setting, problem, and solution—and also to record stories using pictures, mark-making, or writing. Settings were also expected to make Tales Toolkit resources available for children to access outside of the sessions.

The evaluation aimed to examine whether Tales Toolkit shows evidence of promise (supports the outcomes for practitioners and children identified in the theory of change), is feasible and acceptable for early years staff to deliver and could be scaled and tested as part of an efficacy trial. A multiphase, mixed methods evaluation design was used to examine the research questions. It included a review of programme materials, surveys of practitioners pre- and post-delivery, observations of programme delivery, interviews with a selection of practitioners and setting managers, and interviews with the delivery team at the end of delivery.

Programme delivery and the evaluation activities occurred from September 2023 to July 2024.¹ As part of the [Department for Education's Early Years Recovery Programme](#), the EEF has worked 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 aimed to support education recovery following the pandemic while also developing our understanding of effective professional development in the early years sector. The EEF worked with Thames Valley Early Years Stronger Practice Hub to fund settings' access to Tales Toolkit and evaluate the programme through this pilot study. Kate Shelley of Tales Toolkit Ltd originally developed Tales Toolkit; she has since collaborated with Dr Alice Jones Bartoli, director of the School and Family Studies at Goldsmiths University of London, to test and refine the model further.

Table 1: Summary of pilot findings

Area of research	Finding
Evidence of promise	<p>Most practitioners who responded to the endline survey self-reported being highly motivated to deliver the intervention and reported increased confidence and changes in practice regarding supporting children's language, quality interactions, collaborative play, storytelling, and child-led learning.</p> <p>Practitioners also reported perceived positive outcomes for children in CLL and PSED, including creativity, storytelling, writing, vocabulary, speech, problem-solving skills, turn-taking, and sharing. Some practitioners reported especially positive outcomes for children with SEND and EAL.</p>

¹ Some settings reported that staffing issues associated with the expansion of funded childcare impacted on engagement with the pilot. The policy was announced in April 2023, after recruitment to the pilot had started, so staff would not necessarily have considered the potential impact of this policy when signing up. The expansion of provision began in April 2024, part-way through the pilot.

<p>Feasibility of implementation</p>	<p>Staff who responded to the endline survey reported that the intervention delivery was straightforward and did not require much time to plan or deliver. Fifty-eight percent of practitioners who completed the survey reported that they personally had delivered at least once per week, and all said that practitioners and children used the language of problem and solution outside of directed sessions. Half of the settings visited and 89% of staff who responded to the endline survey reported that resources were also available to children outside of the directed sessions.</p> <p>Ten of the 30 settings that took part withdrew from programme delivery over the course of the pilot, mainly due to staffing issues. There were difficulties with training completion in some settings, which staff also explained were due to problems with staffing and time pressures. Fifty percent of staff responding to the endline survey had completed all five training sessions and 64% had completed at least three. Staffing issues were particularly acute for PVI settings and seven of the ten settings that withdrew during the pilot were PVIs that cited staffing issues.</p>
<p>Readiness for trial</p>	<p>Evidence suggests that the training and delivery materials were high quality, well-structured, easy to use, and feasible to implement. However, there were issues around training completion for staff delivering Tales Toolkit and variations in dosage for recipient children that will need to be addressed for a future trial. The delivery team could try a different strategy to ensure a clear understanding of the core and optional elements of the training, add additional monitoring and support to ensure engagement with training and delivery, and consider changing the mode of training delivery to live webinars or in-person sessions, although this would have cost and resource implications for scale-up.</p> <p>Practitioners responding to the endline survey reported perceived positive outcomes in the CLL and PSED areas of the EYFS in which there are a range of validated assessments that are standardised with early years children in the U.K. Engaging practitioners in all types of settings in research activities is feasible. However, high attrition of PVI settings in the pilot means any future trial would need to consider how to more effectively engage PVIs. A larger sample size would be required to account for potential attrition. Careful research design can reduce the burden on settings, and a number of EEF efficacy trials have been delivered with early years children in pre-school and reception.</p>

Additional findings

A benefit of Tales Toolkit is its flexibility and adaptability for use with different children. Practitioners commented that the training was insightful. They suggested the training could be further enhanced by the inclusion of more examples of Tales Toolkit being delivered with children with differing needs and ages and in different size groups. Children in early years settings can have varying patterns of attendance. In the pilot, this meant that those children attending the setting for longer hours could potentially participate in multiple sessions across the week. For a future trial, the delivery team may need to specify a maximum number of directed sessions per week to ensure all children receive the same ‘dose’ of the intervention.

Despite Tales Toolkit being a relatively low-cost intervention, there were mixed findings regarding its perceived affordability. While some staff surveyed in the pilot considered it affordable, some setting managers indicated they would not have been able to afford it in the current economic climate characterised by tight budgets and staffing issues. Nonetheless, most managers interviewed also commented that, in light of the benefits they perceived for children who had taken part, it would be worth the investment.

One limitation of this pilot was the low involvement of settings in deprived areas (based on the Income Deprivation Affecting Children Index), which limited the ability to explore perceived outcomes for disadvantaged children. Another limitation was the low response rate to the endline survey (which received 56 responses from 21 settings compared to 156 responses from 30 settings for the baseline survey). This reduces the strength of the findings and the ability to understand the experiences of less engaged settings.

Introduction

Background evidence

Research shows that the attainment gap between disadvantaged pupils and their peers starts in the early years, grows over time, and has long-lasting impacts (Tucket et al., 2024; EEF, 2018b; de Sylva et al., 2010). Although the attainment gap was evident in the data before 2020 (Hutchinson et al., 2020), the most recent data shows a widening of the gap following COVID-19 due to the educational impact of the pandemic and school closures (Tucket et al., 2021; González and Bonal, 2021). The Early Years Foundation Stage Profile (EYFSP) results show that the gap between the proportion of pupils eligible for means-tested free school meals achieving a 'good level of development' and their peers grew from 19.9 percentage points in 2022/2023 to 20.5 in 2023/2024. The most recent data also shows a widening gap between boys and girls (14.3 percentage points, up from 13.2 in 2021/2022), with a higher proportion of girls identified as having a 'good level of development' compared with boys while the percentage of children assessed with any Special Educational Needs having a good level of development remained at just under one-fifth (19.7%) (DfE, 2024; Cumisky, 2024).

The development of language skills in the early years is an important determinant of outcomes later in life, including academic achievement, employment, and mental health (Law et al., 2017). Early language skills provide a foundation for learning. One in four children who struggle with language at age five do not reach the expected standard in English at the end of primary school, and children with a poor vocabulary at age five are more than twice as likely to be unemployed at age 34 (Kerr and Franklin, 2021). Early intervention has enormous potential for eliminating the attainment gap in the early years (Stewart and Waldfogel, 2017) and the evidence indicates that strengthening language and communication skills is crucial to this.

Guidance from the Department for Education in England emphasises the importance of developing children's early language skills:

'High-quality early years education, with a strong focus on communication, is good for every child. It is especially positive for disadvantaged children. Children who have lived through difficult experiences can grow stronger when they experience high-quality early education and care' (DfE, 2023).

As such, the EEF launched a funding round focused on early language through which it commissioned evaluations of a number of interventions focused on improving children's language and communication skills (EEF, 2023). Play-based learning is integral to Tales Toolkit and this pilot has the potential to contribute to understanding this approach and how it can be applied in early years settings. While play-based learning is widely acknowledged as important, as highlighted by the work of Bruner (2020) and Vygotsky (1967), the evidence base is limited (EEF, 2021b). Approaches such as Tales Toolkit, which include practitioners responsively interacting with children through storytelling-based oral language approaches, have shown promise in supporting children's language development (Hussain et al., 2021). The literature shows the importance of professional development to enable practitioners to learn how to apply and embed high-quality interactions in their early years practice (Husain et al., 2020). Approaches like Tales Toolkit, which involve practitioners engaging with children using responsive oral language prompts (that follow what the child is focused on or respond to their verbal or non-verbal communication), have demonstrated potential in enhancing children's language skills. A quasi-experimental study (QED) including approximately 500 two- to five-year-olds in ten settings found that those receiving Tales Toolkit made more progress than a business-as-usual comparison group of pupils in all seven areas of the EYFSP (Jones Bartoli, 2018). The overall effect size was 0.16 (equivalent to two month's additional progress). This was a small-scale QED with some evidence of imbalance between the two groups at baseline but this was accounted for in the analysis. The study provided some initial evidence of promise that warranted further research to explore which outcomes, such as language and communication, have particular evidence of promise and could be assessed using validated standardised measures in a future efficacy trial.

As part of the Department for Education's (DfE) Early Years Recovery Programme, the Stronger Practice Hubs (SPH) and the EEF worked together to fund early years settings' access to evidence-informed programmes and study the influence on practice and children's outcomes. The aims were to support education recovery following the pandemic and to develop an understanding of effective professional development in the early years. Tales Toolkit was one of the interventions funded through the SPH. It aimed to close the attainment gap for disadvantaged children through

improvements in language, communication and literacy. This pilot evaluation was designed to help Tales Toolkit Ltd refine its intervention and allow the EEF to assess the possible benefits for practitioners and children and whether it could be scaled up to be provided to more settings and children. The research focused on how Tales Toolkit was delivered in practice, the feasibility of delivering the intervention, early years practitioners' experiences of using it, practitioner and setting support needs, and any perceived initial benefits for staff practice and children. Additionally, the pilot evaluation was anticipated to enhance the existing evidence-base concerning play-based learning.

Intervention

This section provides a detailed description of the intervention, with its components, described in line with the TIDIER checklist.

Intervention name

Tales Toolkit.

Why—theory and rationale

Tales Toolkit is a package of training and resources supporting early language and social development through child-led oral storytelling activities. Relatively little planning is required for the sessions, which are structured around practitioners guiding children through the process of creating their own stories. The emphasis is on the power of the training to prepare staff to deliver the programme. Children are supported to identify ideas for four key parts of their story: character, setting, problem, and solution. Over time, practitioners support children to become more independent in creating their stories and, in later stages of the intervention, children are supported to write or draw parts of their story. Through storytelling, staff have opportunities to responsively interact with children, developing their language and communication skills. The activities are also expected to develop children's confidence, creative thinking, problem-solving, and self-regulation skills. The language and narrative skills children develop are also expected to support children's writing development.

Who—recipients

For this pilot, the intervention was aimed at all children aged three to five (although Tales Toolkit Ltd designed the programme to be used by practitioners working with children from aged two to five). Tales Toolkit activities can be delivered with whole classes or groups, smaller groups, or individuals. Settings were advised to provide all children with an opportunity to take part in a session at least once a week.

Early years staff from 30 settings from the private, voluntary, and independent sector (PVI) or maintained nursery schools (either standalone or within an infant or primary school) took part in the pilot. Settings in the following local authorities in the Thames Valley Early Years SPH's region were invited to take part in the research:

- Bracknell Forest;
- Oxfordshire;
- Reading;
- Slough;
- West Berkshire;
- Windsor and Maidenhead; and
- Wokingham.

Settings that were maintained nurseries within an infant or primary school were also encouraged to deliver the programme to their reception class(es) as well as nursery class(es). When settings signed up to take part in the evaluation, nine nursery settings within an infant or primary school indicated that they could also deliver to reception age children.²

² For ease of reporting, 'school-based' is also used to indicate settings that were maintained within an infant or primary school.

What—materials

Each setting received access to:³

- five 50-minute online training videos along with clear instructions on the recommended schedule for viewing them; these included interviews with experts, footage of the physical resources in use, and timers for activities and discussion;
- an optional package of online support and resources including examples of delivery from other settings, relevant research, downloadable images for stories, webinars with experts in the early years (including Dr Julian Grenier), and access to a Facebook group where staff could access peer support ('Talking Tales Toolkit');
- login credentials to access training videos and online resources;⁴ and
- a kit of physical resources to use in the delivery of Tales Toolkit sessions—big bags, small kit, apron, hanger, shuffle book, and big book.

What—procedures

The instructions for the training that were provided to the settings by the delivery team⁵ stated that:

- the training sessions were expected to take place over a three- to four-month period, with flexibility to accommodate the schedule to the setting's needs and preferences;
- the first and second training sessions could be delivered either a week apart or together in the same session;
- the remaining three training sessions could be completed at least one month after the preceding session; and
- staff working with the specified age group had to complete *all five* training sessions.

Also set out in the information the delivery team sent to settings was that staff would attend the training together as it was expected that they would work through the training sessions as a group, in order to facilitate discussion, collaboration, and the completion of a short exercise collectively. Settings are asked to complete the training together, either all staff together or, where that is not possible, in smaller groups, to ensure the interactive elements were fully utilised.⁶ Settings were also asked to identify a lead practitioner who would be the main point of contact. Their role was to organise the group training sessions (gathering the necessary resources, distributing handouts, facilitating the discussions, and completing the short exercise) and plan for delivery. These lead practitioners (facilitators) were expected to watch each video in advance of the training session. This role was promoted as an opportunity for the leading staff member to monitor the level of staff understanding of their group and identify areas for support.

It was recommended that training sessions should be watched in their entirety the first time; after that, staff were welcome to review particular segments of training sessions as needed. Sessions were typically expected to run during staff meetings or in the form of a dedicated training session, for example, as part of an in-service education and training (INSET) day and cover the following topics:

1. 'Welcome and quality interactions'—the process of developing quality interactions through child-led activities.

³ The evaluation team was given access to all of the training videos and online resources but did not receive the kit of physical resources.

⁴ The option was for all staff members to have a log-in, with the lead person having responsibility to share the link to set this up via email.

⁵ In two pre-training documents: pre-training information and story resources for training.

⁶ In these circumstances, the lead practitioner was expected to check all relevant staff had completed the training.

2. 'Big Kits'—the principles of Tales Toolkit, its theoretical base, and how activities can meet its key aims of language, social and emotional, and creativity development.
3. 'All the TTK resources'—developing independent storytellers, evaluating your class-space, and using resources to their full potential.
4. 'Muscles and motor skills'—mark-making and writing.
5. 'Best practice'—reflection on conveying impact to stakeholders, considering next steps, and celebrating successes.

In Tales Toolkit sessions delivered to children, practitioners guided all children through the process of creating their own story and supporting them in identifying ideas for its four key parts (character, setting, problem, and solution). Sessions varied, depending on a number of factors such as the age of the children, how many were in the session, and whether any had English as an additional language (EAL) or special educational needs and disabilities (SEND). A typical session for three-year-olds would start with the practitioner using each of the relevant 'big bags' to let participating children select objects. Firstly, using the character bag to select a character—for example, a princess—and then asking the children to describe what they could see. The practitioner would then use the 'setting bag' to let children select an object such as a castle and ask the group to talk about this. The next step would be to select an object from the next bag, representing a narrative problem (for example, a drawbridge and moat) and ask them what they think a problem might be for the princess in the castle with a moat. The final bag would be the 'solution bag' and children selecting the object here would be shown or asked to discuss how the problem might be solved—for example, a boat for the princess to escape in across the moat. The practitioner might then lead a group writing session—as a continuation of the oral session, as an add-on, or at a different time—where they write up what they have been talking about on the board, showing the four elements of the story. Depending on their age and ability, the children tell the story orally or draw, make marks, or write the story using the varied resources supplied.

The training also set out an expectation that settings would make the Tales Toolkit physical resources available for the children to use when they chose.

Guidance in the frequently asked questions section of the Tales Toolkit website stated that it can be used alongside other interventions, such as Helicopter Stories and Talk for Writing (Pie Corbett), and that using different methods can be complementary.

Who—implementers

Tales Toolkit training was provided by Tales Toolkit Ltd. Early years staff in settings who had participated in Tales Toolkit training delivered the intervention.

How—mode of delivery

Practitioners were expected to deliver Tales Toolkit sessions at least once a week to all children in the room/class. Each session required a space for all children to sit together, either inside or outside, and the Tales Toolkit kits, each containing practitioner-chosen prompts. Practitioners asked children to reveal the prompts from the chosen kit in turn and together they create a story based on the 'character', 'setting', 'problem', and 'solution' prompts. Through completion of all the tasks during the training session, practitioners provided support for children to tell stories independently and enabled children to independently access Tales Toolkit resources along with props they could use to create stories.

Where

Participating children received the intervention from their early years setting practitioners during their usual attendance at their early years settings.

When and how much

The intervention happened during term time or during ongoing provision at settings that operated throughout the year and consisted of at least one weekly session of Tales Toolkit storytelling for all children in the room/class alongside opportunities for children to access Tales Toolkit independently in the provision. Although there was no specified duration

for the sessions they were generally up to around 30 minutes. With younger children, the sessions could be shorter, depending on their age and when on the day they were being run.

Tailoring and adaptation

No tailoring or adaptation was planned. However, Tales Toolkit was set up so that settings had the flexibility to adapt the programme to their children (and the wider community). For example, a setting could choose to lead groups that were differentiated according to need and so could scaffold the storytelling to the group.

How well—planned

Additional support included research articles linked to each session, a social media group for participants, an archive of webinars with early years experts, and downloadable resources. The delivery team offered online drop-in sessions and settings could contact the delivery team directly for support via email or an online call.

Theory of change

The theory of change (ToC) was initially developed with facilitation from the EEF in a theory of change workshop with the delivery team. This was subsequently adapted during an ‘implementation, delivery, and evaluation analysis’ (IDEA) workshop (Humphrey et al., 2016) with the EEF, the evaluation team, and the delivery team at set-up stage. As shown in Figure 1, the ToC had a target population of children between three and five years, in particular, those who were at risk of poorer outcomes for personal, social, and emotional development (PSED) and communication, language and literacy (CLL).⁷ This included children receiving the Early Years Pupil Premium (EYPP) and those with SEND or EAL. The ToC describes the intervention inputs and outputs as well as the anticipated short-, intermediate-, and longer-term outcomes (see Appendix A for the assumptions and causal mechanisms).

The ToC suggests that the intervention changes and improves practitioner behaviour around interactions and storytelling and that because of these changes, children will have improved attainment in communication and language.

For practitioners, the inputs section specifies the training and resources they will receive and the outputs show how they go on to deliver the intervention to children in their setting.

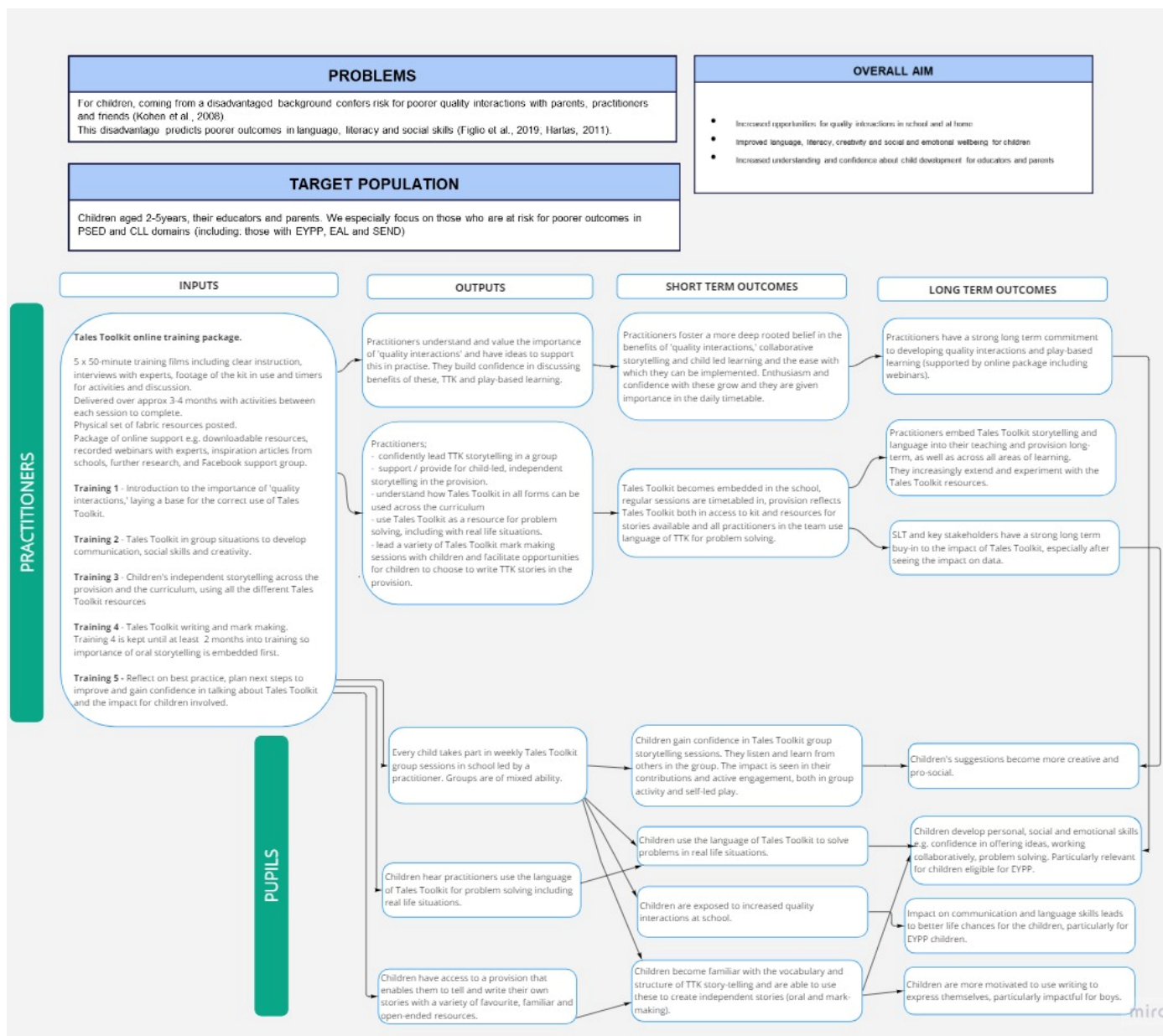
The outcomes for practitioners in the short term are expected to lead to a change in the way they engage with children due to them having a greater understanding of quality interactions and how to support these, modelling language and interaction skills, and using strategies to improve language. A further short-term outcome specified in the ToC is that practitioners embed the Tales Toolkit sessions in the school through regular timetabled sessions and that provision reflects the intervention both in terms of access to the physical resources and practitioners using the language of Tales Toolkit for problem solving. Longer term outcomes for practitioners comprise a greater understanding of the importance of language and communication and quality interactions and play-based learning.

For children, the outputs are taking part and enjoying the weekly Tales Toolkit sessions, being exposed to story-telling opportunities, and experiencing the foundations of character, setting, problem, and solution and associated activities in the setting. This allows children to experience models of communication and story-writing with positive feedback, and to practice these new skills.

The short-term outcomes for the children are improved confidence in storytelling, vocabulary, and language and improved social interaction and behaviours around solving problems in real life. Longer term outcomes include accelerated progress and attainment in language and communication, improved personal, social, and emotional skills, and increased creativity.

⁷ The Early Years Foundation Stage statutory framework was revised in 2023 and 2024. Prior to 2023, communication, language and literacy were combined (denoted by CLL). The trial was designed using CLL, hence this is still used in the report. The current areas of learning and development are split into three prime areas of (a) communication and language, (b) personal, social, and emotional development, and (c) physical development and four specific areas of literacy, mathematics, understanding the world, and expressive art and design, which help strengthen and develop the prime areas.

Figure 1: Final Tales Toolkit theory of change model



Research questions

The evaluation aimed to examine whether the pilot programme showed evidence of promise, was feasible for settings to deliver, and could be replicated, scaled, and tested as part of an efficacy trial. It therefore aimed to answer the following detailed research questions. These were used to inform the collection of evidence to ensure the evaluation could report on whether these conditions were met.

Is there evidence to support the theory of change?

RQ1 Does the programme lead to a change in practice (as per the short-term outcomes in the theory of change)?

RQ1.1 Is there evidence of collaborative storytelling, quality interactions, and child-led learning?

RQ1.2 Is there evidence of practitioners using the resources in their classrooms or other learning environments (for example, in an outside play area, other shared areas in the setting)? How do they use them, and does this vary by type of setting (for example, PVI versus maintained settings)?

RQ2 Do practitioners perceive the intervention has led (or will lead) to positive outcomes for children (for example, their language and communication skills, PSED, creativity, and writing), especially for children in receipt of EYPP or with EAL and SEND?

RQ3 Do practitioners have confidence and motivation to deliver the intervention?

RQ4 Are there any unintended consequences or negative effects—for example, pushing out other programmes, activities widening the attainment gap between disadvantaged and non-disadvantaged pupils, or a perceived burden?

RQ5 How do pupils engage with the different activities and aspects of the programme?

Is the approach feasible to implement and deliver?

RQ6 Is the intervention feasible for practitioners to engage in and implement?

RQ7 Can settings deliver with fidelity and within routine practice? Is there variation in delivery across settings—for example, is the intervention embedded into the curriculum, planning, practice, resources, and language or in separate sessions and activities?

RQ8 Does the delivery method of the training (context and mode) have an impact on engagement with the programme (practitioners)?

RQ9 What are the potential barriers and facilitators of the programme?

RQ10 What support is needed and how can a useful support model be developed?

Is the programme ready to be evaluated in a trial—and if so, what would be needed?

RQ11 Do the training and resources constitute a standard intervention that can be replicated and scaled up?

RQ12 What, if anything, should be updated or changed for a trial evaluation (for example, ToC, materials, delivery method)?

RQ13 Are there any key contextual factors that appear to facilitate or impede successful implementation at a larger scale, for example, setting types?

RQ14 What will the cost of the intervention be to settings? Are the costs considered affordable? What is the willingness of settings to pay for the programme and commit the time, beyond the pilot? How does the cost and affordability vary by type of setting?

RQ15 Are changes in key outcomes (for example, communication and language) measurable?

Is the evaluation feasible to implement?

RQ16 Is it feasible to involve practitioners in research activities? In particular, can their support be secured for the collection of primary outcome data and for depth interviews (for example, due to time-burden)?

RQ17 Would it be feasible to collect child outcome data? What outcome measure(s) might be used?

Ethical review

The IES applied for ethics review through the internal IES Research Ethics Committee in early 2023. The committee comprises two senior staff within the IES and the chair of the ethics board—a member of the Institute Management Team. Depending on the nature of the research and the perceived level of risk, projects undergo either an expedited review of the submitted approach, research information, and research tools (with scrutiny by the committee Chair) or a full review by the sitting Ethics Committee. As the study did not include direct data collection from children and research

materials for the adults taking part were well-drafted, accessible, and comprehensive in addressing the various dimensions of good ethical research practice, a full review by the IES ethics committee was not deemed necessary. The study received approval on 3 March 2023. Settings were provided with a memorandum of understanding information document (see Appendix B) explaining the intervention, the study, what would be involved for the setting, how their data would be used, stored, and protected, and the responsibilities of the evaluators, the delivery team, and participating settings. The document contained a link to an online memorandum of understanding setting agreement form hosted and managed by the delivery team, which settings could complete where they wanted to take part in the study.

Participating settings shared an information sheet supplied by the evaluation team with the parents and carers of participating children. This explained the study and provided contact details for the evaluation team so parents and carers could get in touch if they had any questions or concerns (see Appendix C).

Settings that opted into the qualitative research and staff at those settings who were invited to take part in evaluation activities were provided with an information sheet and a privacy notice. The staff information sheet explained the research, what it would involve, and provided contact details for the evaluation team (see Appendix D). The privacy notice explained the legal basis for processing personal data, set out how their data would be used, stored, and protected, and communicated to participants their right to withdraw from data processing (see Appendix E). At the start of interviews, interviewees verbally consented to indicate that they understood the research aims, agreed to the interview being recorded, and were given assurance of anonymity. Setting staff had the chance to opt out of taking part in the study when they received the invitation to respond to the survey or arrange an observation.

Data protection

This research adhered to the General Data Protection Regulation (GDPR) and the Data Protection Act 2018. The IES's legal basis for processing personal data was 'legitimate interests'. The evaluation of Tales Toolkit fulfils one of the IES's core business purposes (undertaking research, evaluation, and information activities) and was, therefore, in its legitimate interest: that processing personal information is necessary for the conduct of the evaluation. For the purposes of conducting the evaluation to assess evidence of promise for Tales Toolkit, the IES and Tales Toolkit Ltd (the delivery team) were each data controller for personal data of school staff collected during the evaluation process. Personal data was shared with trusted processors, such as transcribers and members of the delivery and evaluation teams, solely for the purposes of proper delivery, management, and evaluation of the project. The IES securely deletes all personal data within six months of any project finishing, that is, once the final draft of the evaluation report has been submitted.

As noted, a privacy notice was made available to settings and staff (Appendix E). This explained in detail what data would be collected and why as well as how data would be stored, transferred, and deleted. It also provided information to participants regarding their right to withdraw from data processing. This was available online, with the link provided in the MoU information document, and attached with email briefings to take part in the surveys and interviews. Contact details were also provided in case of queries or complaints. A data-sharing agreement was signed between the IES and Tales Toolkit stating data to be shared, by whom, how, and why, to ensure full data security, for example, sharing name, job role, and email address for key staff at settings.

In addition, information sheets for staff provided an overview of the research, which included basic information about what data would be collected, how it would be stored and processed, and contact details for further queries (Appendix D). An information sheet for parents gave an overview of the research and provided reassurances that children's personal data would not be collected as part of the study (Appendix C).

Project team

IES evaluation team

Becci Newton	Principal investigator, responsible for leading the evaluation.
Clare Huxley	Project director: supported the PI with leading the evaluation.
Ceri Williams	Project manager: responsible for managing all research activity, liaising with the EEF and Tales Toolkit, and coordinating members of the research team.
Olivia Garner	Deputy project manager: responsible for drafting research tools, supporting the coordination of the research team, qualitative research and analysis.
Jade Talbot	Research fellow: responsible for quantitative and qualitative research and analysis.
Louisa Illidge	Research officer: responsible for qualitative research and analysis.

Tales Toolkit implementation team

Kate Shelley	Project lead (Tales Toolkit).
Dr Alice Jones Bartoli	Project lead (Goldsmiths, University of London).
Rhys Bevan	Project manager (Tales Toolkit).
Kara Leatherbarrow	Project support officer (Goldsmiths, University of London).
Georgie Matthews	Project support officer (Tales Toolkit).

EEF team

Daniela Alvarado	Evaluation manager (2024–2025).
Aoife Duff	Senior programme manager (2022–2025).
Guillermo Romero	Evaluation manager (2022–2023).

Methods

Recruitment

Settings were recruited to the pilot through Tales Toolkit Ltd (the delivery team). The pilot sample did not intend to be representative of the early years population in England, however, the target number of 30 settings enabled the intervention to be delivered and evaluated across a variety of settings, allowing the exploration of acceptability and delivery across different setting types. The sample size was relatively large for a pilot study to account for some anticipated attrition. The delivery team recruited 30 settings comprising 13 PVI settings and 17 maintained nurseries—the latter, eight standalone and nine school-based settings in the Berkshire and Oxfordshire areas.⁸ However, ten settings withdrew from programme delivery over the course of the pilot. Characteristics of the recruited, withdrawn, and final samples are discussed in the Participants section under Findings.

Early years practitioners received fee-free access to the training and resources set out above, which would usually cost approximately £800 excluding VAT per setting. Each of the settings received at least one pack of the physical resources and settings with a larger early years cohort or those operating over multiple rooms or sites received up to five sets. Additional sets of physical resources could be purchased for £100, excluding VAT. Settings could keep all of the physical resources at the end of the project. Settings also had access to Tales Toolkit resources online for the year the pilot was running (September 2023 to August 2024). They could also continue to have access to the online platform and resources after the end of the pilot, at a reduced re-subscription rate of £200 per year.⁹

In recognition of the staff time costs associated with undertaking the PD, additional funding was made available (at the end of the project) to cover approximately 50% of staff time costs for attending training: settings could claim £7.50 per hour per staff member who completed the five hours of training, up to a maximum of ten staff members.

Furthermore, settings that took part in the qualitative exploration of delivery received £100 in the evaluation year (2023–2024) once the visit and follow-up interview were complete. In addition, practitioners who responded to the surveys and provided their details were entered into a prize draw. This featured one prize of £50 and two prizes of £25 each at both baseline and endline. In addition, to further boost responses to the endline survey, settings were offered £10 for each survey submitted up to a maximum of £50.¹⁰

There were some exclusion criteria, meaning settings were not eligible to participate in the study:

- Early years settings could only sign up to receive one programme funded through the SPH initiative between 2022 and 2025.
- Settings taking part in the evaluation of the DfE Early Years Professional Development Programme were not eligible to take part.

Recruitment had initially been due to start in January 2023, however, this was delayed (due to reasons outside the control of the delivery and evaluation teams). The Tales Toolkit team was informed at the end of March that it could start recruitment. They were expecting to be supported by the Manager of the SPH, however, they were not in post until April and on starting they had a number of competing priorities in addition to working with new council/SPH leads and settings. SPH staff were supportive, running engagement events that the delivery and evaluation teams attended and making direct contact with settings. However, in this shortened period of time recruitment was difficult. The EEF did agree to expand recruitment to include Oxfordshire to the pilot as it was an authority the Tales Toolkit team had worked with previously.

⁸ The maintained nurseries within an infant or primary school could also deliver to a reception class. In the report, these are also called 'school-based settings' for short.

⁹ A small number of settings were allowed to access the online resources for longer, in cases where they had genuine reasons for not being able to complete all the training before the end of the 2023/2024 academic year.

¹⁰ Twenty-one settings were eligible for the additional payment for responding to the endline survey (six PVIs, eight maintained standalone, seven maintained school-based) but only 14 settings provided payment details (three, five, and six, respectively) and one setting (maintained standalone) declined the payment.

It is important to note the timing of the (delayed) recruitment and the potential knock-on effects of this on a setting's ability to engage with the pilot. For example, most school-based settings were not recruited until May (at the earliest) by which time many will have planned professional development and INSET days for the coming autumn term. The timing may also have been important as settings may have signed up to the pilot before being fully aware of the impact of the expansion of funded childcare on their provision. With the first phase starting in April 2024, settings may have had to start changing their plans for CPD from the autumn term of 2023.

Data collection

Drawing on the EEF IPE guidance (Humphrey et al., 2016; EEF, 2022), a multiphase design was used to examine the research questions. This was based on a triangulation of mixed methods to help fully develop and explore a theory of change model and then test the three main EEF pilot research areas as outlined in the Introduction.

The final theory of change model and description of the key mechanisms, outputs, and changes made is provided in the Findings section under Theory of Change.

The evaluation comprised:¹¹

- one 'intervention delivery and evaluation analysis' (IDEA) workshop reviewing the theory of change and programme materials, plus two set-up meetings to agree on the evaluation design;
- surveys of practitioners at two timepoints—pre-treatment (baseline) and post-treatment (endline);
- qualitative exploration of delivery—a subsample of settings was visited while the intervention was being implemented to observe delivery of the programme and conduct interviews with practitioners and the setting manager(s); staff in these settings were also invited to take part in a follow up interview at the end of the intervention;
- interviews with staff in settings that withdrew; and
- interviews with the delivery team at the end of delivery and collection and analysis of data to examine fidelity.

Table 6 provides a summary of the research, data collection methods, and analysis.

Surveys of practitioners

Online surveys were conducted with practitioners and managers in all recruited settings, at baseline (pre-treatment) and endline (post-treatment). This meant that settings that were involved in the qualitative exploration of delivery (see below) were invited to take part in the surveys but there was no intention to link survey data with the qualitative data, and the surveys did not feature as an element of this data collection method. In addition, a longitudinal cohort formed where settings appeared in both surveys was established to allow comparisons over time, as high attrition rates were observed (see section on analysis). Table 2 shows the total number of settings for which data is available; a detailed description of the process behind each survey data collection is presented below. The surveys were developed using online survey software, SNAP, which allows completion on both desktop and mobile devices via hyperlink or QR code.

¹¹ A detailed timeline is shown in Figure 2.

Table 2: Number of survey participant settings by type of setting

	Baseline survey	Endline survey	Longitudinal cohort
PVI	13	6	6
Maintained nurseries within schools	9	7	6
Standalone maintained nurseries	8	8	6
Total (settings)	30	21	18

Baseline survey

The baseline survey (in Appendix F) aimed to capture information on:

- practitioner characteristics, experience, and knowledge;
- usual practice, including broader approaches to supporting language and communication; and
- child-led play.

The survey included a range of questions and attitude statements using a five-point Likert scale. It drew on the Practitioner Confidence and Skills questionnaire that the IES co-developed with researchers at the University of Oxford as part of the Coaching Early Conversation, Interaction and Language (CECIL) study (Dawson, Huxley and Garner 2022) designed to capture staff confidence, knowledge, and practice to support children’s language and communication development (Appendix G).

The baseline survey was relatively short, taking approximately five to ten minutes on average to complete and contained 12 questions. It was emailed to the lead practitioner in 30 settings in early September 2023 to capture data before starting the training. The lead practitioners nominated themselves during the recruitment stage, and these were usually setting managers or owners or early years leads. The lead practitioners were asked to complete the survey and cascade it to all practitioners who would be involved in the Tales Toolkit.

At baseline, 156 surveys were received from the 30 settings, with at least one practitioner responding in every setting. Responses were received from 13 PVIs, nine maintained nurseries within an infant or primary school, and eight maintained standalone nurseries.

Table 3 shows the distribution of survey responses at each data collection point by job role, including staff who reported having mixed roles.¹² These are grouped by the most senior position as practitioner, teacher, or manager where respondents ticked more than one role. At baseline, 85 respondents performed early years practitioner roles, 42 had teaching roles, and 21 management roles. Eight respondents at baseline noted other roles. The mixed roles varied and included the room lead and teacher or teaching assistant, teacher (QTS) and early years lead, deputy manager and early years teacher.¹³ Given the variety of roles performed by staff, it is a limitation that the survey did not capture their primary role.

¹² This survey question was multi-code and mixed roles could combine one or more other role types as the survey did not collect data on their primary role.

¹³ A description of the group job role categories is in Appendix L.

Table 3: Number of survey respondents by job role

Job role	Baseline survey	Endline survey	Longitudinal cohort
Practitioner	80	23	17
Mixed role—practitioner	5	1	1
Teacher	35	17	15
Mixed role—teacher	7	4	3
Manager	15	4	2
Mixed role—manager	6	2	2
Other	8	5	1
Total (responses)	156	56	41

Endline survey (post-treatment)

The endline survey was longer than the one used at baseline, with 30 questions taking, on average, ten minutes to complete. A number of questions asked at baseline were repeated to report change over time, for example, around staff confidence, knowledge, skills, and areas of child development (in CLL, PSED, and creativity). A number of new questions were also added to enable staff to reflect on their experiences of the training (including reasons for not completing the training) and delivering the intervention (Appendix H). Staff were asked about their engagement with different training sessions, how they were delivering Tales Toolkit with children at their setting, whether any adaptations were made (beyond expected differentiation to meet the needs of individual children), the time and the resources required to participate, and any perceived outcomes for the children.

The survey was sent out in June 2024 to capture findings towards the end of the intervention period. An individual link to the survey was emailed to all 156 staff who had submitted a baseline survey the previous autumn—including staff in settings that had withdrawn from delivery or that had reported partial engagement—to try and capture their views on the programme and their reasons for not completing the training or delivery. In addition, a general link was sent to the lead practitioners, which cascaded to staff who could or did not respond to the baseline survey.

As shown in Table 2, at the endpoint there were 56 responses from 21 settings: six PVI, eight standalone-maintained nurseries, and seven maintained settings within an infant or primary school, resulting in a setting-level attrition rate of 30% (n = 9). Of these responses, 40 were from the individual link (response rate of 26%) and 16 from the general link. However, some practitioners who were sent an individual link responded via the general link instead (this leads to an effective response rate of 36% relative to the response rate obtained in the baseline). As a consequence, the number of staff involved in Tales Toolkit significantly decreased affecting the reach of endline survey participants endline as shown in Table 2.

Attrition between baseline and endline surveys

Since nine settings did not complete the endline survey, equivalent to 30% of the recruited sample, the level of attrition is considered high. The evaluation team recorded insights into reasons for this as part of the recruitment activities to secure the settings taking part in the qualitative exploration of delivery (which started in March 2024). During this process, managers in many settings reported significant staffing issues impacting their ability to engage with the intervention and evaluation. They also provided possible reasons for attrition at the setting level and the low response rate to the endline survey. Among the reasons suggested for either withdrawing from the pilot or partial delivery included changes in management since baseline (for example one PVI business was sold and the manager who had signed up to the pilot left; in another, the pilot was not taken up by a new manager, in another the new manager did not know anything about the pilot), staff leaving the setting, and staff illness. Additional demands on staffing were also being

reported due to the expansion of funded childcare. This new policy came into place during the pilot in April 2024 causing staffing issues for the early years sector as a whole. The timing of the endline survey could also have been a factor, as settings (especially school-based ones) can be very busy towards the end of the academic year.

To address the low response rates at the endline, the EEF agreed upon several mitigations (in addition to the prize draw and additional incentive already discussed). Each member of staff who had submitted a baseline survey was emailed by the evaluation team at least three times and invited to complete the endline survey. The evaluation team also reminded staff in the case-study settings, both in the visits and follow-up interviews, to respond to the survey and remind colleagues to respond to it. The EEF also agreed to the endline survey staying open longer than originally proposed, beyond the end of the school summer term, to allow practitioners more time to respond, particularly PVI settings that do not close for the summer. Despite these mitigations, the response rate to the endline survey was lower than expected.

Qualitative exploration of delivery

The IES initially sampled 12 settings in which to explore delivery qualitatively using information from settings themselves via fields in the MoU and additional local intelligence from the Thames Valley Early Years Stronger Practice Hub. The key characteristics used for selection were setting type, location, and a range of levels of deprivation, but with a focus on those in areas of higher deprivation (based on the Income Deprivation Affecting Children Index (IDACI) scores). Although a small number of potential replacement settings had also been identified (in the event of any of the initial 12 being unable to take part), these were unable to commit to taking part and the evaluation team were able only to secure the participation of 11 settings. The reasons given by the settings were primarily staff shortages and lack of availability,¹⁴ reflecting wider issues reported by pilot settings in terms of delivering the intervention and, ultimately, some withdrawing.

Table 3 shows the sample information for the settings that were involved. The sample allowed the evaluation team to explore experiences from a range of setting types, with a skew towards those in areas of higher deprivation, as expected. In terms of location, originally it had not been intended that settings in Oxfordshire would be included in the pilot. However, recruitment difficulties meant that EEF agreed to settings in Oxfordshire being invited to take part. In the end, the number of settings recruited in this area was quite small, which made it harder to secure their participation.

The qualitative research was designed to capture data from settings at two-timepoints: during the intervention in the form of a one-day in-person visit between March and May 2024 and at the end of the intervention in a follow-up virtual interview (in June and July 2024).

The visits collected data on the implementation and delivery of Tales Toolkit. In addition to interviews with key staff, such as setting managers and practitioners, observations of staff delivering Tales Toolkit with children were also carried out. When arranging visits, evaluators gave information about planned data collection (for example, observations, interviews) and the setting, then suggested best timings for these from their perspective (for example, attending morning and afternoon sessions separately across two days). Evaluators also sought to be responsive and adaptive on the day so as not to interfere with the setting's normal delivery.

Table 4: Distribution of the qualitative data collection sample by type of setting, county/region, and IDACI percentile

	Number of settings
Setting type	
Private or for-profit nursery (PVI)	3
Maintained nursery school—standalone	4
Maintained nursery within an infant or primary school	415

¹⁴ This was also the reason given by the potential 12th setting that had initially been selected.

¹⁵ All of these four settings delivered to nursery and reception class.

County/region	
Berkshire	10
Oxfordshire	1
IDACI¹⁶	
Deprived 30%	3
Average 50%	5
Least deprived 30%	1
Least deprived 20%	2

Table 5 provides an overview of the settings participating in the qualitative research and the data collected at the two timepoints. This includes the number of interviews and observations at each stage and information on the interviewee's job role.

Table 5: Overview of the qualitative data collection by timepoints

Qualitative data collection	Timepoint 1: during the intervention	Timepoint 2: at the end of the intervention
Settings	11	10
Observations	11	N.A.
Individuals interviewed	36	17
By job role		
Managers	13	9
Practitioners	23	8

Interviews with staff

Two rounds of interviews were carried out as part of the qualitative data collection, the first during the implementation of the intervention in a setting visit and the second in a virtual follow-up interview.

The plan for the visits was to interview two staff members in each setting, for example, the setting manager or head of EY and a practitioner delivering the intervention. However, a number of settings offered the opportunity to interview more than two members of staff, which we took up where possible. All interviewees were offered separate interviews to ensure open and honest discussions, although some settings opted for joint staff interviews (due to time pressures and availability of staff on the day). This presented risks that interviewees might not speak as freely about their own experiences. Balancing this, paired interviews could provide peer support increasing confidence to share experiences and the opportunity to triangulate between experiences in discussions. Interviews lasted around 30 to 45 minutes and were scheduled flexibly to allow all staff to participate outside of ratio restrictions.

For the follow-up interviews, towards the end of the intervention, all the leads and staff who had been interviewed in the first round and provided us with their email addresses were invited again. Although most of these interviews were conducted separately, some were carried out as joint interviews (for example, the manager and practitioner together) as this was requested (reportedly mainly due to time pressures or staffing schedules and ratios). It should be noted that

¹⁶ Income Deprivation Affecting Children Index (<https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019>)

joint interviews may have limited the staff's ability to be honest in front of colleagues, especially where a power imbalance may exist such as between practitioners and managers. This may mean the findings could be biased in some way although researchers did not detect signs that discussions were constrained by power imbalances.

Despite the multiple attempts of the evaluation team to conduct follow-up interviews, it was not possible to interview the same people as in the first visits—sometimes put down to staff changes. For example, in one setting the lead practitioner had moved to another role within the broader organisation and did not have time; in others the practitioners had left the setting. Further information is available in Table 6.

The interviews explored the following topics (Interview topic guides are in Appendix I):

- experience and engagement with training, as well as materials;
- workload and time requirements of delivery;
- implementation costs;
- facilitators and barriers to implementation;
- adaptations and reasons—including views on how children with SEND, EAL, or in receipt of EYPP responded and engaged;
- how Tales Toolkit compares to the usual practice;
- perceived outcomes for children;
- costs incurred by settings, including time and whether this was manageable and acceptable;
- issues which would need to be considered for wider roll-out; and
- suggested improvements.

Observations of practice

During visits, the evaluators also carried out semi-structured observations of practitioners delivering the intervention with children (see Table 6). The observations were designed to assess fidelity, allowing for data triangulation (with the surveys and interviews) and to shape questions in the follow-up interviews. The observations focused on the feasibility of delivery and child engagement during Tales Toolkit sessions. The approach involved the development of a semi-structured observation framework based on the AEIOU model (Robinson et al., 1991). This model, rooted in ethnographic methodology, organises information into five key elements: 'activities', 'environments', 'interactions', 'objects', and 'users' (see Appendix J). By using this model, evaluators could identify key components of the intervention and observe how children were engaging. The evaluators aimed to be as unobtrusive as possible by explaining to the staff before the session what they would be doing (verbally and in the staff information sheet) then sitting quietly at the back of the room. The researchers did not engage with the children unless they engaged with them in which case, they told them their name and explained they were just watching their teacher. However, it must be acknowledged that there may have been an observer effect where practitioners or children may not have behaved exactly as they usually would. Evaluators gave assurances that they were not there to judge or assess practitioners' performance. However, practitioners may have felt slightly nervous or anxious regardless, especially in the context of the sector where settings are externally assessed by organisations such as Ofsted. Data collected using the model was qualitative and incorporated into the analysis framework with the interview data from the visit to enable thematic analysis.

Interviews with settings or practitioners that withdrew

Ten settings withdrew from programme delivery during the pilot, throughout the delivery: seven were PVIs and three were maintained settings within a school. In most cases, the reasons for withdrawal were due to staffing pressures. For example, one early years lead/reception teacher at one of the withdrawn schools who completed the endline survey reported that they had to withdraw from programme delivery (after completing the first session of training) because the staff had to be redirected due to child behaviour issues: 'We started but then due to extreme behaviour we had to redirect staffing elsewhere.' The evaluation team was also able to conduct interviews with four of these ten settings' managers to explore their reasons for withdrawal, including any barriers to facilitating the training or delivering the intervention, as well as intentions to deliver the Tales Toolkit in the future. All interviews were with PVI settings and carried out by

telephone during the summer term of 2024 (interview guide in Appendix I). The managers in all four reported staffing issues as the main reason for ending their involvement in the pilot, and three reported that staffing issues would need to be resolved for them to deliver Tales Toolkit in the future. The managers in the remaining six settings were contacted (multiple times) and invited to take part in the interviews: their lack of response might suggest that they had been experiencing similar issues with staffing to the ones reported by those interviewed in the four settings.

Delivery team data

No observations of the intervention training sessions were conducted as these were led online at a time that suited the setting. Instead, the evaluation team carried out a review of the learning materials; this enabled the team to gain an understanding of the intervention and served as a foundation for preparing observation protocols, interview guides, and surveys.

Evaluators also conducted an online interview with the Tales Toolkit delivery team at the end of delivery and explored training delivery, setting engagement and participation, as well as enablers and barriers to successful implementation and scaling up of the programme (interview guide in Appendix I). We had also planned to collect and analyse data on fidelity from the delivery team, however, the team had received very few requests for support or delivery queries from study settings so there was very little data to analyse regarding the support provided. In addition, the Tales Toolkit website was hacked during the study period and so had to be completely rebuilt. This meant that the training and other online resources were very briefly unavailable and this made it difficult to track viewing data over the course of the study.

Table 6: Methods overview

Evaluation dimension	RQ addressed	Research methods	Data collection methods	Planned sample size and sampling criteria	Achieved sample size	Data analysis methods
Evidence of promise	1, 1a, 1b, 2, 3, 4, 5	Practitioner survey	Online questionnaires (baseline and post-treatment/endline)	Baseline N = 60; Endline N = 60. At least 2 in each of the 30 settings at both timepoints.	Baseline N = 156 (from all 30 settings). Endline N = 56 (from 21 settings).	Descriptive frequencies, cross-tabs if sample size allows
	1, 1a, 1b, 2, 3, 4, 5	Session observation	Semi-structured observation using the AEIOU tool	12: at least 1 in each of the 12 settings.	11: at least 1 in each of the 11 settings.	Thematic analysis
	1, 1a, 1b, 2, 3, 4, 5	Interviews	Semi-structured interviews	Timepoint 1: 24 (12 setting managers/EY Leads and 12 practitioners). Timepoint 2: 24 (same staff as at timepoint 1).	Timepoint 1: 36 setting managers/EY Leads and practitioners. Timepoint 2: 17 (out of 36 carried out at timepoint 1).	Extraction framework, thematic analysis
	1, 1a, 1b, 2, 3, 4, 5	Delivery team interviews	Semi-structured interviews	2/3 members of the Tales Toolkit delivery team (programme managers and developers).	4 members of the Tales Toolkit delivery team (programme managers and developers).	Extraction framework, thematic analysis
	1, 1a, 1b, 2, 3, 4, 5	Monitoring Information	Analysis of intervention data collected by Tales Toolkit	30 settings.	30 settings.	Thematic analysis, frequency counts
	1, 1a, 1b, 2, 3, 4, 5	Withdrawal interviews	Interviews with practitioners or managers/settings that withdrew from the study	Up to 10.	4 managers.	Extraction framework, thematic analysis
Feasibility of implementation	6, 7, 8, 9, 10	Practitioner survey	Online questionnaires (baseline and post-treatment/endline)	Baseline N = 60. Endline N = 60. At least 2 in each of the 30 settings at both timepoints.	Baseline N = 156 (from 30 settings). Endline N = 56 (from 21 settings).	Descriptive frequencies, crosstabs if sample size allows
	6, 7, 8, 9, 10	Session observation	Semi-structured observation using the AEIOU tool	12: at least 1 in each of the 12 settings.	11: at least 1 in each of the 11 settings.	Thematic analysis

	6, 7, 8, 9, 10	Interviews	Semi-structured interviews	Timepoint 1: 24 (12 setting managers/EY Leads and 12 practitioners). Timepoint 2: 24 (same staff as at Timepoint 1)	Timepoint 1: 36 setting managers/EY Leads and practitioners. Timepoint 2: 17 (of 36 carried out at timepoint 1).	Extraction framework, thematic analysis
	6, 7, 8, 9, 10	Delivery team interviews	Semi-structured interviews	2/3 members of Tales Toolkit delivery team (programme managers and developers)	4 members of Tales Toolkit delivery team (programme managers and developers)	Extraction framework, thematic analysis
	6, 7, 8, 9, 10	Monitoring Information	Analysis of intervention data collected by Tales Toolkit	30 settings	30 settings	Thematic analyses, frequency counts
	6, 7, 8, 9, 10	Withdrawal interviews	Interviews with practitioners or managers/settings that withdrawn from the study	Up to 10	4 managers	Extraction framework, thematic analysis
Readiness for trial	11, 12, 13, 14, 15, 16, 17	Practitioner survey	Online questionnaires (baseline and post-treatment)	At least 2 in each of the 30 settings (60min) at both timepoints.	Baseline N = 156 (from all 30 settings) Endline N = 56 (from 21 settings)	Descriptive frequencies, crosstabs if sample size allows
	11, 12, 13, 14, 15, 16, 17	Session observation	Semi-structured observation using the AEIOU tool	12 - at least 1 in each of the 12 settings.	11 - at least 1 in each of the 11 settings.	Thematic analysis
	11, 12, 13, 14, 15, 16, 17	Interviews	Semi-structured interviews	Timepoint 1: 24 (12 setting managers/EY Leads and 12 practitioners) Timepoint 2: 24 (same staff as at timepoint 1)	Timepoint 1: 36 setting managers/EY Leads and practitioners Timepoint 2: 17 (out of 36 carried out at timepoint 1)	Extraction framework, thematic analysis
	11, 12, 13, 14, 15, 16, 17	Delivery team interviews	Semi-structured interviews	2/3 members of Tales Toolkit delivery team (programme managers and developers)	2/3 members of Tales Toolkit delivery team (programme managers and developers)	Extraction framework, thematic analysis
	11, 12, 13, 14, 15, 16, 17	Monitoring Information	Analysis of intervention data collected by Tales Toolkit	30 settings	30 settings	Thematic analyses, frequency counts
	11, 12, 13, 14, 15, 16, 17	Withdrawal interviews	Interviews with practitioners or managers/ settings that withdrawn from the study	4 managers (up to 10 planned)	4 managers (up to 10 planned)	Extraction framework, thematic analysis

Data analysis

The quantitative data from the surveys and the qualitative data (from the interviews and observations) were used to triangulate learning. The data was analysed as follows.

Qualitative data analysis

Interviews were digitally recorded with the participants' agreement and transcribed verbatim. Researchers analysed the data using a 'framework' approach. This approach involves reviewing the transcripts and summarising (coding) the data and key quotes deductively under themes covered in the interviews. Themes can then be analysed inductively and deductively to pick up on nuances and different dimensions beneath the headline theme. Data from the observations of Tales Toolkit sessions was also coded into a framework under the headings used in the AEIOU tool. Additional information relevant to the evaluation was also coded, including any critical reflections identified by the researcher. The framework approach is an Excel-based qualitative analysis tool that ensures that analytical processes and interpretations are grounded in the data and tailored to research questions.

Frameworks allow for analysis within settings (looking in detail at each case) and between settings (comparing settings across the sample) as well as thematic analysis. This facilitates analysis to understand impacts and experiences within a setting context and to compare across specific characteristics (for example, role, setting type). By using this approach, researchers compared what was observed in the Tales Toolkit sessions with what staff reported in the interviews (in the 11 settings taking part in visits). Analysing the interview data from the two timepoints also allowed a comparison of change over time (for the staff who participated in both). However, during the iterative experience of undertaking fieldwork and analysis it also became clear that setting type and contextual factors had a stronger influence on implementation than the internal dynamics in individual settings. This led to reporting primarily on the different sources of qualitative data rather than a full integrative analysis by setting. This shift away from full case integration may affect the perceived rigour and transparency of the analysis. The approach prioritised comparison of staff data (interviews and observations) across settings with coding and iterative interrogation leading to the identification of recurring patterns, themes, and meanings from across these settings rather than first exploring these dynamics within a case. This approach allowed key themes in the delivery experience to emerge more clearly. However, the limitation is reduced insight into common factors in the settings' dynamics that may have led to these themes being present.

Analysis of survey data

Quantitative analysis of the survey data, using SPSS, was purely descriptive and no statistical tests were conducted. In the main, descriptive frequencies are reported, however, where appropriate and the sample size allowed, some cross-tabular analysis of different groups is also noted, for example, by job role. Some key points about the presentation of the survey data are made below.

- Due to the low number of survey respondents at endline, the percentage, number of responses ('n ='), and the valid (base) number of responses ('N =') for each question are presented for transparency.
- The endline survey results include new staff—who had not responded to the survey at baseline—as well as some who had completed it so that as much data as possible can be used to inform findings.
- If respondents did not answer a question, the data was coded as 'missing'.
- Due to this missing data and the low response rate at the endline, the findings should be interpreted with caution.
- The two most positive options on questions using a five-point Likert scale (for example, 'a fair amount' and 'very much/frequently') are combined into 'agreed' unless otherwise stated.¹⁷

¹⁷ The scale used throughout is: 1: 'not at all'; 2: 'slightly'; 3: 'somewhat'; 4: 'a fair amount'; 5: 'very much/frequently'; 6: 'not applicable/don't know'.

Longitudinal cohort

As a consequence of the high level of attrition, a longitudinal cohort was created in the analysis phase (August 2024). This contained only data from staff who responded to both surveys (at baseline and endline). This approach was not proposed in the study plan as the level of attrition was not anticipated. As the longitudinal sample size is necessarily smaller than the endline as a whole, this brings risks of the data having limited reliability. To manage any such risk and to promote transparency and interpretation, descriptive data (of the whole cohort) is presented alongside the full endline survey findings.

In addition to the descriptive analysis of this cohort, the results from the longitudinal cohort were also analysed to explore changes over time among staff involved throughout the entire evaluation. This was likely to portray more clearly any change dynamics since new entrants to the sample (at endline) could not comment on practice before starting delivery (as documented at baseline).

This analysis was carried out across four dimensions (see Evidence of Promise), however, only those in the longitudinal cohort who responded to all questions (in all four dimensions) at both baseline and endline were included in this analysis. This enabled a more robust analysis of change over time than if the entire longitudinal sample had been included. As a result, five cases were excluded from the analysis as they had not provided valid responses to all of these questions at baseline and endline (N = 36).

Percentages for this analysis were calculated using this fixed subsample and are rounded up to the nearest whole number. It is important to note that, unlike the baseline survey, the endline survey included a 'not applicable/don't know' (NA/DK) response option for certain questions. These responses are retained within the complete case dataset and included in the denominator when calculating percentages in order to preserve a consistent base for comparison across timepoints. While this approach may slightly understate the proportion of respondents providing a substantive answer, it ensures that time-based comparisons are made on a like-for-like basis. NA/DK responses are shown separately, where relevant, for transparency ranging from zero to four. For these questions, using a five-point Likert scale, the average (mean) response between one and five was calculated. The means were compared and only the questions with the biggest change between baseline and endline are presented in the main report (along with each base number and percentage responding most positively). The mean results for questions with smaller or no changes in the mean are reported in Appendix K, together with the associated charts. However, these results are only an indication of the change between baseline and endline and must be interpreted cautiously as the sample size did not allow any significance testing (which would need to be tested in a future evaluation). The sample size did not allow for any analysis of the longitudinal cohort by setting type (due to the small number of PVIs).

Timeline

Figure 2 shows the timeline of activities related to the evaluation and intervention delivery.

Figure 2: Timeline of activities

Date	Activity
November 2022	First set-up meeting and IDEA workshop
January 2023	Second set-up meeting
January–March 2023	Ethical approval, developing recruitment materials
March–July 2023	Recruitment of settings
September 2023	Baseline survey of practitioners
September 2023–Spring 2024	Recommended period for training delivery
September 2023–Summer 2024	Intervention period
Autumn 2023–Summer 2024	Collection of programme data from the delivery team
March–May 2024	Setting visits—observation of delivery to children, interviews with practitioners
June–July 2024	Setting interviews with practitioners (end of intervention), delivery team interviews
June–August 2024	Endline survey of practitioners
September–December 2024	Analysis and reporting
October 2024	Emerging findings presentation
January 2025	First draft of evaluation report

Findings

Participants

In total, 30 early years settings were recruited to take part in the Tales Toolkit pilot and signed an MoU to participate. As this was a small pilot study, the sample was intended to include a range of setting types and contexts rather than to be representative of the early years population in England. Table 7 presents the distribution of recruited, withdrawn, and participating settings by three main characteristics: type, county/region, and deprivation levels using the Income Deprivation Affecting Children Index (IDACI).¹⁸

Key features of the recruited sample of 30 settings are:

- 13 were private, voluntary or independent (PVI); the others were almost evenly split between maintained linked nurseries (9) and maintained standalone (8);
- 23 were located in Berkshire, the remaining seven in Oxfordshire; this was broken down further into sub-areas to ensure geographic coverage within each region; and
- seven settings were in deprived areas (IDACI Deciles 1 and 4).

Table 7: Distribution of settings (recruited, withdrawn, and participating) by type, county/region, and deprivation deciles

	Recruited settings	Withdrawn settings	Participating settings
Setting type			
Private, voluntary, or independent (PVI)	13	7	6
Maintained nursery school—standalone	8	0	8
Maintained nursery within an infant or primary school	9	3	6
County/region			
Berkshire	23	7	16
Oxfordshire	7	3	4
Deprivation - IDACI deciles *			
Decile 1: most deprived 10%	1	1	0
Decile 4: deprived 30%	6	2	4
Decile 6: average 50%	7	0	7
Decile 7: least deprived 40%	3	1	2
Decile 8: least deprived 30%	2	1	1
Decile 9: least deprived 20%	9	4	5
Decile 10: Least deprived 10%	2	1	1
Total	30	10	20

¹⁸ The Income Deprivation Affecting Children Index (IDACI) is a subset of the Income Deprivation Domain, which measures the proportion of the population in an area experiencing deprivation relating to low income.
<https://opendatacommunities.org/def/concept/general-concepts/imd/idaci>

Settings that withdrew were still asked to complete the evaluation activities. Five staff from three settings that withdrew (two PVIIs and one school-based) responded to the endline survey.

* No settings were recruited in Deciles 2, 3 or 5.

Attrition occurred throughout the intervention period: of the 30 settings initially recruited to the pilot, one withdrew prior to taking part and nine notified the delivery or evaluation team that they had withdrawn from programme delivery part-way through the academic year, leading to an overall attrition rate of 33%. Seven of the ten settings that withdrew were PVIIs (70%), reflecting difficulties reported by such settings participating in the pilot. The reasons given to the evaluation team by setting managers for dropping out were not to do with the intervention but instead with their ability to engage with the pilot primarily due to staffing pressures (which have been reported across the early years sector as a whole). Reasons given by PVIIs included unforeseen circumstances and changes in management: for example, the business was sold, the manager who had signed up to the pilot had left and it was not taken up by the new manager or the new manager did not know anything about the pilot.¹⁹ The expansion of funded childcare was specifically mentioned by two managers/owners:

'We are not able to deliver (Tales Toolkit) currently due to staffing shortages, as well as being at full capacity with child places in their nursery. This has been the case since the expansion of the free childcare hours so since April 2024 this has been acutely felt in their setting. We are also short-staffed more widely across our nursery chain. We have chosen to focus on embedding core curriculum with new children in the nursery and this has left us with no time to use Tales Toolkit. I am hesitant to stress or burden staff further with a new programme. We're not sure if we will implement this next academic year as it's unlikely our resourcing/staffing issues will alter in this time' (manager, PVI).

'Staffing has been a big issue for us this year. Being a rural setting has impacted our staff hiring as they need to be able to drive to get to us. The nursery expansion has impacted on us, the number of children is growing this year too (partially due to long waiting list and partially due to hours expansion). We have issues with literacy in our setting, and Tales Toolkit will potentially help with that, but we won't be able to deliver anything until June/July at the earliest' (manager, PVI).

The three maintained settings that withdrew also reported issues around staffing including one manager who reported that changes in staff meant all the new staff would need to be trained (which was not possible), and another setting where staff were seriously ill.²⁰

The final participating sample comprised six PVI settings (down from 13 at recruitment), eight maintained standalone settings, and six maintained nurseries within an infant or primary school (down from nine at recruitment). The number of participating settings in deprived areas (IDACI Deciles 1 and 4) dropped from seven to four (as three withdrew). This limits the representation of settings serving children in deprived areas (and potentially, the ability of this pilot to explore how the intervention worked for this group). However, the proportion of settings in deprived areas in the sample only dropped slightly (from 23% to 20%) due to high attrition among those in the least deprived areas (from 16 to nine).²¹ Also, the number of settings in the average area (Decile 6) remained at seven, leading to an overall increase in their representation (from 23% to 35%). The balance of settings by region was similar to the recruited sample.

Feasibility

This section investigates whether it was feasible for settings to deliver the Tales Toolkit programme and whether they were able to deliver it with fidelity. Information from the case studies and surveys is used to explore delivery, any changes or adaptations made, staff views on the training, intervention and resources, any facilitators and barriers and how they

¹⁹ This feedback was collected while contacting settings for recruitment to the case studies, from March 2024 onwards. The information was discussed with the EEF and the delivery team at the time. The delivery team also reported having problems making contact with some settings to arrange backfill payments for staff time spent participating in the Tales Toolkit training, which the EEF provided funding towards.

²⁰ All settings that withdrew from the study were invited to take part in a short interview with the evaluation team or send a brief email response to discuss their reasons for withdrawing.

²¹ In Deciles 7 to 10.

were supported to deliver. The following research questions will be covered in the sections below. Where necessary, the RQ has been split so different components can be analysed separately.

RQ6 Is the intervention feasible for practitioners to engage in and implement?

RQ6.1 Is the training feasible for practitioners to engage in?

RQ6.2 Is the delivery feasible for practitioners to implement?

RQ7 Can settings deliver with fidelity and within routine practice? Is there variation in delivery across settings, for example, embedded into the curriculum, planning, practice, resources, and language or in separate sessions/activities?

RQ8 Does the delivery method of the training (context and mode) have an impact on engagement with the programme (practitioners)?

RQ9 What are the potential barriers and facilitators of the programme?

RQ10 What support is needed, and how can a useful support model be developed?

RQ6.1 Is the training feasible for practitioners to engage in?

The information sent to settings that participated in the pilot set out that each training session was 50 minutes long and the requirement was that delivery staff should complete all five training sessions as a group, ideally simultaneously. However, the delivery team also allowed for staff to watch the sessions in smaller groups to ensure the interactive elements were fully utilised. In these circumstances, the lead practitioner was expected to check that all relevant staff had completed the training. Although the training sessions were asynchronous, meaning settings could choose to deliver the training when it suited them, the pre-training information provided to lead practitioners outlined an expected timeframe for the five sessions: the first and second training session could be close together (ideally a week apart but at the same time if necessary), leave at least a month between Sessions 2 and 3, leave another month between Sessions 3 and 4, and, finally, leave at least a month until running Session 5 (at the end of term) so that staff can reflect on the intervention.

The findings suggest a mixed picture of engagement. As presented in Table 8, of the 56 staff who responded to the endline survey, 50% (n = 28) reported completing all five training sessions, while 64% (n = 36) had completed at least three. The interviews with staff confirmed the results of the survey: that the reasons for the variations in training completion were primarily due to staffing issues. Staff who had participated in the training were very positive about it, reporting that the resources were helpful and gave them sufficient skills and knowledge to deliver the programme.

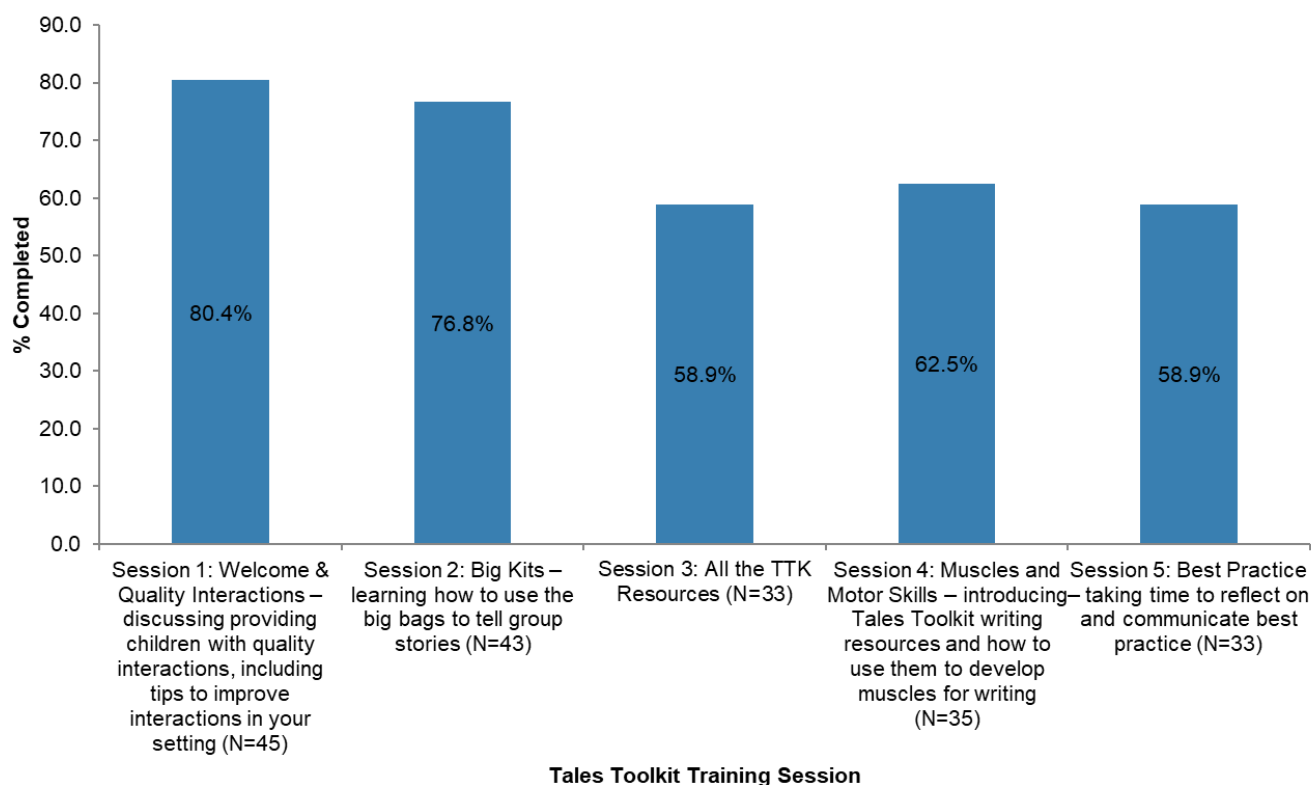
Training completion

The training completion rates are explored in two ways below using the endline survey data—completion rates of each session (Figure 3) and number of staff reporting completion of the various sessions in the endline survey (Table 8).

Figure 3 shows the participation rates for each of the five training sessions. For 22 of the 56 staff who responded to the endline survey, the first two had the highest levels of involvement (80%, n=45 and 77%, n=43), followed by Session 4 at 63% (n = 35), and then Sessions 3 and 5 (both 59%, n = 33) (N = 56).

²² Participation rates were asked as separate questions with staff asked to tick the box if they had done the session. There was also an option for staff to say they had not taken part in any Tales Toolkit training. The remaining cases in each question were system missing values.

Figure 3: Completion rates of the five Tales Toolkit training sessions



Source: Evaluation endline survey 2024.

Table 8 shows the breakdown of the total number of Tales Toolkit training sessions completed by endline survey respondents. Looking at all staff, 28 (50%) reported completing all five and 36 (64%) had completed at least three sessions. This reflects a moderate engagement with the training for all respondents.

Table 8: Total number of Tales Toolkit training sessions completed

Total number of Tales Toolkit training sessions completed	All	
	N	%
1	4	7.1
2	9	16.1
3	5	8.9
4	3	5.4
5	28	50.0
At least 3 sessions	36	64.3
No sessions	7	12.5
Total	56	100

Source: Evaluation endline survey, 2024.

Seven staff (12.5%; three practitioners and four managers) reported doing no training at all (N = 56). They came from a range of setting types: four from PVIs, two from school-based settings, and one from a standalone setting. However,

this included only one practitioner who had not been offered the training but was delivering (this member of staff worked in a PVI setting where staff were expected to watch the training on their own, which is described in more detail below). The reasons given by the other staff for not attending any training were: Tales Toolkit not being delivered to the age group they teach (1), non-teaching headteacher (1), unable to attend the training (2), joining the setting later in the year (1), and staffing changes and a lack of staff (1).²³

It was expected that someone in the setting, either the lead practitioner or another member of staff, would facilitate the training. The endline survey found that eight managers (of 13) and nine practitioners (of 43) actively supported the training sessions as facilitators (N = 56). In the interviews, two of the staff were training facilitators and they reported that they watched the sessions before the rest of the staff so they could lead the training and associated discussions and activities.

In the study interviews, staff reported that most training happened as expected, being shared in groups led by the facilitator in staff meeting times. However, there was still wide variation in how this was done on a setting-by-setting basis. Where training was done in groups, this was mostly done in team meetings (generally after school). The size of the group varied, depending on the size and type of setting. Examples include:

- a maintained, school-based setting—all the teaching staff in nursery and reception watched the first four videos in team meetings, as did the SENCOs (but not the nursery nurses or LSAs); the fifth video everyone watched separately in their own time, then came together to discuss it;
- a maintained school-based setting—around nine out of the 11 staff had watched the videos together, mostly after school in team meeting times; there were tech issues where some staff who missed sessions could not go back and catch up: new staff had not watched them yet; and
- a maintained standalone setting—12 staff members had done training in team meetings but had sometimes skipped some of the activities due to time constraints; new staff members had not completed training (had been delivering after watching other staff) and will watch videos new academic year.

Some settings ran group sessions for some of the training but then asked staff to complete other sessions on their own, for example, in one of the PVIs (that was part of a chain) the manager and practitioner had done the training, at different times, some in staff training sessions and some outside nursery time in the evening.

Two settings only ran the shorter training sessions including a maintained standalone setting where all staff (except one-to-one support staff) watched the shorter videos for the first four sessions during staff meetings, however, no staff watched the fifth session. In another maintained standalone setting, ten members of staff (not TAs) watched the 25-minute versions of training in staff meetings (staff who were not there were given logins to watch them in their own time).²⁴

Staff views on the training sessions

In the endline survey, 70% of staff (38 of 54) agreed that participating in the training fitted into their working day and routine. Managers and practitioners who were interviewed were also positive about the sessions, finding the videos well-presented, clearly broken-down, insightful, and valuable. Some staff also specifically mentioned the balance of research, practical usage, and modelling. Positive feedback from staff who had done the training included:

'I like videos because that's kind of my way of learning as well, I like to see other people doing things for me to do it. Rather than reading notes, if I read notes, I find it really hard to implement. So, I much prefer a video to show you what to do, so yeah it was very good' (practitioner, maintained standalone).

²³ The endline survey contains responses from some staff who were new to the setting (and did not complete a baseline survey). In one school-based setting, the reception teacher had only done session one before having to stop the training as, due to extreme behaviour, they had to redirect staffing elsewhere.

²⁴ The shorter sessions were designed for staff to complete the training on their own and did not include time for the group discussion or exercises.

'Training was really good, not too long, enough balance of research, practical ideas and really good to see it in practice. Liked that it showed age ranges (two years and reception children). Good to hear from practitioners and the Tales Toolkit lady, as well as Julian Grenier. Good mixture of theory, practical, and watching how a session was being delivered. More realistic than other training—normal classroom where children are shouting out' (manager, maintained standalone).

'The videos are really good and informative, the length of videos was fine and it's good to have time to discuss things within the videos. I liked how they showed practitioners doing it with the children involved, and it really helped to show me how to do it. The third one made me think of new ideas, you know, how to expand it' (practitioner, PVI).

Managers at four setting visits reported that Session 1, 'Quality Interactions', was particularly useful and that they had received feedback from practitioners who had found it helpful.²⁵ The reasons given by staff included that the video showed specific evidence of how using quality interactions benefitted children and that it reminded practitioners of the importance of engaging and interacting in this way with the children. Managers also said the session was useful because it allowed practitioners to reflect on what they found difficult in this area and to identify ways to overcome this and ultimately improve the quality of interactions.

'I think it's improved the quality of interactions. It's really highlighted to us as well who's got the potential for really enjoying those sessions, and therefore who can cope with interactions in a different way. I think it's made some staff reflect on if they find those kind of things particularly stressful ... that thinking about that for their other interactions ... I think doing the training this way has made staff reflect a bit more on why they find that difficult to do, and that has opened avenues for us to talk to them about that a little bit more, and be able to say that actually that's what we need to be doing all the time, so it's been useful from that staff training point of view as well' (manager, maintained within a school).

Session 1 was so valued by one PVI setting that they planned to show the video to all of their staff, regardless of whether they were going to deliver Tales Toolkit, because of how useful it was. One early career teacher (ECT) had found the first sessions particularly helpful because their ECT training did not have a specific focus on the early years.

Staff views on the training resources

The Tales Toolkit training included downloadable resources to be used alongside the training, such as facilitator notes and reference materials for delivering the programme, such as the summary sheet. In the endline survey, 82% (n = 44) agreed that the training (videos and resources) had given them sufficient knowledge and skills to be able to deliver the programme, and they had found the training videos and other resources helpful (N = 54).²⁶

The 11 staff interviewed who had used the downloadable resources found them helpful, especially the summary sheet. Staff also felt that the time in the training provided for discussion was helpful, giving them time to think of ideas for delivering the children at their setting. It also gave practitioners time to ask each other questions and learn from each other.

'Having other people in the group, you all have ideas, and talk about it, and actually then it all starts to click and you think, oh yeah, I didn't think of it that way, or that way, etc.' (Practitioner, PVI).

Not all staff who participated in the training experienced this element, either because they viewed the shorter training videos or viewed training at home or separately. They could not discuss with colleagues and may have skipped these activities entirely. The topic of peer support in delivering the Tales Toolkit is explored further in Facilitators to Implementation. Almost half the managers interviewed had chosen not to use the downloadable resources or the videos with the discussion time. They reported that this was mainly because they did not have time to watch the full-length videos, although the manager and practitioner in one setting said they could not find the downloadable resources.

²⁵ It is important to note that more staff engaged with session one compared to later sessions.

²⁶ A number responded 'not applicable or don't know' to these questions, indicating that they may not have completed some or all of the Tales Toolkit training sessions.

Two managers, who had acted as facilitators for the training, commented that they found the facilitator notes useful and had planned out which members of staff would be best to support each other during the training, based on the knowledge of the practitioners, and their strengths and weaknesses.

Reasons for variations in training

As reported, the setting interviews discovered wide variations in training completion within and between settings. The reasons for this were not generally explained by the type of setting but related to wider issues affecting the early years sector as a whole, particularly around staffing. The evidence indicates that bringing all the relevant staff together at the same time, either as a whole team or in smaller groups, was a barrier to settings completing all of the training within the expected timeframe of the pilot. Although a particular issue for PVIs—that rarely had opportunities for whole-setting meetings due to staffing and ratio levels—this was found to also be an issue for standalone maintained settings that were not always able to create these opportunities. School-based settings also had difficulties fitting the training into the autumn term (as expected) as they reported that plans for professional development are often made a long time in advance and there a number of competing priorities in the school calendar. Managers in a few settings also reported that the length of each session could be an issue as normal staff meetings were not always long enough to include the training if there were other agenda items. Due to this and other staffing issues, some managers used the shorter training sessions.²⁷ In two settings, all staff had completed the training in their own time. In others, staff who were unable to complete it with everyone else (for example, they were absent on the training days) were asked to do it on their own.

There were also issues caused by changes in staffing during the pilot that impacted on engagement with the training. In settings where staff had been employed after the training, some had been able to do it later (sometimes in small groups). However, some new staff had not been asked to complete the training or had been asked to complete it in their own time; in these cases, it was unclear whether the training facilitators in the settings had—or were able to—check that this had happened.

The setting interviews found some examples of staff who had not done any of the training but were delivering Tales Toolkit. Most, but not all, of these cases were of staff who were new to the setting. In a few instances, staff observed colleagues delivering Tales Toolkit while a couple were planning on doing the training in the coming academic year. However, the managers in three maintained settings (two school-based and one standalone) reported that it was common for some staff, such as teaching or nursery assistants, not to have been asked to do any of the training but still deliver the programme.

The interviews also found that not all settings had completed the training in the required timeframe and, again, time pressures were reported to be the main reason for this. For the maintained school-based settings, the difficulty came from trying to fit the training into the (busy) school timetable. Two reported they could not start the training until January 2024 while another did not run Sessions 4 and 5 until after Easter (two using an INSET day and the other paying for staff cover).

'We were slow to start with Tales Toolkit because of the fact that we needed to do the training ... if it was just the nursery setting, you'd use an INSET day to do that but obviously, as a school, we've got lots of different priorities and tasks. So, in the end, we ended up booking a supply teacher, to cover nursery for the day and then we were all able to do the training' (manager, maintained within a school).

A few settings that were delayed with the training (due to IT/login or staffing issues) reported they were planning to either complete the last training session or start the training again from the beginning of the following academic year (2024/2025). The technical issues reported when trying to access the online training included not having the required logins or finding that they were unable to log in. The Tales Toolkit website was hacked in November 2023. Although a new website was created, settings had to be moved over, which affected access to the training for some staff during this time.

²⁷ The shorter sessions were designed for staff to complete the training on their own and did not include time for the group discussion or exercises. Tales Toolkit no longer offers the shorter sessions on its online platform, although they were available during the pilot.

Impact of delivery method of the training on engagement

RQ8 Does the delivery method of the training (context and mode) have an impact on engagement with the programme (practitioners)?

This section reports on the perceived impact of the delivery method of the training on teachers' engagement. The findings of the evaluation show that it has been mixed.

Sixty-one percent of staff responding to the endline survey (n = 33—across the range of settings) agreed that the online mode (through videos and downloadable resources) was 'better for them' (N = 54). The interviews also found that managers appreciated the flexibility of the online mode, which was easily accessible and could be shared with any number of staff at no extra cost. The structure of the training also meant it could be fitted around their specific context and business-as-usual activities, while the discussion activities and facilitator notes allowed staff to take ownership of their own training.

However, the online and asynchronous mode of delivery—together with the contextual factors faced by the early years sector as a whole around staffing and the expansion of funded provision—potentially offered too much flexibility and led to some non-completion as discussed above. Results from the endline survey show that while 50% of staff (n = 28) had completed all five training sessions, 5% (n = 3) had completed four, 9% (n = 5) had completed three, 16% (n = 9) had completed two, and 7% (n = 4) had completed only one (N = 56). By the time of the follow-up interview (summer 2024), four settings had not delivered all five training sessions.

The Scaling Up of the Intervention section below discusses these issues further in the context of a larger trial.

Tales Toolkit—delivery

RQ6.2 Is the delivery feasible for practitioners to implement?

RQ7 Can settings deliver with fidelity and within routine practice? Is there variation in delivery across settings, for example, embedded into the curriculum, planning, practice, resources, and language or in separate sessions/activities?

This section investigates the feasibility of delivery (RQ6.2) and whether the Tales Toolkit was delivered with fidelity and within routine practice (RQ7). Evidence from case studies and surveys is used to explore how the programme was delivered across different settings, how it fits with business-as-usual practice, any adaptations made, staff experiences of delivery, and their views on the programme.

Overall, most staff who engaged with the endline survey or interviews felt that they were able to engage with delivering Tales Toolkit, implement it in their classroom and that the sessions had fitted into their working day and routine. Most settings could deliver sessions in line with the 'character, setting, problem and solution' structure at least once weekly, for the required time, and within routine practice. However, there was quite a lot of variation in delivery between and within settings (in particular around the Session 4 activities) some of which may have been caused by variations in training uptake (as discussed above). Staff were generally happy with the resources provided and made them available outside of the directed sessions. However, they were less likely to make resources available in outside areas, primarily due to the weather. Any adaptations staff made to support delivery were considered in line with expectations.

Frequency of delivery

It was recommended that the Tales Toolkit sessions should be delivered to all children *at least* once a week; settings were also encouraged to start delivery as soon as they could after the second training session and continue delivery until the end of the summer term (2024). In the endline survey, all staff were asked, 'How often *on average* do you personally deliver Tales Toolkit to *all* the children in your room/class or group?' Table 9 shows the results broken down by job role. Looking at practitioners, as they would be the staff usually expected to deliver Tales Toolkit, 58% delivered at least once a week:²⁸

²⁸ Combining once a week with twice, three times, and more than three times.

- 42% (n = 18) delivered sessions once a week with all children in their group as recommended;²⁹
- 16% (n = 7) delivered sessions more often—twice, three times or more per week;
- 30% (n = 13) delivered sessions less than once a week—once a fortnight or less; and
- 9% (n = 4) never delivered Tales Toolkit.

Table 9: Frequency of delivery with all children in their class or group

Frequency of delivery	Practitioner		Manager		Total	
	N	%	N	%	N	%
Never	4	9.3	7	53.8	11	19.6
Less than once a fortnight	4	9.3	5	38.5	9	16.1
Once a fortnight	9	20.9	1	7.7	10	17.9
Once a week	18	41.9	0	0.0	18	32.1
Twice a week	3	7.0	0	0.0	3	5.4
Three times a week	3	7.0	0	0.0	3	5.4
Over three times a week	1	2.3	0	0.0	1	1.8
Other	1	2.3	0	0.0	1	1.8
Total respondents	43	100	13	100	56	100

Source: evaluation endline survey 2024. It was possible that a practitioner and a manager could both deliver to the same class, meaning the children in that class could receive the intervention more than once a week. No maximum number of sessions was set in the pilot.

Staff who indicated in the endline survey that they delivered sessions more or less than once a week (that is, above or below the minimum dosage for the intervention) were asked to select the reasons why. Reasons for delivering more included:

- delivering three times per week to ensure all children with different attendance patterns could take part;
- because it was linked to ‘group time’, which happened twice per week; and
- because they used it during free-play provision, where children could choose what to interact with.

Of those delivering less frequently than recommended:

- seven were managers or in specialist roles who did not usually teach;
- four reported that they did not have time;
- four shared responsibilities for delivering sessions with a colleague—such as working part time or as part of a job share;
- two had not delivered Tales Toolkit with their whole class or group but were delivering to smaller groups;
- one had only participated in some of the training; and

²⁹ Although we did ask if practitioners deliver to all children in their room/group/class once a week, this did not necessarily mean all children in the group receive the programme once a week due to other factors such as absences, group variations, and other changes.

- four gave an ‘other’ reason—including having one day of teaching time per week, teaching in other areas as well, staffing issues, and additional needs of children in the classroom.

Of the two respondents who delivered to smaller groups rather than the whole class, one delivered to smaller groups twice per week and the other explained that all groups of children had one session once per fortnight. Instructions for the Session 2 task stated that delivering to smaller groups was permitted; however, to remain compliant with the delivery requirements, settings would need to ensure that children in these smaller groups participated in at least one Tales Toolkit session per week.

Of the staff who were not delivering Tales Toolkit sessions with all children in their class/group (N = 11):

- four were managers or in specialist roles who did not usually teach;
- two had only completed part of the training;
- one said they did not have time;
- one only delivered sessions with small groups rather than all the children in the class;
- one was new to the setting and had not done any of the training; and
- two specified ‘other’, one of whom had not accessed any training due to staff issues so had not delivered and the other explained that they had started delivery but ‘stopped due to needing to redirect staffing elsewhere because of issues with extreme behaviour’.

In the endline survey, managers, deputies, and specialist leads were also asked if staff had been able to deliver Tales Toolkit at least once per week, as recommended. Five reported that staff had been able to deliver in full at least once a week, four said that staff had been able to deliver in part at least once a week, and four reported that their staff had not been able to deliver once a week (N = 13).

Findings from the case-study interviews concurred with these survey findings, with most practitioners delivering at least once a week but with the frequency varying between daily and fortnightly. Staff who delivered twice a week reported doing so because of variations in attendance (either 15 or 30 hours per week) to ensure all children were receiving at least one session a week. Variance in child attendance patterns was more common in PVI settings. Staff also reported varying delivery due to class or group size. Practitioners in just under half the settings also reported delivering sessions during free play with at least two settings (one standalone and one within a primary school) delivering this daily, while staff in three other settings (including a mix of setting types) delivered Tales Toolkit in free play if children requested it. As expected, most staff continued delivery until the end of the pilot, however, a few staff reported reducing the frequency of delivery (which was not expected) and one reason given was that they had done most or all of the training or embedded delivery (discussed in the next section).

The follow-up interviews also found some variations in delivery by the age of the children. Some staff working with children in reception were not carrying out the activities associated with session four as expected. Even though the activities were designed to be used with their age group, one teacher thought the children were too old for them as they were generally expected to be writing by the end of reception. In another case, the teacher was running longer writing sessions because they wanted to give the children opportunities to independently practice what they had learnt and fully record the stories they had created:

‘I don’t know how long [Tales Toolkit] recommend, but once you’ve helped the children to write the story, you don’t want them to leave with the idea they haven’t finished. So, I tend to take a lot of time and then depending on the level they’re at, they would record everything, or they would start, and I would finish, but they all record a whole story ... it can take up to 45 minutes. If I’ve got six, I’m working with this one, they think they’re playing, and then I work with the next one, they’ve got a sentence, I always ask them, “Tell me the sentence.” We focus on the sentence together, that we record. So, it takes a lot of time to balance’ (reception teacher, maintained setting within a primary school),

In addition, a few staff working with younger children had also not followed the recommendations for the Session 4 activities because they thought the children were too young. In one setting (maintained standalone), a practitioner working with three- to four-years olds reported that they had waited until the summer term to do this because they did

not think the children were ready before then. Staff in this setting had only done the shorter training sessions, which may explain their confusion over this session. However, another practitioner working with three-year-olds in a maintained school-based setting (who did the fourth training session later in the academic year) realised that it would have been helpful earlier to have done this earlier, as their children were already beginning to draw people by the time staff did the training.

The case-study interviews also found some variation in delivery within settings, with some practitioners delivering the Tales Toolkit less frequently than others. As noted, some practitioners, such as those working part-time or as part of a job share, were personally delivering sessions less often than once per week. A few in their initial interviews reported delivering Tales Toolkit less often due to a lack of confidence; however, by the time of the follow-up interviews, they generally reported feeling more confident in delivery.

Length of delivery

The length of a Tales Toolkit session varies depending on factors such as the size of the class or group and the age of the children.

In the endline survey, all staff were asked: 'Approximately how many hours per week do you (or your staff) spend delivering Tales Toolkit sessions? (Please tick one).' Table 10 shows that most of the respondents reported dedicating around 30 minutes to an hour each week (43%), however, from the survey questions we are unable to determine the number of sessions this time relates to, so we cannot calculate an average session length from this data alone.

Most respondents indicated that they delivered one session per week (see Table 9) so it is tempting to assume that the reported weekly duration reflects a single session, suggesting that most staff generally achieved the expected 30-minute length. However, this assumption cannot be confirmed with certainty as data does not directly capture details of individual sessions. For those who reported spending more than an hour per week, it is likely that they delivered multiple sessions or included related activities. This seems plausible as the attention span of three- to five-year-olds typically does not allow for single sessions of that length. Overall, while the data suggests that delivery durations generally aligned with the developers' recommendations, we cannot state this definitively given the available information.

Table 10: Weekly time spent delivering Tales Toolkit

Time spent each week	N	Percentage
Between 30 minutes and 1 hour	12	22.6
Around 30 minutes	11	20.8
1 to 2 hours	10	18.9
Less than 30 minutes	9	17.0
2 to 3 hours	3	5.7
3 to 4 hours	1	1.9
Other	7	13.2
Total	53	100.0

Source: evaluation endline survey, 2024.

Data from the observed sessions shows that 12 of the 15 observed were 30 minutes or less. In the two cases where they were longer, one (with ten children aged three to four) was 40 minutes because it included a drawing activity (PVI) and the other was 45 minutes (with 11 children aged three to four) and included a mark-making (maintained standalone).

The main reasons given by staff who delivered for less than 30 minutes included the children's age or attention span, time constraints, and other curriculum areas to teach. One early years lead who delivered a session fortnightly said the nursery children had access to the resources at all times. The main reasons given for spending more than 30 minutes delivering Tales Toolkit each week were that staff delivered multiple sessions, including delivering in group time, occurring twice a week, to smaller groups daily, in morning and afternoon attendance groups (in both nursery and reception), and in free play or continuous provision. One of the respondents who took between 30 minutes and an hour was an ECT who needed time to prepare and write up the session afterwards (an expectation of the induction process, not the intervention).³⁰ Thirteen percent (n = 7) of respondents selected an 'other' amount of time each week and reasons given were that they were not delivering yet (because they had not done the training) or that it varied depending on class size.

Embedding delivery—feasibility

Overall, most staff felt that they were able to engage with delivering Tales Toolkit, implement it in their classroom, and that the sessions had fitted into their working day and routine. Findings from the endline survey, in terms of the responses and additional answers to open-text questions, are presented below (the process of embedding is explored in more depth in Embedding Delivery: Practice).

- 76% of staff (n = 42) agreed that they had been able to engage with Tales Toolkit. Staff reported finding the intervention interesting, enjoyable, and fun to deliver; resources are good and easy to use; children engage well; the intervention fits well with others such as Helicopter Stories, Drawing Club, and the 'Share Attention, Respond, Expand, Conversation' (ShREC) approach; and that they enjoy engaging children's imaginations (N = 55).
- 74% of staff (n = 39) agreed that they had been able to implement Tales Toolkit in their class or room, finding it easy to implement, that there had been especially good interaction with boys in particular, some excellent interactions from children with SEND, and good engagement with children overall. Some reflected that having a Tales Toolkit area worked well, using it whenever they could, and supporting Tales Toolkit activities within free play and child-initiated learning worked as well as using it in small groups. However, one staff member was unsure whether all Tales Toolkit activities at their setting were child-led and another noted that children's initial enthusiasm for engaging with the Tales Toolkit resources during play had reduced over the course of the year (N = 53).
- 70% of staff (n = 38) agreed that planning the Tales Toolkit sessions fitted into their working day and routine (N = 54) while 69% (n = 37) agreed that delivering the sessions fitted into their working day and routine (N = 53).

Embedding delivery—the Tales Toolkit sessions

This section presents evidence of Tales Toolkit becoming embedded at settings and how it was embedded. Further information around collaborative storytelling, quality interactions, and child-led learning is provided in the Outcomes section under Changes in Practice and evidence of practitioners using the resources and embedding them in their classroom and other environments is included under Using Resources Outside of Tales Toolkit Sessions in the Outcomes section.

In the endline survey, 62% of staff (n = 34) agreed that over time the Tales Toolkit sessions had become more child-led and 51% (n = 28) agreed that storytelling had become more collaborative (N = 55).

The visits allowed the evaluation team to look in more detail at how settings were delivering key elements and observe the intervention in 'real-life'. The observations showed that all 11 settings were using the Tales Toolkit structure of 'character-setting-problem-solution' in the sessions as well as the 'big bag'. Three also used other physical resources including the apron and the strip in the sessions that were observed (see Resources for more details). Evaluators also observed children in all settings engaging in the Tales Toolkit sessions, such as picking out objects or volunteering ideas.

³⁰ This was not an expectation for Tales Toolkit.

In addition to this, in five settings, children from a mix of ages were engaged in mark-making across a mix of setting types.

Embedding delivery—outside the Tales Toolkit sessions

It was expected that staff would embed Tales Toolkit into their practice and more widely in the setting. The endline survey found that 65% of staff (n = 35) agreed that Tales Toolkit had been embedded into their classroom (for example, it is part of their planning, curriculum, practice, and the language is used in other activities) while 54% (n = 29) felt it had been embedded in the wider setting (for example, as part of the wider planning, curriculum, practice, the language used, and in other activities) (N = 54).

Staff in all 11 settings gave numerous examples of how Tales Toolkit had become embedded outside the directed sessions, including:

- talking about problems and solutions in the context of day to day play and conflict;
- using the language and storytelling aspect of Tales Toolkit in other areas of the curriculum;
- providing opportunities for smaller group work for children who felt less able to speak up in the whole group sessions (so they could fully engage with Tales Toolkit);
- making Tales Toolkit part of their long-term plan to promote language and communication skills;
- holding Tales Toolkit sessions during library time so children could enjoy storytelling in a room with lots of books;
- having story time twice a day which uses the language of Tales Toolkit;
- giving children access to the resources in free play, for example, the 'big bags' for children to play with themselves, character and setting ideas spread around the setting; and
- using Tales Toolkit to discuss things that the children were learning about in other lessons.

'We've been doing about pollution, pollution of the seas, and I was listening to somebody this morning doing one about a turtle getting stuck in a net and that's a problem, how are we going to solve it?' (manager, maintained standalone setting).

Tales Toolkit resources

In the endline survey, 87% of staff (n = 47) agreed that the Tales Toolkit resources were easy to use with 80% (n = 43) agreeing very much (N = 54). Findings from the interviews confirmed this, with most staff finding them easy to use and able to use them as an integral part of their delivery of Tales Toolkit. Staff were pragmatic and creative when engaging with the resources, adapting what they used to children's needs and preferences as well as what was efficient or best suited to the context.

Additional survey findings suggest that the online format was accessible and convenient for most staff, with 78% (n = 42) agreeing that having resources available online was easiest for them (N = 54). The most common benefits cited for hosting resources online were accessibility for all staff, including new staff, so staff could access the resources at different times, and in their own time, and all materials being in one place. Being able to review materials and rewatch videos was also beneficial as this helped with understanding and identifying tips. However, a few surveyed staff commented that it would be nice to have physical resources such as a handbook and one respondent reported that they had printed out the training summaries with prompts for use in the classroom.

In the Tales Toolkit sessions that were observed as part of the setting visits, practitioners used either the Tales Toolkit symbols or the big bags (when the big bags were being used, the separate symbols were not needed because they were on the bags). In one setting, staff started packing the bags for each other's sessions so that the objects would be a surprise for the practitioner as well as the children.

'The reason we started doing it was because some of the staff said they found it hard because they would put objects in with the idea of how the story was going to go, and they found it hard not to try and lead the story. So that's why we started doing it for each other: [it] was so actually there was no preconceived idea

of the story—you were delivering and letting the children lead it. I think that's been really helpful and has worked really well' (practitioner, maintained setting within a school).

The observations also identified staff in two settings using boxes in addition to the big bags. When interviewed, staff said this was because it enabled the children to see in and pick out their own character, setting, problem, and solution.

Some practitioners reported feeling unsure about how or when to use some of the resources, such as the big book. The Tales Toolkit big book is introduced as part of the Session 2 training task, which staff are expected to complete between training Sessions 2 and 3. There are also examples of big books from other schools in the 'Inspiration' section of the downloaded resources. However, as the big book is not introduced in the session handouts, where staff training did not include the activities, this introduction may have been missed and evidence from surveys and interviews showed that very few practitioners engaged with the online resources.

'We've used the hanger, the aprons, the big bags, but [not] the two books, the book side of things, because I don't know exactly how to use them yet, but I just thought naturally it was the next step, so we haven't got to that point yet' (practitioner, maintained setting within a primary school).

Children in reception classes were more likely to use the recommended proformas for recording their stories after a session and it was less common for nursery-aged children to be given proformas. While children in reception will, for the most part, be more confident with mark-making and writing than children aged three to four, mark-making is a part of the curriculum for three- to four-year-olds and the Tales Toolkit resources include pictures and symbols to be used alongside mark-making and suggestions for group activities for recording stories.

Adaptations to Tales Toolkit

During the training sessions, practitioners were encouraged to tailor the Tales Toolkit sessions to the needs of their children, where needed. The adaptations observed in setting visits or reported in the interviews were considered in line with the intended delivery of the programme as the main narrative structure of sessions (character, setting, problem, solution) and practitioner-child interactions have been retained. Therefore, the adaptations observed and reported below are defined as within the scope of delivery as intended rather than outside of it.

The interviews indicated that many settings adapted delivery for children who could not engage. For example, for children with SEND (who could not work in a group setting) their support worker did a version of Tales Toolkit one on one. Adaptions for children with EAL included singing familiar songs, using actions so these children could engage, and including them in groups with children they did not know so that they could pick up new words.

These findings are reflected in the endline survey where 47% (n = 28) of staff agreed that they had made some changes or adaptations to how Tales Toolkit was delivered (N = 55). The most commonly reported change was to change the items in the bag for variety, matching children's interests, or to use more robust items, such as weatherproof material for the bags for use outside. Examples of other changes included adding characters into the bag, using items linked to restorative practice, letting children lead sessions while practitioners facilitate, putting up pictures with Tales Toolkit language, changing delivery location or environment, adaptations to support children with SEND or EAL, and recapping ideas and structure as a class before recording individual stories. Changes such as adding new characters or objects into the bags and changing the location or environment of Tales Toolkit delivery sessions were encouraged by the delivery team as part of embedding Tales Toolkit in the setting and building children's and practitioners' confidence with the sessions and activities.

Barriers and facilitators

RQ9 What are the potential barriers and facilitators of the programme?

This section considers some of the key barriers and facilitators to implementing the Tales Toolkit programme. The main barriers identified were related to the needs and number of the children, and staffing issues (particularly for PVI's). Key facilitators included support (from senior management and colleagues), and the resources provided.

Barriers to implementation

In the endline survey, staff were asked to select which barriers they or their colleagues had experienced when delivering Tales Toolkit. The barriers identified by the highest number of respondents out of a total of 46 valid responses are shown in Table 11. The most common were the high needs of children (54%) followed by staff absences (43%), children's behaviour (35%), and class sizes (24%).

Table 11: Barriers to implementing the Tales Toolkit programme

Type of barrier	Number of respondents	Percentage of respondents (%)
High needs of children	25	54
Staff absences	20	43
Children's behaviour	16	35
Size of classes	11	24
Difficulties recruiting staff	6	13
Difficulties maintaining staff child ratios	6	13
Staff turnover	5	11
Lack of training	1	2
Other	7	15
Total number of responses	97	211

Source: evaluation endline survey, 2024.

Given that respondents could select multiple options, the total number of respondents is more than 100%.

'Other' included no barriers (1), having the time to delivery to multiple groups (1), only delivered to nursery (1), lack of age-appropriate resources (1), not completed training (1), confidence of staff to implement something new (1) and new children joining each term (1).

The interviews and observations also collected information on barriers affecting delivery and the findings confirmed those of the survey: that staffing issues and time pressures were significant. While time pressures were felt across all types of setting, staffing shortages were particularly acute for PVI's.

The main staffing issues reported by settings are summarised below.

- **Staff turnover** could lead to a lack of consistency in delivery. Where staff joined later in the year, they sometimes found themselves delivering the programme without having watched or participated in the training but instead learnt by observing other staff or reviewing materials. Tales Toolkit does not include a cascading element, as all staff are expected to take part in the training, so newer staff did not always have a full understanding of what they were delivering and why, nor were they fully confident in their delivery.
- Similarly, even where staff attrition was not an issue there could also be consistency, quality, and confidence issues where **staff who did not participate in the training**, such as teaching assistants, were delivering the programme. Evidence from interviews indicated that this was primarily due to challenges finding time for staff in these roles to participate in training, such as attending staff meetings, releasing them from one-to-one support with children, or where training occurred after school hours when they were not contracted to work. Setting managers interviewed wanted teaching assistants and staff in specialist roles to be able to deliver Tales Toolkit during free play when a child asked or in the regular group sessions.
- **Management changes** could also affect implementation, especially where there was a change in senior leadership or in staff delivering the intervention. One manager reported they had joined their

setting in September and had immediately been given responsibility for the Tales Toolkit programme, despite having never heard of it and being unsure what it entailed; in another case, a manager took over the early years role as maternity cover and was tasked with setting up Tales Toolkit in their setting. In both cases, the managers wanted to implement the programme because it promoted literacy but had little to no knowledge or context for the programme before they started supporting colleagues with implementing it at their setting.

- **Staff shortages** were cited as the main reason in four of the ten settings that withdrew from the study prior to delivery. Staff interviewed had not wanted to add further pressures to practitioners at their setting by introducing a new programme. Three of these settings said that staffing would continue to be a barrier to participation in the future as they did not think issues would be resolved in the following academic year (2024/2025). The fourth PVI said it would like to implement Tales Toolkit, and it may be possible next year, but it would still be very difficult. One PVI manager fed back that their setting was part of a chain, which was also short-staffed.

Key factors around managing numbers and needs of children at the setting that could affect implementation are summarised below.

- One setting reported they were expanding their provision to two-year-olds, which meant they were more pushed for time than they would be normally to deliver Tales Toolkit sessions. Similarly, a practitioner at a setting that had withdrawn from the study commented that the expansion in funded provision (from April 2024) meant they were at full capacity and so had to focus on the core curriculum rather than adding in programmes. These issues may affect more settings as the free childcare entitlement expands to 30 hours in September 2025.
- Other settings took in children once every term, so there were often new and younger children who needed to be introduced to Tales Toolkit, which led to some disruption in delivery. It is not uncommon for children to join and leave throughout the year, meaning that some children have less exposure to the intervention across the academic year.
- Variations in child attendance patterns, such as mornings only or 15 hours (rather than 30), meant some settings delivered multiple sessions each week at different times in order to ensure that all children received at least one session per week (although this was not always monitored).

Finally, some settings were delivering other communication, language, and literacy (CLL) interventions such as Helicopter Stories, Early Talk Boost, a Tom Sherrington Walk Thru, and phonics programmes.³¹ This may have led them to have less time for Tales Toolkit, however, staff more often felt that Tales Toolkit encouraged their children and had a greater impact on them because it extended their storytelling and understanding of how to structure a story more than other interventions.

Facilitators to implementation

Facilitators to delivering Tales Toolkit were mostly explored through interviews and observations. This section highlights some of the key facilitators.

The biggest overall facilitator of implementing Tales Toolkit for the settings was buy-in from senior managers. Many reported that it fit the ethos of their setting and in some cases that they had been the one to suggest or make the decision to apply to participate in the Tales Toolkit study. This was typically because the programme supported goals they already had for children and staff at their setting.

'When I saw this came up, I was thinking it would tick all my boxes on my school development plan, as improving communication and language, which has been a priority post-Covid, along with PSED' (manager, maintained standalone).

³¹ The Tales Toolkit website states that programmes such as Helicopter Stories can be used alongside it.

Multiple managers said during interviews that they felt Tales Toolkit fitted in well with the curriculum. They also liked that it would encourage children's writing skills and their ability to structure a story, which is something they could see would benefit literacy skills in the years ahead.

Support from colleagues

Several practitioners highlighted the importance of having support from colleagues and working as a whole team. They gave examples of discussions they had during training, of coming up with delivery ideas together, problem-solving and making adaptations for their setting, and observing each other's delivery; in one setting the team created new story props together and would pack the story bags for each other so that practitioners and the children could both be surprised by what came out of the bag and be more creative. Some managers commented that discussion sessions during the training also provided time for less experienced staff to ask questions and learn from more experienced staff. In all interviews, practitioners and managers said they felt they had supportive colleagues who also wanted to work on Tales Toolkit. One maintained standalone setting was part of a federation where several nurseries were taking part in the Tales Toolkit programme: they described how they were able to share thoughts on training and delivery across the federation. These findings suggest that including the activities in the training was a key facilitator for peer support and an important enabler for both learning and embedding delivery in settings.

Findings from the endline survey showed that, in the main, staff were positive about their experience of support from senior staff and peers. The majority (76%, n = 41) agreed that they had received the support they needed from their managers to deliver the programme and from their colleagues (N = 54).

Resources

Some staff, including managers and practitioners, said the resources, especially the physical props, made it easy and straightforward to deliver Tales Toolkit with little preparation. Managers who acted as training facilitators reported that they appreciated having downloadable resources readily available too, as this meant the training was quick for them to set up and deliver with their staff. This was helpful as the training model for Tales Toolkit specified that the training sessions would run alongside delivery of Tales Toolkit sessions for the first few months and later training sessions introduced new activities and resources.

'Because it's so open ended and because there's so little prep, that definitely is an enabler. Once you've done the training, that's the hardest thing to get done, once you've actually done that, implementing it is super easy' (manager, maintained setting within a school).

Support

What support is needed and how can a useful support model be developed?

This section explores the support made available as part of the Tales Toolkit programme to practitioners and settings, how practitioners engaged with this, what further support might be needed, and how the support model could be developed further. Respondents to the endline survey were reasonably satisfied with the support they received from the Tales Toolkit team although few had taken up the support available. The interviews confirmed this with low uptake of the support offered but many staff expressed a desire for more support from delivery team, particularly in feedback of how they were progressing. The Tales Toolkit team confirmed that few staff had contacted them or participated in the support sessions offered to settings.

Alongside the training videos with discussion prompts, Tales Toolkit provided a package of online materials on its members' website to support staff. These included downloadable resources, recorded webinars with experts, inspirational articles from schools, further research, and a Facebook support group where ideas could be shared between settings. Setting staff could also contact the Tales Toolkit team via email or online chat with any queries or support needs. The endline survey found that 53% (n = 28) of respondents agreed very much that they had received the delivery support they needed from the Tales Toolkit team (N = 53).³² Further feedback given in a follow-up open-text

³² 8% of respondents indicated 'somewhat agree', 2% indicated 'not at all', and 38% responded 'N/A or don't know' to this question, indicating that they did not request support from the Tales Toolkit team.

question included that support from the delivery team had been timely and helpful, the Facebook group (where used) had been useful, and that resources and guidance had been very helpful, so they had not needed further support (N = 12). Staff in nine of the 11 settings reported engaging with the online resources: four settings had at least one member of staff who reported that the online resources were useful and a helpful place to go for advice, three settings had at least one member of staff who had watched at least one webinar, one maintained standalone setting said it had accessed information about Tales Toolkit on Dr Julian Grenier's website,³³ and another promoted use of the online resources in its weekly newsletter to staff, including them as professional development.

However, most staff interviewed at case-study settings were unaware there was a Facebook group and often said they did not have time to look at the additional online resources. This is reflected in the endline survey findings where 57% (n = 30) reported having accessed the members' page and agreed that it was useful but 34% (n = 18) answered 'not applicable/don't know', indicating they had not accessed online resources (N = 53). It may have been the case that only the lead practitioner or training facilitator was aware of the Tales Toolkit member's page and Facebook support group and that other staff may have been unaware of this online support. As discussed above, the Tales Toolkit website was hacked during the delivery period, which possibly also affected participants' ability to access the online information.

Many staff in settings commented during interviews that they would have liked to have received some feedback from the delivery team so they could be sure that they were delivering the programme correctly. This was not something that the delivery team proactively offered as part of their support package, but they also did not receive any requests of this nature from staff at the participating settings. In addition, the delivery team confirmed that they had received very few requests, overall, for support from participating staff and settings during the pilot, which suggests that there was a mismatch where some staff were not aware of what was available on the member's page and Facebook group so felt that they would have benefitted from support from the delivery team but did not make contact to request this support.

Models of support are discussed further in *Scaling up the Intervention*.

Overall feasibility

- Lower-than-expected levels of engagement with the training may have affected fidelity: not all staff who were delivering Tales Toolkit had completed all the training and not all settings had completed the training within the recommended timeline or used the full version of the training videos with all staff together.
- The on-demand video training format and online resources were convenient, but their asynchronous nature may have affected completion.
- Single logins per setting made compliance tracking difficult and they were not always shared with all staff (as expected).
- Tales Toolkit sessions were not always delivered weekly.
- All settings were observed using the Tales Toolkit structure of character, setting, problem, and solution, however, delivery of the sessions varied within and between settings.
- Practitioners found the resources easy to use but use of the resources varied (especially outside of directing the sessions).
- There was evidence of adaptations for children who had SEND or EAL.
- Barriers to training and delivery included staff issues and time pressures as well as large class sizes and high needs of some children (especially with the ongoing impacts of the Covid-19 pandemic).
- Enablers to delivery included buy-in from senior staff, support from colleagues, and participating in the training directly rather than learning from observing colleagues.
- Practitioners often reported that they were delivering Tales Toolkit in the way they understood it but were not sure if this was what was intended (and would have liked some feedback from the delivery team).

³³ Dr Julian Grenier is an expert in early years education and Senior Content and Engagement Manager (Early Years) at the EEF.

- Time pressures reported by many settings may have affected their ability to engage, especially for settings facing severe pressures such as staffing issues.
- Asynchronous training may have meant settings had a more distant relationship with the delivery team (compared to face to face or live online training).
- Few staff were aware of the member's area or the Facebook support group where they could contact Tales Toolkit or were unsure of what support they could get (possibly due to lead practitioners not sharing this information), which suggests a more proactive model of support could help practitioners to deliver Tales Toolkit confidently and ensure that the core elements are delivered correctly. The issue around logins could be addressed by ensuring it is clear to facilitators and lead contacts that they are expected to share these with all staff. The delivery team could build a check-in or reminder about this into their contact with settings.
- Although staff in many settings said they would have liked to reach out to Tales Toolkit with their questions, the delivery team indicated they had not received many questions, and few settings had attended the drop-ins.

Evidence of promise

This section seeks to explore whether the pilot intervention showed evidence of promise and supported the theory of change. It reports on the following research questions.

- RQ3** Do practitioners have confidence and motivation to deliver the intervention?
- RQ1** Does the programme lead to a change in practice (as per the short-term outcomes in the ToC)?
- RQ2** Do practitioners perceive the intervention has led (or will lead) to positive outcomes for children (for example, their language and communication skills, PSED, creativity, and writing), especially for children in receipt of EYPP or with EAL, and SEND?
- RQ5** How do pupils engage with the different activities and aspects of the programme?
- RQ4** Are there any unintended consequences or negative effects—for example, pushing out other programmes, activities widening the attainment gap between disadvantaged and non-disadvantaged pupils, or a perceived burden?

Practitioners' confidence and motivation

RQ3 Do practitioners have confidence and motivation to deliver the intervention?

Findings from the surveys and interviews indicate that confidence and motivation among staff to deliver Tales Toolkit improved during the intervention. Eighty-three percent (n = 45) of those who responded to the endline survey felt motivated to deliver the programme; 80% (n = 43) felt confident (N = 54). The interviews suggested that confidence levels varied by how much of the training staff had completed and how often they had delivered Tales Toolkit. As expected, confident practitioners were those who felt that the training had prepared them for delivery and some managers concurred that the training helped practitioners develop their confidence. Variations in delivery also affected confidence levels, for example, where some staff did not deliver regularly (in the case of part-time staff or job shares) or where they needed to adapt delivery to meet the needs of their class in terms of age or size.

Evidence from follow-up interviews indicated that confidence in delivery had increased over time as practitioners completed more training and gained more experience of delivery. Peer-to-peer learning also improved confidence levels and practitioners gained from observing other staff delivering and having a confident practitioner who took the lead. Staff in two PVI settings reported that a lead practitioner, confident in delivering Tales Toolkit, supported colleagues in delivery, including modelling sessions for them (particularly if they had not done all of the training). In one PVI, this was seen as

a positive experience for all staff, including the lead practitioner, whose confidence also grew as a result of being observed by colleagues:

'It's like anything you do, with the more practice you naturally feel more confident with doing it and so I've had a few more people come in and watch me, since we've spoken last, a colleague in April, I've had a lot more people come in and kind of observe me delivering it, so I guess that's made my confidence grow a little bit' (lead practitioner, nursery and reception).

However, in the other setting, the lead practitioner reported this less as a positive and due to the fact that some colleagues had not done all the training.

Supporting language development

Supporting language development is a core aim of the programme and in the interviews, staff reported that they had gained more knowledge, tools, and confidence in supporting children's language and communication development including, in some cases, supporting storytelling and creativity skills. This was attributed to the training, including the structure of the group sessions, which enabled the incorporation of good practice. Practitioners reported that the programme had given them more ideas for language development, the ability to support the language skills of children with SEND, and a knowledge of how to encourage children to tell stories and develop their skills in language and literacy:

'So, I think for me, also it's given me more resources to use, more knowledge about, you know, how to sort of engage children in telling stories and developing their story language and their literacy skills' (practitioner, standalone maintained nursery).

Some practitioners reported that they had become more confident at letting children lead and tell stories independently. Several described how their confidence had grown over time as the programme brought out the children's imaginations and capacity for coming up with ideas (which the practitioners were sometimes surprised by) as well as seeing the effectiveness of child-led learning once they had practiced it. Other staff felt they needed more practice to develop this confidence.

In the endline survey, staff were asked to reflect on any changes in their confidence and knowledge around supporting language development since the beginning of the Tales Toolkit study: 80% of respondents (n = 32) agreed that their confidence in supporting language development among children aged three years and above had improved with 48% (n = 19) agreeing 'very much' (N = 40).³⁴ Just one, a headteacher with over ten years' experience working with children in the early years sector, indicated that their confidence had not increased at all, which is perhaps not that surprising given their length of service. Staff were also asked to what extent their knowledge of language development had improved since beginning Tales Toolkit: 76% (n = 25) agreed this had been the case; 33% (n = 11) agreed 'very much' (N = 33). Two respondents reported that their knowledge had not increased at all, one was the headteacher mentioned in the previous point and the other was another early years practitioner with over ten years' experience.

Longitudinal cohort—changes in confidence

Staff in both baseline and endline survey were also asked how confident they felt in their knowledge and skills supporting children in a range of different areas of development.³⁵ Results presented in this section are for the areas of development which saw the largest increases in staff confidence between the two timepoints (based on changes in the mean). These results match some of the main outcomes of the Tales Toolkit intervention for staff: supporting children with EYPP to make good progress, supporting collaborative thinking and play, supporting children with SEND to make good progress in language and communication, and supporting children with creative thinking.

- Supporting children receiving Early Years Pupil Premium to make good progress in their language and communication skills: the average mean score (on a five-point scale) increased from 4.0 to 4.38. At

³⁴ This section combines all the 'positive' answers: 'slightly', 'somewhat', 'fairly', and 'very much'.

³⁵ 'How confident are you in your knowledge and skill at each of the following?' Scale used: 1, 'not at all'; 2, 'slightly'; 3, 'somewhat'; 4, 'fairly'; 5, 'very much'; 6, 'not applicable/don't know'. The final option (6) was included in the endline survey but not the baseline survey.

baseline, 28% (n = 10) of staff reported being very confident (N = 36), compared to 53% (n = 19) at endline (N = 36; DK/NA = 4).

- Supporting children with collaborative thinking and play with adults and other children: the average mean score increased from 4.14 to 4.49. At baseline, 33% (n = 12) of staff reported being very confident (N = 36), compared to 67% (n = 24) at endline (N = 36; DK/NA = 1).
- Supporting children with SEND to make good progress in their language/communication skills: the average mean score increased from 3.72 to 4.06. At baseline, 14% (n = 5) reported being very confident (N = 36), compared to 28% (n = 10) at endline (N = 36, DK/NA = 3).
- Supporting children with creative thinking and problem-solving skills: the average mean score increased from 4.14 to 4.43. At baseline, 33% (n = 12) of staff reported being very confident (N = 36), compared to 64% (n = 23) at endline (N = 36; DK/NA = 1).

Fuller analysis including questions within the 'supporting language development' dimension where there was a smaller change (differences in means of 0.28 or less) are reported in Appendix K Table 1.

Changes in practice

RQ1 Does the programme lead to a change in practice (as per the short-term outcomes in the ToC)?

This section investigates evidence for change in practitioners' practice, including whether there is evidence of collaborative storytelling, quality interactions, and child-led learnings (RQ1a). It also explores whether there is evidence of practitioners using the resources in their classrooms or other learning environments, such as outside, and whether there are any differences by setting type (RQ1b). Findings from the case studies and surveys suggest that most staff felt that participating in the programme did result in changes to their practice.

The endline survey reflected this: 72% of those surveyed (n = 39) agreed that participating had resulted in changes or improvements to their regular practice (N = 54). Similarly, interviews and observations of Tales Toolkit sessions found evidence of practitioners across all types of settings engaging in quality interactions, collaborative storytelling, child-led learning and strategies to support children's development. Several staff at three settings explicitly attributed the changes in their practice to the training, which gave staff opportunities to consolidate and reflect on good practice, as well as increasing knowledge and confidence in supporting children's development (see Practitioners' Confidence and Motivation to Deliver the Tales Toolkit). The following sections provide more detail on the types of changes reported.

Quality interactions

In the endline survey, 74% of staff agreed (n = 40) that delivering Tales Toolkit had led to more quality interactions between staff and children (N = 54). Staff reported in the interviews that, due to the balance of theory and practice in the training and, where this happened, the opportunities to reflect on practice in group discussions, those who took part reported developing a good understanding of the benefits of quality interactions and child-led learning. Evidence from the observations in settings corroborated these findings, with examples of staff engaging in quality interactions during Tales Toolkit sessions such as waiting for children to respond, making eye contact, and using open-ended questions. However, a few staff did report having difficulties implementing some approaches due to them not attending all the training, in addition to staff shortages and working with children of mixed ages or abilities. An example of one such setting is provided below.

Observation: quality interactions

One PVI setting was experiencing severe staffing shortages where staff were unable to leave the room to speak to colleagues or answer the door even in order to maintain staff-child ratios. This appeared to have an impact on the ability of the practitioner to engage in quality interactions when delivering Tales Toolkit. This practitioner had taken part in the training, watching the training videos in groups alongside other staff at the setting. The practitioner was observed delivering a session to a group of ten two- to four-year-olds, with one additional member of staff to support. The researcher noted the practitioner did not use open questions or 'I wonder' statements and left little time for children to respond. It was also not always clear to the researcher whether children were understanding the language being used. The researcher reflected that the practitioner appeared to be finding it difficult trying to support children of different age and need levels, and with multiple activities taking place before and during the session they were likely fatigued and overstretched. The manager at this setting also said that it took some time for practitioners to build their confidence and were unsure about how to deliver and use resources initially. (PVI)

Longitudinal cohort—changes in practice

Staff in both baseline and endline surveys were also asked how often they engaged in different behaviours to facilitate quality interactions with children. Results presented in this section are for the areas of development which saw the largest increases in frequency between the two timepoints (based on changes in the mean).³⁶

The largest changes practitioners reported in this area centred on modelling a wide variety of words and questions, imitating the child's language and behaviour, and having extended back and forth interactions with children. These elements of interaction are all focused on building up language, either through taking what the child has said and expanding on it through modelling further language or through building it into a longer conversation or interaction.

- Modelling a wide variety of words, including adjectives and connectives: the average mean score (on a 5-point scale) increased from 4.33 to 4.75. At baseline, 47% (n = 17) of staff frequently did this (N = 36), compared to 81% (n = 29) at endline (N = 36).
- Using a wide range of questions: the average mean score increased from 4.14 to 4.53. At baseline, 44% (n = 16) of staff frequently did this (N = 36), compared to 69% (n = 25) at endline (N = 36).
- Imitating what the child has said or done: the average mean score increased from 4.06 to 4.42. At baseline, 39% (n = 14) of staff frequently did this (N = 36), compared to 69% (n = 25) at endline (N = 36).
- Having extended back and forth interactions with children, allowing the child time to listen, process and reply: the average mean score increased from 4.44 to 4.78. At baseline, 47% (n = 17) of staff frequently did this (N = 36), compared to 83% (n = 30) at endline (N = 36).

Fuller analysis including questions within the 'quality interactions' dimension where there was a smaller change (differences in means of less than 0.3 or no change) are reported in Appendix K Table 2.

Creativity and collaborative storytelling

Evidence from the observations showed staff engaging in collaborative storytelling such as encouraging children's story development and using open-ended questions. Evaluators observed staff making efforts to engage all children in the group in collaborative storytelling, even where sessions were practitioner-led, by offering one to one attention to

³⁶ 'Please could you rate the extent to which you do the following things when normally interacting with a child or group.' Scale used: 1, 'not at all'; 2 'slightly'; 3, 'somewhat'; 4, 'fairly'; 5, 'frequently'. 6, 'not applicable/don't know' set as missing for analysis. N varies due to these responses at endline only.

individual children, drawing out children who tended to speak less, and asking children to join in with actions, noises, and songs. An example from an observed session is summarised below.

Observation: collaborative storytelling

An observation was carried out of a reception teacher delivering Tales Toolkit to their class (four- to five-year-olds). This practitioner had taken part in all five training sessions, watching the training videos in groups alongside other staff at the setting. Staff had split the reception class into two groups because some children struggled to engage in whole-group settings without an adult supporting. 'I think the smaller groups are so much better; but what things work better for us is actually having an adult there ... so they can sit with them.' There were 17 children in the observed class, all with mixed confidence levels and some with English as an Additional Language. Although the session was practitioner-led, they encouraged active participation in the story and used elements of humour to engage the children and add to the storytelling approach. The story involved a dragon losing his teeth when he lands in a swamp, and the practitioner asked the children to make 'whooshing' noises and mime actions. The practitioner asked different children in the group for their input, asking a few different children each question, and drawing on less vocal ones, collating a range of responses and different ideas. The solution was a magnifying glass and magic wand which were used to find the teeth, and the practitioner asked the children to repeat, 'Abracadabra, teeth come back.' In the interview, the teacher said that she likes that the SEN children can now use their imagination with their objects, such as using a block as a baby. She did admit to struggling with getting the quieter children involved as a few children in the class are very loud. But she thinks the sessions give the quieter children the opportunity to talk about how to share ideas and now these children are putting their hands up because they know she sometimes picks children to share their ideas. *Maintained nursery within a school*

Longitudinal cohort—changes in creativity and collaborative storytelling

Staff in both baseline and endline survey were also asked how often they engaged in different behaviours to facilitate creativity, play and storytelling with children.³⁷ Results presented in this section are for the areas of development which saw the largest increases between the two timepoints (based on changes in the mean).

The largest differences in behaviour focused on supporting children to engage with sustained sharing thinking, providing access to materials to make marks or tell their own stories and helping children who find it difficult to join in pretend play or storytelling. Another change in practice suggested that Tales Toolkit has facilitated practitioners to practise dialogic reading with children, when children are encouraged to talk about what is happening and give their own ideas and helping children to extend their ideas through sustained discussion.

- Supporting children to engage with sustained sharing thinking: the average mean score (on a five-point scale) increased from 4.03 to 4.66. At baseline, 31% (n = 11) of staff frequently did this (N = 36), compared to 67% (n = 24) at endline (N = 36, DK/NA = 1).
- Providing access to apps or hand-made books where children could make marks or add pictures or videos to tell their own stories: the average mean score increased from 3.17 to 3.76. At baseline, 19% (n = 7) of staff frequently did this (N = 36), compared to 33% (n = 12) at endline (N = 36; DK/NA = 2).
- Helping children who found it difficult to join in pretend play or storytelling: the average mean score increased from 4.08 to 4.56. At baseline, 25% (n = 9) of staff frequently did this (N = 36), compared to 58% (n = 21) at endline (N = 36).
- Discussing children's ideas and responses when reading books (dialogic reading), encouraged children to talk about what is happening and give their own ideas: the average mean score increased

³⁷ 'Please could you rate the extent to which you do the following things when normally interacting with a child, group or in your classroom, answering as honestly as you can.' Scale used: 1, 'not at all'; 2, 'slightly'; 3, 'somewhat'; 4, 'a fair amount'; and 5, 'very much'. 6, 'not applicable/don't know' set as missing for analysis. N varies due to these responses at endline only.

from 4.11 to 4.58. At baseline, 33% (n = 12) of staff frequently did this (N = 36), compared to 64% (n = 23) at endline (N = 36).

- Helping children to extend their ideas or narrative through sustained discussion: the average mean score increased from 4.17 to 4.58. At baseline, 31% (n = 11) of staff frequently did this (N = 36), compared to 69% (n = 25) at endline (N = 36).

Fuller analysis including questions within the ‘creativity and collaborative storytelling’ dimension, where there was a smaller change (differences in means of less than 0.42), are available in Appendix K Table 3.

Supporting children’s personal, social, and emotional development

In the interviews, staff reported that delivering Tales Toolkit provided opportunities for practitioners to support children’s personal, social, and emotional development (PSED)—in particular by encouraging turn-taking, listening, and sharing. They also reported that using the language of Tales Toolkit offered opportunities for problem-solving and enabled staff and children to use and develop this language. Opportunities for turn-taking and sharing were also seen in the Tales Toolkit sessions observed in the setting visits.

A commonly reported change in practice among staff across all types of settings was using the language of Tales Toolkit (character, setting, problem, solution) when interacting with children, within and outside of the sessions, especially for problem solving. Practitioners described modelling the language in real-life problem-solving scenarios (for example, a problem arising between two children) as well as in storytelling. This is explored further in Perceived Child Outcomes—PSED. One manager at a PVI explained during a follow-up interview how the programme had extended their existing practice around problem solving:

‘I think it’s the knowledge that we’re really good at problem-solving during play. So, [saying] “How do you think we could do that?”, “Why do you think that happened?” But actually, stories aren’t something that we did before. So actually we have put that problem-solving into stories ... it’s given us another tool to use’ (manager, PVI).

Other examples of practice supporting PSED, from interviews and observations, included ensuring children take turns to choose an object, which, in larger groups, sometimes meant inviting a child who was ‘sitting nicely’ to take their turn. Where sessions were more child-led (and children were older), practitioners also encouraged children to take turns in telling and listening to each other’s stories. An example from an observed session is given below.

Observation—supporting PSED

In an observation of a session delivered with six reception children (two girls and four boys), the practitioner encouraged turn-taking, asking each child to pick an object for a problem and a solution out of the box and take it in turns to explain why it was a problem or solution, encouraging them to listen to each saying, ‘Did you hear [child’s] amazing story?’ The practitioner then asked children questions about the problems and solutions to extend their thinking. One problem was a bouncing frog who was jumping on a nearby snail and the solution that the child identified was a seashell that the snail could hide under. The practitioner asked this child, ‘Does frog mean to hurt snail, or did he just not notice because the snail is so small?’ This practitioner had taken part in Sessions 1 to 4 of the training alongside other staff prior to the observation visit. In the interview, they said ‘I had big surprises with some of the quietest boys, suddenly, there were a lot of stories in those heads.’ *Maintained nursery within a school*

Longitudinal cohort—changes in supporting PSED

Staff in both baseline and endline surveys were asked how often they engaged in different behaviours to facilitate children's PSED. Results presented in this section are for the areas of development which saw the largest increases in frequency between the two timepoints (based on changes in the mean).³⁸

The findings reflect the central structure of Tales Toolkit sessions where stories are based around finding solutions to problems: talking through how to resolve problems, using stories with challenges to talk to children about feelings, encouraging children to talk about a problem, and teaching children ways to solve conflicts.

- Talking through how they, you, and others have resolved problems or difficulties: the average mean score (on a five-point scale) increased from 4.22 to 4.74. At baseline, 42% (n = 15) of staff frequently did this (N = 36), compared to 81% (n = 29) at endline (N = 36, DK/NA = 1).
- Using dialogical story time to discuss books that deal with challenges, explaining the characters' feelings and how they solved those problems: the average mean score increased from 3.81 to 4.26. At baseline, 25% (n = 9) of staff frequently did this (N = 36), compared to 50% (n = 18) at endline (N = 36, DK/NA = 1).
- Encouraging children to talk about a problem together and come up with ideas for how to solve it: the average mean score increased from 4.25 to 4.61. At baseline, 44% (n = 16) of staff frequently did this (N = 36), compared to 69% (n = 25) at endline (N = 36, DK/NA = 1).
- Teaching children ways of solving conflicts such as modelling how to listen and agree a compromise: the average mean score increased from 4.44 to 4.75. At baseline, 58% (n = 21) of staff frequently did this (N = 36), compared to 83% (n = 30) at endline (N = 36).

Questions within the 'supporting children's personal, social and emotional development' dimension where there was a smaller change in the response (differences in means of less than 0.31) are available in Appendix K Table 4.

Using resources outside of Tales Toolkit sessions

RQ1.2 Is there evidence of practitioners using the resources in their classrooms or other learning environments (for example, in an outside play area, other shared areas in the setting)? How do they use them, and does this vary by type of setting (for example, PVI versus maintained settings)?

This section investigates whether there was evidence of practitioners using the resources in their classrooms or other learning environments as outlined in the training, such as outside (RQ1.2). The findings show that practitioners were utilising the resources outside of the sessions, however, only a minority of practitioners had made the resources available

In the endline survey, 89% of staff (n = 49) reported that Tales Toolkit resources were freely available for children to engage with outside of the staff-led sessions: 47% (n = 26) indicated that this was the case both inside and outdoors at their setting, 42% (n = 23) that they were available indoors only, and 11% (n = 6) that resources were not available outside of the sessions (N = 55).

Observations found that six of the 11 settings had a Tales Toolkit corner set up with items such as story bags, boxes, and symbols so children could play with the resources independently (outside of the adult-led sessions). One maintained setting within a primary school had a corner set up for their reception class but not for their nursery room because they were worried that children might run off with the resources, damage them, or get them dirty. They felt that their nursery children had not yet understood the link between the bags and the structure of a story so were not ready to use them independently. The Tales Toolkit corners that were observed generally had the big bags hanging up at child height. Other settings did not have a dedicated corner but had the big bags hanging up for children to use independently. In one

³⁸ 'Please could you rate the extent to which you do the following things when normally interacting with a child or group, answering as honestly as you can.' Scale used: 1, 'not at all'; 2, 'slightly'; 3, 'somewhat'; 4, 'a fair amount'; and 5, 'frequently'. 6, 'not applicable/don't know' set as missing for analysis. N varies due to these responses at endline only. Base (N) varies due to missing data at endline.

setting, the apron and the strip were accessible for children to use independently. However, other settings had not reached the point of using these items (at the time of the visit):

'Normally [the big bags] are just, like, hanging on the hooks in the classroom, in my classroom, which means that sometimes the children can get them out and start to fill them up' (practitioner, maintained setting within a primary school).

In other settings, however, practitioners put the big bags away after the session had finished. In a couple, the resources were stored in-between sessions because only one member of staff was trained or confident enough to deliver the programme. In all three PVI settings, the resources were not available for children to play with outside of the sessions: one manager told us this was so they could deliver sessions at a specific time.

In terms of having the resources in outside areas, all the settings had wanted to offer this, however, bad weather during the spring meant that only one (maintained standalone) setting was able to do so. They thought it worked well, despite initial concerns that the resources would get dirty or damaged when outside. This was the concern for other settings who had not yet set up Tales Toolkit outside.

Among the settings, around half were observed to have resources such as the big bags and the strips available for children to play with independently inside the classroom. One setting had a Tales Toolkit area set up outside, and another couple reported that their children were able to play with the resources, such as the apron and laminated symbols, outside.

Most of the settings that did not make resources available outside of Tales Toolkit sessions were delivering only to three-year-olds, including all three PVIs, whereas two of the three settings delivering to reception (four to five years) had made the resources available (see Resources for more information).

Children's engagement

RQ5 How do pupils engage with the different activities and aspects of the programme?

Staff reported in the survey and interviews that most children were able to engage with Tales Toolkit and this was confirmed in the observed sessions with children particularly enjoying the opportunities to interact and direct the story-making process using the big bag. However, there were some examples of observed sessions where not all children were engaged, for example, in large reception classes. In terms of children's independent engagement with the resources (outside of the staff-led sessions), although the endline survey found that staff thought this was happening, this was not often observed in the setting visit.

In the endline survey, 82% (n = 45) of staff agreed that children were engaging with the Tales Toolkit sessions (N = 55). This was corroborated in the sessions observed during the visits where practitioners would sit on the floor with the children in a circle and use the big bag and quality interactions to encourage children to think about the character, setting, problem, and solution. It was reported and observed that children especially liked using the big bags because of the excitement of finding out what was in them:

'I think the resources do really help that engagement, like you saw today ... them getting the excitement of opening the big bag and what's going to be in there and pulling something out. So, I think that definitely helped the engagement, and we found that in reception as well, just that sort of awe and wonder around knowing what's going to be in there, that's a magical thing, in itself, isn't it? And, so, that's really helped pull them in, especially, with the those who might find that concentration and attention a bit more difficult' (manager, maintained setting within primary school).

In observations, children also clearly enjoyed:

- being able to pass the items in the bag around in a circle;
- pulling items out of the big bags;
- suggesting ideas about what the character's name could be, what the character liked to do, and how the problem would make the character feel;

- practitioners using their ideas; and
- telling the story back to the practitioner at when asked.

Reception age children also enjoyed being able to write or record their story after they had told it to a practitioner.

However, some of the younger children seemed less sure of the meaning of the 'setting' and 'solution' symbols and needed to be prompted by a practitioner to remember what it was. After the observed sessions, some practitioners reflected that this was because these concepts were more abstract and took longer for children to understand. One practitioner also noted that asking about the setting could sometimes take children out of the story, as they were already focused on the character aspect. During one observation, when the practitioner showed the children the solution bag, the children were less sure what this was, so to remind them, the practitioner asked the children why there was a smiley face on the bag and a child said it was because it 'makes everything happy' (maintained standalone setting).

Most staff felt it was easier to get children to engage in Tales Toolkit when group sizes were smaller. During observations in settings, sessions were generally delivered in groups ranging from six to 15; however, one setting was delivering to a reception class of 30 and another to a group of three children. Staff sometimes found it difficult to keep children engaged in sessions, especially where group sizes were bigger. Having an additional adult was beneficial in these situations to support the key worker by instructing the children and asking them questions (to keep them engaged). Where sessions were delivered with smaller groups, practitioners sometimes strategically selected which children joined the group, although, where these groups took place in the main room, evaluators observed that if a child came over and wanted to take part as well, the practitioner would let them join. Practitioners explained that this selective approach was often based on the confidence levels of the children; for example, they might include less confident children in one group, such as children who had EAL or who were quiet, so that they could take time to explain their ideas and have their thoughts heard. In some observations, children were very enthusiastic about the storytelling and shouted out a solution as soon as the problem came out of the big bag. In these instances, practitioners encouraged the children to 'slow down' and 'listen to other ideas first'.

Evaluators did observe some sessions where another group of children was doing different activities in the same room, which created some distractions for those in the session. The observations also pick up on some instances where at least one child was distracted in the session or drawing unrelated pictures during mark-making sessions. In one case, the practitioner interviewed afterwards commented that sometimes children were not in a 'listening mood'.

In the endline survey, 62% of staff (n = 34) reported that children at their setting engaged with Tales Toolkit resources available beyond the staff-led sessions (N = 55). However, in the settings where resources were made available for children to use independently, staff interviewed reported that although children enjoyed being able to use them, it often took some time using the resources in the directed sessions for children to want to do this.

Perceived child outcomes

RQ2 Do practitioners perceive the intervention has led (or will lead) to positive outcomes for children (for example, their language and communication skills, PSED, creativity, and writing), especially for children in receipt of EYPP or with EAL and SEND?

The following section explores this question.

Language development

The evaluation data indicated perceived positive outcomes for children in language development: 81% of staff responding to the endline survey (n = 43) agreed that Tales Toolkit had a positive impact on the communication, literacy, and language of children at their setting and that it benefitted the development of the most disadvantaged children (N = 53).

The setting interviews confirmed these findings, with staff reporting observing improvements in children's language and communication skills, including improved vocabulary development, and children becoming more verbal or vocal. Staff particularly noted this outcome for children with SEND and EAL (as well as younger children at earlier stages of development). The early years lead at one of the settings with a high proportion of children with SEND reported that Tales Toolkit benefitted children with SEND, EAL, and in receipt of EYPP more than other approaches they had tried.

Managers in some settings also identified distinct benefits for EYPP children: one reported that their EYPP children in particular were doing very well while the manager in another suggested that it might be helpful for EYPP children to be exposed to the language of Tales Toolkit (although they were not able to provide examples of how it had benefited this group specifically in their setting).

Perceived positive outcomes were often attributed to practitioners language-modelling and introducing children to words they would not usually use during play or in other contexts. At a setting where about a third of the children had SEND (with high behavioural or speech and language needs, which were attributed by staff to COVID-19), one practitioner described Tales Toolkit as a 'godsend' for children's language and communication skills. They gave an example of a SEND child who was pre-verbal being exposed to more vocabulary due to the intervention:

'What we saw at the beginning is the vocabulary has improved. Even by just giving key words and the adult just modelling key words and just describing what we were doing and saying, like today, with [child] "Oh it's cold. It's cold!" He's listening to you, so the vocabulary has increased. They're connecting ideas better and we're finding that in their general play as well' (nursery teacher, maintained nursery and reception).

Staff also reported that the clear and familiar structure of the sessions gave children more confidence and reassurance to speak, leading to improvements in vocabulary, language skills, and clearer speech. Listening to other children in the sessions was also seen as beneficial, including for those children receiving speech therapy. A practitioner at a PVI described a child in their group (who was attending speech therapy) who had benefitted from Tales Toolkit, especially group modelling by their peers.

'I've got a little boy who's going to speech therapy at the minute and it's really helped him with his talking. Like using new words, hearing his friends use them and us talking about it ... I think it's the creativeness of letting them create their own stories, letting them have free rein on what they want to do and doing it together with his friends, it's helping him build confidence' (practitioner, PVI).

Several staff interviewed reported children with EAL specifically benefitting in terms of expanding their vocabulary. This was explained by the group modelling, adaptations that could be made to help engagement, and the use of real objects (as opposed to picture books).

'I think it's really benefited the EAL children because you've got real objects ... so, it's really helping them to learn those words, as well because they can see it all happening in front of them. Whereas, in a book they might switch off a bit because it would be too much of an overload ... they might not see the picture clearly and there's often quite a lot in a picture' (practitioner, standalone maintained nursery).

This practitioner also reported that they were able to add in higher-level language over time as the children learned the new words. For example, they might have a purple blanket as a setting, but over time it would become a shimmering, shiny, purple blanket.

Staff also reported benefits for those who participate in sessions with children of different abilities or ages as they learn from each other. In the case studies, some observed sessions were with such mixed groups. Staff explained that the decision to mix groups was due to the importance of children modelling communication skills to each other, as this had a greater impact than practitioner modelling.

'So definitely the children with higher levels of development have really, really responded well to it and you need some of those modelling to some other children. There are some other children that it's just gone above their head a little bit. But it certainly has promoted the kind of talk about characters and settings that probably wouldn't have happened until [they reached] the older age group, which is quite nice' (manager, standalone maintained nursery).

A few staff also noted that quieter children (who may not have had SEND or EAL) were becoming more vocal and confident to speak during the group sessions. This increased confidence in engaging can also be interpreted as a PSED outcome.

'Some of our quieter children, it's definitely brought them out a little bit more ... just got them just saying some normal words or more words, it's definitely helped' (manager, PVI).

Creativity and storytelling

Staff reported perceived positive outcomes around creativity and storytelling in the follow-up interviews. This was particularly evident in older children and often attributed to older children having a better understanding of story structure and sequencing. One commented that some of the older three-year-olds were inventing their own stories while the younger ones repeat the stories from the sessions. Another, in a different setting, reported that the older children would expand the key words into a sentence and story. A reception teacher in another setting reported that the children were now more likely to tell a story with a plot, rather than just list characters and objects, which they were doing previously. Additionally, the manager at a PVI setting reported that Tales Toolkit readily supported the creativity skills of their four-year-olds, although felt the three-year-olds needed some more input to build creativity.

'The four-year-olds have been amazingly creative and inventive. Our three-year-olds need a little bit more scaffolding and they pick up similar patterns and sort of do it, just on a lower level. But our four-year-olds, the ability for them to be creative and imaginative, Tales Toolkit supports that' (manager, PVI).

Practitioners also noted that older children were doing more storytelling independently outside of group sessions, using the language of character, setting, problem, and solution.

Some staff also noted at follow-up how Tales Toolkit improved children's confidence in their own ideas and stories. In one nursery, the practitioner felt that less confident children benefitted in particular because they could see that there was no right or wrong answer. Staff in some settings reported that Tales Toolkit had positive impacts on children with already well-developed language skills. The manager of one setting in a less disadvantaged area (PVI) found that the intervention still benefited their cohort, improving their creativity and imagination skills by extending their learning.

A quarter of the settings reported delivering other storytelling programmes, such as Helicopter Stories, alongside Tales Toolkit. Staff in these settings reported that children's storytelling had improved and perceived this was due to Tales Toolkit (for example, one practitioner explained that children doing Tales Toolkit were giving their stories more structure than children who were taking part in other programmes). However, the limitation of potential interactions between the programmes in settings delivering other similar programmes is acknowledged.

Writing

In the interviews, many practitioners felt it was too early to detect any effects on mark-making and writing, usually due to the stage of delivery they were at (although in some of these cases, the Session 4 training had only been done recently). However, a few expected that the improvements in creativity and storytelling would translate into mark-making and improved writing in the future. At a follow-up interview, one nursery teacher delivering to three-year-olds explained how they felt the intervention would benefit children in terms of writing, going into reception, and beyond.

'It's been good, and I think it's a really good structure and as you go on into schools, they're able to tell, write their stories because they've already got that structure' (nursery teacher, standalone maintained nursery).

At the setting visit, a practitioner in another setting (delivering to three- and four-year-olds) reported that the children were doing more writing and mark-making than before and they '[could not] stop mark-marking and writing' in afternoons after taking part in the Tales Toolkit session in the morning. They also reported that boys in particular were doing more marking and writing than they would have expected.

'A lot of boys are doing more, like marking and writing. Particularly boys that I didn't think would be doing it' (practitioner, standalone maintained nursery)

Personal, social, and emotional development

In the follow-up interviews, most practitioners reported perceiving improvements across all children with respect to social skills, turn-taking, and sharing. They also noted that progress in one aspect of development can often lead to benefits in another area, for example, increased confidence can lead to increased engagement with conversations or the use of more complex vocabulary.

Practitioners highlighted the benefits of children understanding and using the Tales Toolkit language of problem and solution when encouraged by a practitioner and reported children increasingly using this language independently in real-life scenarios. A reception teacher gave an example of a child who had hurt themselves in the playground, then saying, 'We have a problem, and now we need a solution!' In the follow-up interviews, a nursery teacher added that Tales Toolkit helpfully teaches children that there can be more than one solution to a problem, while a practitioner in another nursery reported that the language was used often and that 'encouraging children to think of their own solutions to problems really helps their confidence'.

Some managers also noted that children who were already confident benefited from the intervention. The deputy head in one school reported at follow-up that all children benefited from Tales Toolkit, as learning to listen to other children is important for skills such as turn-taking, patience, and collaboration.

'I think all children benefit, because even the children who are very confident storytellers benefit, because they have to listen to other children's ideas too. And that's as important in terms of sharing ideas and waiting for a turn and listening and being able to stay patient, and to collaborate with the group. And just hearing a good story, the story doesn't go in the direction you would have taken it, but you can understand that you can make your own story that goes in this direction. I think it's something that every child would gain from' (deputy head, maintained nursery and reception).

The findings were less specific for children with SEND, EAL, or EYPP with regards to PSED outcomes. A few practitioners perceived specific PSED outcomes for their children with SEND. One practitioner at a setting with a high proportion of SEND children said they saw improvements in children's social skills including turn-taking, respecting others' ideas, compromising and problem solving together.

'When we've got the little social groups, you're seeing that they are looking at each other and they're responding better. So, they're understanding it's a two-way conversation. That if I have a turn, I've got to wait' (nursery teacher, maintained nursery and reception).

Unintended consequences or negative effects

RQ4 Are there any unintended consequences or negative effects—for example, pushing out other programmes, activities widening the attainment gap between disadvantaged and non-disadvantaged pupils, or a perceived burden?

In the endline survey, 4% of staff (n = 2) agreed (slightly or very much) that taking part in Tales Toolkit had some negative effects while 80% (n = 43) reported no negative effects (N = 54). One reporting a negative effect worked with children in reception and felt that children may have done less writing than expected as they were focusing more on telling stories. In interviews, staff did not identify any negative consequences, however, in settings which were struggling with staff shortages, finding time for all staff to take part in the training could be challenging and some staff at these settings reported using their free time to do training.

Similarly, staff participating in interviews did not specifically identify any unintended consequences of delivering the programme. However, staff at a few settings where children had different attendance patterns reported delivering sessions more frequently to children with 30 funded hours (which may risk creating or widening inequalities). For example, one practitioner scheduled Tales Toolkit sessions twice a week and delivered with whichever children were at the setting at the time so children on 30 hours would have two sessions per week. At a couple of these settings, staff noted that they perceived stronger positive outcomes in the children who had taken part in more Tales Toolkit sessions. A manager at a maintained nursery reported that the children at nursery all day get two sessions a day (as they deliver in the morning and afternoon), and they said they had noticed more impact with the 30-hour children, saying how these children would go and act out the story afterwards. At a setting where delivery had been more consistent in the morning (due to staffing issues affecting afternoons), the nursery teacher noted seeing more impacts with the morning children, including improvements in language and communication, imagination, and social skills (nursery with about a third of children with SEND or behavioural needs).

While 61% of staff responding to the endline survey (n = 33, ranging from 'slightly' to 'very much') felt that there had been some effects of delivering Tales Toolkit that they did not expect (N = 54), all the open-text responses were positive and centred on expected outcomes of the programme, such as children engaging with the sessions and telling their own

stories, articulating problems or emotions, and improvements in language use and confidence. Some wider benefits that were described included quieter children ‘finding their voice’ and boys engaging well with the programme.

Readiness for trial

This section explores considerations for future scale-up of the Tales Toolkit intervention and evaluation as part of a trial, including contextual factors which may affect successful implementation, any changes suggested, and costs to settings and affordability. Research questions covered are:

- RQ11** Do the training and resources constitute a standard intervention that can be replicated and scaled up?
- RQ12** What, if anything, should be updated or changed for a trial evaluation (for example, ToC, materials, delivery method)?
- RQ13** Are there any key contextual factors that appear to facilitate or impede successful implementation at a larger scale, for example, setting types?
- RQ14** What will the cost of the intervention be to settings? Are the costs considered affordable? What is the willingness of settings to pay for the programme and commit the time, beyond the pilot? How does the cost and affordability vary by type of setting?

Scaling up the intervention

- RQ11 Do the training and resources constitute a standard intervention that can be replicated and scaled up?**
- RQ13 Are there any key contextual factors that appear to facilitate or impede successful implementation at a larger scale, for example, setting types?**

Whether the training and resources constitute a standard intervention that can be replicated and scaled up (RQ11) is first considered, including whether there are any key contextual factors that may facilitate or impede successful implementation at a larger scale (RQ13).

For the most part, the Tales Toolkit design was well-structured and fairly standardised. The five training videos were pre-recorded and available online on-demand so all settings could use and access these along with instructions for accompanying activities to run within the group participating in the training. Each setting received a set of Tales Toolkit physical resources, such as story bags, an apron, and a record book, as well as access to online resources such as webinars, blogs, and other materials. The Tales Toolkit sessions are well structured and the story structure of ‘character’, ‘setting’, ‘problem’, and ‘solution’ supports practitioners to embed the Tales Toolkit approach to engaging with children and problem-solving into their wider ways of working. In addition to this, practitioners can access ongoing support from the Tales Toolkit team via email, chat, or drop-in sessions and join a practitioner network with online meetups (although the pilot found that not many took up this offer). The online mode of training delivery and online resources hub would be scalable to a larger number of settings and feedback from the delivery team indicated that they had capacity for scaling up the number of kits to send to settings (as long as they had sufficient time to prepare). However, the evaluation identified some compliance issues regarding how staff and settings engaged with the training that may have affected the extent to which the intervention was implemented in a standardised way.

Training and resources

In the current study, half of the practitioners had completed the five training sessions over the course of the year. As some elements of the programme are introduced in later training sessions, not all settings were delivering all elements of the programme, such as mark-making, for example, which is introduced in Session 4. Also, the feedback from practitioners indicated that the discussion activities during the longer training sessions were beneficial for planning and embedding the Tales Toolkit, as well as for communal learning among staff, so there is a concern that staff participating in the shorter training sessions, without these activities, could be at a disadvantage. Some issues were also raised around accessing the resources and support that the delivery team made available. There are several options the delivery team could consider to facilitate compliance with the training (and therefore fidelity in delivery), but these would have implications for resourcing and therefore scale up.

Considerations for the delivery team

- The delivery team should consider increasing or clarifying messaging and information around the required timings and core, mandatory elements, such as in a protocol/guidance document, so that facilitators and practitioners are clear as to what is necessary and by what timeline in order to deliver Tales Toolkit effectively. In particular, emphasising the importance of completing all of the training by a certain time. The research found examples of settings running the Sessions 4 and 5 quite late, primarily due to finding the time to get all staff together but also partially due to what they thought would be appropriate for the children. Some training and delivery expectations were set out in the MoU but as settings started signing MoUs several months before the pilot delivery phase started, some may have forgotten these requirements. The training documents, such as the pre-training information sheet, do specify the timings but they are also open to interpretation (for example, at least a month between Sessions 3 and 4). It may be necessary to include a cut-off for when all or some of the training (Sessions 1 to 4) needs to have been completed to allow sufficient time for settings to use the resources (and to show mark-marking outcomes).
- The delivery team should also consider adopting a more proactive approach to supporting settings in order to check that they are engaging with all of the training sessions, delivering appropriately, and to assist with any queries. Feedback from the team indicated that they had capacity for offering further support; however, this was rarely accessed. The delivery team could monitor access to and engagement with the training more consistently through developments to the members area of the online Tales Toolkit platform: ideally, this could be used to monitor things like viewing of training videos and accessing of resources by individual setting accounts. Unfortunately, over half of the settings in the pilot reported challenges around accessing the members area, some of which were due to the website hack (which were resolved), however, there were also issues with setting leads not sharing the log-in details with all staff (as expected), which meant staff other than the facilitator could not always access the training or resources online. The delivery team currently implements end-of-unit questionnaires for training participants and has suggested that it could increase expectations around completing these to monitor training completion and the use of resources and additional support. This may be strengthened by building in regular check in points throughout the first few months of delivery, which the delivery team could specify as mandatory in the MoU. A more proactive approach would increase costs and resources needed for the delivery team when scaling up delivery.
- The delivery team should consider providing more targeted support for those in the lead practitioner or training facilitator role at each setting to ensure that they understand the training requirements and to encourage use of the team's role as advisor and troubleshooter. The delivery team could support settings who cannot get all staff together and find an alternative approach. Targeted support could be a separate training course for training facilitators or there could be support calls at the start and throughout the training and delivery period to monitor and support training and delivery. The approach would have implications for resource and cost for the delivery team and would be an additional engagement requirement for settings.
- Changing the format of the training should be considered. Moving the training delivery method from asynchronous online delivery to either synchronous online delivery, where training sessions are delivered online at set times, or in-person training could provide a firmer structure in terms of timing of training and consistency in delivery and allow for direct monitoring of attendance. However, there would be practical and cost implications to both approaches. Synchronous online training would require settings to make staff available at set times and would require staff and technical resource for the delivery team. In-person training would lead to greater costs for both the delivery team and the settings, such as venue hire and travel; it would also take longer if travel is required and may be more difficult for staff to attend or for settings to be able to provide cover for their absence. Cost and resource requirements would increase significantly during for scale-up if an in-person training model was used.
- Consider designating the first three training sessions as the essential core that all settings need to complete—as these explain quality interactions and how to deliver Tales Toolkit sessions—and then later sessions could be additional, depending on the age group of the children: for example, the mark-making component of the programme might not be considered mandatory for both younger and older children

(some staff thought these children were too young for mark-making while some reception teachers thought their children were too old). This could then be communicated to settings in training guidance so that settings short of time or delivering to certain age groups know which elements of the training are required for them. The endline survey found that more staff reported completing Session 4 than Session 3 and so had missed a core session (and had not necessarily watching them in chronological order as might be expected), perhaps where they had been absent and not caught up. A short pilot may be helpful to assess a core and options training offer.

- Providing top-up training sessions throughout the year to introduce staff to Tales Toolkit would be helpful in cases where staff had joined a setting or a class or group during the course of the year. This could be especially useful for supporting PVI settings as those in the pilot were more acutely affected by staff turnover than the maintained standalone nurseries and maintained nurseries within a school. This approach would have cost and resource implications for the delivery team. Alternately, the delivery team could provide advice and guidance on how to support and incorporate staff who join during the year.

Delivery

Evidence from the evaluation suggests that staff found it straightforward to deliver Tales Toolkit sessions using the character, setting, problem, solution structure and that children were using this language outside of settings too but not all settings were delivering to all children at least once per week as required and not all settings made resources available to children outside of the sessions. Some of these issues could be addressed via one of the support or guidance options suggested in the training section above, or the delivery team could identify some elements of delivery as core and some as optional. A short pilot may be helpful to assess the core and options for the delivery offer. However, there were additional factors affecting consistent delivery which could be addressed to ensure standardised delivery.

Differences in child attendance patterns led to some settings delivering Tales Toolkit sessions several times per week to ensure that all relevant children took part in at least one session per week. However, this could mean that children attending for more hours were taking part in more sessions relative to their peers and, therefore, receiving multiple 'doses' of the intervention. In addition, different staff delivering the sessions at different times could mean that some classes or groups did more than one session. Although this was acceptable in the pilot (where staff were expected to deliver a minimum of one session per week) this could be an issue at scale-up. Child attendance patterns varied across all types of settings but PVI settings in our study were more likely to be offering wraparound provision, for example, from 8 a.m. to 6 p.m., so may have wider variance in attendance patterns. This may not be an issue for fidelity, but the delivery team may wish to consider whether to set an upper limit on sessions per week to ensure a more equal dosage across all settings, better enabling comparison of outcomes. However, interviews with staff at settings that had adopted this delivery strategy suggested that it would be challenging to track participation for each participating child and add an additional burden (when most settings were already struggling with time pressures). Any future trial would need to measure compliance, so tracking child-level participation would be crucial for measuring child outcomes. For future evaluations, one option might involve settings committing to monitoring (and building this into MoUs and recruitment documents). Alternatively, the intervention could be delivered to children in reception class as they have more structured days with registers taken at the beginning of the morning and afternoon sessions, enabling better tracking of participation. Guidance around children joining throughout the year would also assist with consistent delivery.

Also, some settings were using Tales Toolkit alongside other language interventions such as Helicopter Stories. The Tales Toolkit website states that the programme can be used alongside other interventions. However, in an efficacy trial this may need to be monitored or prohibited.

Contextual factors

While the delivery model of Tales Toolkit did not involve direct engagement from the delivery team after the initial set-up phase when settings received resources and access to the website, they did engage with settings to administer funding for staff cover where relevant. The delivery team noted some difficulties engaging and communicating with settings, which was time consuming, and would have implications for recruitment and continued engagement of settings with the programme at scale-up. They found it particularly challenging to engage with PVI settings. In the pilot, settings of all types reported time pressures but staff shortages were particularly acute among the PVIs. Seven of the ten settings that withdrew from programme delivery over the course of the pilot were PVIs and all withdrew because of a staffing issues. This suggests that recruiting and retaining PVI settings may be a challenge. If they were to be included in future trials,

a larger geographical area would need to be sampled from and include a large reserve sample to account for withdrawals. A further challenge in the recruitment of settings reported by the delivery team was difficulty in recruiting settings in disadvantaged areas. Only seven recruited settings and four retained settings in the final sample were based in an area with a deprived IDACI score (and only in Decile 4). This limited the extent to which it was possible to explore how Tales Toolkit worked for disadvantaged children. This could be addressed in a future efficacy trial by focusing recruitment in areas of highest deprivation (using IDACI) and/or widening the geographical area of the intervention.

The delivery team also reported how issues in the summer term of 2023 led to a relatively short recruitment phase for the pilot. For some pilot settings, this meant that they were unable to start Tales Toolkit training early in the following academic year as they had planned their professional development and INSET days in advance, so were unable to schedule the training until later in the academic year (in some cases, not until the January or April). A longer window for recruitment should address this, together with the early identification of a lead practitioner to maintain consistency in communication (and provision for situations where that may change, including a secondary contact).

One further consideration for a future efficacy trial is that the free childcare entitlement expands to 30 hours from September 2025, which may increase challenges for settings on staffing, maintaining staff-child ratios, and other resources. In turn, this may create a barrier to engaging in interventions, trials, and CPD more generally as settings adjust to the impacts of this, especially for PVI settings, which often reported experiencing staffing issues in the current study.

Suggested changes to the programme

RQ12 What, if anything, should be updated or changed for a trial evaluation (for example, ToC, materials, delivery method)?

This section investigates any changes that could be made for a future trial evaluation; this includes changes to the Tales Toolkit materials and the theory of change. The previous section presents options for changes to the delivery model to facilitate consistency of delivery and possible implications for scale-up. Overall, feedback on the training and resources was positive, however, staff interviewed at settings had some suggestions for how the training and resources could be further improved.

Resources

The majority of staff interviewed felt that the training was useful. They commented that they had found the current example videos really helpful. Nonetheless, several suggested that it would be helpful to have further video examples of Tales Toolkit being delivered with different groups and contexts, such as groups of different sizes, younger (pre-school) age groups, mixed-age groups, children with different abilities or needs, different classroom environments, or parents delivering the programme. In terms of delivery, some practitioners delivered the sessions with smaller groups (of six to ten) rather than the whole class or group (15 to 30) as they struggled to keep a large group of children engaged. Several suggested that having advice or ideas on adapting the programme for nursery-aged children or children who have EAL or SEND would be useful, and a few practitioners would have liked advice for supporting parents, such as delivering in their home language. Practitioners were able to access support from the Tales Toolkit delivery team via emails or chat, but the delivery team reported that few practitioners did this and no examples were found of staff doing this in the case studies. However, several practitioners reported that they would have liked the opportunity to get feedback from Tales Toolkit to ensure they were delivering the programme correctly. During interviews, the delivery team reported that they had sufficient capacity to provide more ad hoc support. In light of this, the delivery team could consider new ways to communicate the support available and suggest to settings the option of sending in videos of delivery for feedback, although this could require considerable resources at a large scale and there are also safeguarding and GDPR considerations to consider. The delivery team could consider using existing modes of communication, such as email or the Facebook group, more frequently to check-in and remind settings what support is available, or they could try other approaches such as WhatsApp groups or adding an FAQ section to the online resource hub.

The majority of staff were also positive about the delivery resources but there were some suggestions for changes to the physical resources. In some instances, these were changes that the delivery team were in the process of making or were planning to make when interviewed towards the end of the research period. These included:

- changing the Velcro on the bag to a zip or popper to make it easier to open for some children;

- making the apron bigger for reception children to use and adding pockets; the delivery team was in the process of doing this at the time of writing;
- adding activity sheets to the kit and sheets that explain what the different resources included are as it is not always clear to practitioners how to use them; the delivery team commented that it might also be helpful to include a note about the training in case the emails get lost;
- laminating materials to make them more durable;
- using a colour other than white for the fabric materials as it showed dirt quickly, especially when being used outside;
- settings being able to put recordings of practitioners delivering sessions on their online platform in order to encourage parental engagement with modelling Tales Toolkit language at home for parents to see; and
- possibly reviewing the design of the 'character' symbol as some children interpreted it as a gingerbread man.

Theory of change

Updates could also be made to the theory of change to inform a future evaluation of the programme in an efficacy or other randomised controlled trial. These might include adding more detail to specify what delivery should look like and which elements are core or mandatory, for example, whether it is mandatory for practitioners to complete all five training sessions within a fixed timeline to be compliant, whether practitioners are expected to undertake mark-making activities with children of all ages, or whether this is required only for younger children. It may be helpful to create an activities section between the inputs and outputs section to specify this. For a trial context, it would also be important to identify the key outcomes for children's CLL and PSED, and areas where the delivery team would expect to see the most meaningful change or impact. For example, staff in settings reported CLL outcomes around vocabulary, length of sentence, and narrative language; PSED outcomes included problem-solving and turn-taking. It may not be necessary to go to precisely this level of detail in the theory of change model, but it will be useful to identify key areas or aspects in order to inform how a future trial would measure impact. Recent changes to the Early Years Foundation Stage (EYFS) would also need to be accounted for in any new design. Currently, the three prime areas are language and communication, PSED, and physical development. Further detail around key areas or aspects of practitioner behaviour that would be expected to change would facilitate the implementation and process evaluation to investigate whether the programme is working as expected and how.

Costs and affordability

RQ14 What will the cost of the intervention be to settings? Are the costs considered affordable? What is the willingness of settings to pay for the programme and commit the time, beyond the pilot? How does the cost and affordability vary by type of setting?

Most settings reported no intervention costs other than cover for staff training—and this was funded as part of the pilot—however, although some of the costs for cover could be claimed back at the end of the year, one setting suggested that this could cause cashflow issues as they set their budget at the beginning of the year so the cover constituted an unanticipated spend. Most settings incorporated training into staff meetings, so they did not require cover. One school did have to pay for training cover and another had to fund TA hours to attend the training outside of their normal school hours. Some schools and PVI's reported associated training costs—for example, offering time off in lieu for training has an impact on staffing and using staff meeting time for training meant other planned activities for these meetings were lost. As discussed in the Feasibility section, staff surveyed agreed that delivering Tales Toolkit fits into their working day and routine. Similarly, interview feedback suggests that staff did not find preparing and delivering sessions particularly burdensome.

A couple of settings identified additional costs specifically around the resources, including the cost of materials and practitioner time to make additional resources. One setting reported needing to be creative with the resources as they could not afford to buy more. These findings were reflected in the endline survey, where all staff were asked whether the cost of delivery was affordable. Of managers, 73% (n = 8) agreed (a fair amount and very much) that it was affordable; none thought it was not affordable and 27% (n = 3) answered 'not applicable/don't know' (N = 11). For

practitioners, 44% (n = 17) agreed (somewhat, a fair amount, and very much) that it was affordable to deliver; 1 (3%) said it was not affordable: 'There is no budget—we took part in a funded trial.'³⁹ It was not possible to meaningfully explore differences by setting type due to the high proportion of respondents selecting 'not applicable/don't know'. However, staff were asked to explain their answers and the responses from those who said it was (very much) affordable were mainly around the ease of using resources they had anyway (7): 'We use existing resources we have around', 'Props can be inexpensive items that, connected with imaginative minds, produce great results.' Although one did qualify that by saying, 'We have the resources, it's just staffing that is an issue at the moment.' Teachers who said costs were affordable 'a fair amount' said:

'All training costs are hard to find in schools at the moment. The cost of the training and being able to revisit it plus how fantastically organised and presented it is, gives huge value for money. This is especially true as it also covers high quality interactions and physical development training elements too.'

'Costs are linked to training. You need to train all staff working with children for it to have highest impact.'

In the follow-up interviews towards the end of the pilot, most managers and headteachers reported that they thought the programme was affordable and value for money (and would recommend it to others) as they could see the positive outcomes from it. The early years lead in one school-based setting (that had needed to pay for supply) said the costs had been worth it, 'yes, we could have made it work within their budget', while the manager in a standalone said:

'In terms of value for money and, like we said, the impact of just being able to have all the staff together and do it all together, hearing it and seeing the proper videos, rather than getting it second hand [from] someone else who's been out on the course, I think it's really worth the money, even if we had paid for it. The impact has outweighed what the cost would have been I would say for us' (manager, maintained standalone).

However, a couple of managers commented that they may not have invested in buying Tales Toolkit if they had not known about the positive impacts. The manager of a PVI, who felt that their setting could have afforded to buy Tales Toolkit, might not have done so before the pilot:

'So we're really fortunate here, we don't have a budget. So would we have been able to afford it? Yes. Would I have chosen to spend that money on it? Probably not. But doing it now, I would. It's a lot of money but actually I think the impact it's had on the children has been massive and it's been worth it. But I don't know if I'd have, I'd have said, "Yes, it's definitely worth it, let's go for that"' (manager, PVI).

The headteacher at a maintained standalone setting who thought it was affordable and worthwhile suggested that a 20-minute introductory video that people could watch before buying the programme which explained the process and included some examples from nurseries about the impact might help to convince settings of the benefits upfront. They explained that they had not previously understood how Tales Toolkit differed from other interventions such as Helicopter Stories.

'We could have afforded it with some fundraising, but because we were using Tales Toolkit symbols with Helicopter Stories, we probably wouldn't have bought it, because we thought they were doing it anyway. In hindsight, we would definitely have bought it. We would have liked an introductory video to view without buying it—20 minutes explaining the process with some examples from nurseries about the impact it has' (headteacher, maintained standalone).

The managers in a school-based setting reported that receiving the payments towards staff backfill costs that were provided as part of the pilot earlier in the pilot year would have been better for them financially.

'We accessed the additional funding for staff, but this came at the end of the year. It would have made a difference to have had this upfront, (early years lead, maintained school-based).

³⁹ 54% (n = 21) said 'not applicable/don't know' (N = 39) as practitioners were not necessarily expected to comment on the affordability of the programme.

The managers in three of the maintained settings (two standalone and one school-based) said that although they would have liked to, they would not be able to afford the market cost of the programme (£800 excluding VAT) on current budgets. One manager reported they would have liked to buy more kits (£100 without VAT) and resubscribe to the online resources, however, they would not be able to afford them, even at the discount provided to settings involved in the pilot (£200). The Tales Toolkit website states that the costs for the average setting work out to £100 per staff member and £5 per child.⁴⁰ The delivery team felt that PVI settings specifically would not have signed up without funding. This could, in part, be due to differences in funding models between maintained and PVI settings: some PVI settings may be part of a charity or other organisation that part-funds their delivery whereas other settings will need to cover their costs through charging fees. There was an increase in nursery closures as an impact of the Covid-19 pandemic and the high rate of closures has continued since with many nurseries struggling with staffing costs.⁴¹

In the endline survey, 13 managers or deputy managers (including specialist roles) were asked whether they would have found the full cost of Tales Toolkit affordable: the majority (11) indicated that they would not. All eight managers or deputy managers from PVI settings indicated that these costs were not affordable, two managers or deputy managers from standalone nurseries considered them to be affordable, but two others said they were unaffordable, and one specialist from a nursery within a school also indicated that the costs were not affordable. Reflections from the managers and deputy managers surveyed included that budgets were very tight at the moment, including for releasing staff. Now knowing the benefits of Tales Toolkit, managers said it was worth investing in but they were not fully aware of these benefits before participating in the study, and the cost would have affected their decision if they had to invest upfront. This suggests the cost of the programme would need to be funded for a future trial, and it may be a barrier to some settings adopting the approach outside of a funded study. Unless, as suggested by the setting managers involved in this pilot, the process and the positive impacts of Tales Toolkit could be made clear to them.

Evaluation feasibility

This section explores the feasibility of engaging practitioners in research activities (RQ16) and collecting child outcomes data (RQ15 and RQ17).

Engaging practitioners in research

RQ16 Is it feasible to involve practitioners in research activities? In particular, can their support be secured for the collection of primary outcome data and for depth interviews (for example, due to time-burden)?

As reported under Barriers and Facilitators, settings faced issues around staffing and time pressures. These were particularly acute within PVI settings. This may have implications for practitioners' and settings' capacity to engage with and support evaluation activities. Challenges around staff turnover and the need to maintain staff-child ratios mean that particular thought and care needs to be taken around evaluation activities to ensure that these do not become too burdensome for settings. Interviews with staff from four of the ten settings that withdrew from programme delivery showed that staffing issues or burden on staff were key reasons for ceasing. At one, the manager explained that they had struggled with such staffing issues since the expansion of the free childcare hours in April 2024. It is possible that the increase to 30 hours in September 2025 may create further pressure on settings. The majority that withdrew were PVIs (despite comprising 45% of the sample recruited to the pilot) suggesting that this setting type is most at risk of withdrawing from future evaluations.

In contrast to the challenges described above, overall participation in the study interviews and surveys was good. The study aimed to include twelve case studies: a sample of 11 was achieved with a balance of around one-third from each of the three setting types. The baseline staff survey received 156 responses across 30 settings and the endline survey received 56 responses across 21 settings. This is an average of five respondents per setting at baseline and three at endline. A fall in responses between baseline and endline staff surveys is not unusual for studies in education⁴² but the decrease for Tales Toolkit was larger than anticipated (64%). This can be partly attributed to the attrition in settings from 30 to 20 (33%) over the course of the evaluation but was also affected by staffing issues (including illness, turnover, and

⁴⁰ <https://www.talestoolkit.com/pricing> (accessed 19/01/2025)

⁴¹ <https://ndna.org.uk/news/nurseries-continue-to-close-as-government-funded-childcare-expands/> (accessed: 19/01/2025)

⁴² For example, the REAL Programme (Speight et al, 2022) saw a 29% decrease in responses between baseline and endline surveys with early years practitioners (52 to 37 respondents) and attrition in settings from 53 to 43 over the course of the trial.

recruitment), the timing (coming up to the summer holidays), and possibly other evaluation activities overlapping with the endline survey such as the follow-up interviews (although these were used to encourage completion) and the delivery team contacting settings to administer refunds for staff cover. The fieldwork period for the endline survey was extended but there was limited further uptake as some settings were finishing provision for the summer. Many of these factors can be accounted for in future research design to reduce the risk of attrition between surveys. Overall, these findings suggest that it is possible to engage practitioners in early years settings in research activities relevant to implementation and process evaluation, although PVIs may find this more difficult due to staffing levels.

Child outcome data and measures

RQ15 Are changes in key outcomes (for example, communication and language) measurable?

RQ17 Would it be feasible to collect child outcome data? What outcome measure(s) might be used?

Practitioners reported perceived outcomes for children participating in Tales Toolkit covering CLL and PSED. These included improved skills in vocabulary, creativity and storytelling, mark-making, problem-solving, behaviour, and language. Practitioners perceived different effects with the different year groups, pre-school or reception, because of their different stages of language development. The current pilot study explored children aged three to five in reception and pre-school years but it may make sense for a future trial to focus on one year group. For example, focusing on children in reception would enable an evaluation to capture a fuller range of outcomes, such as mark-making and writing, more complex language in storytelling, and communication. The findings around dosage (with some children receiving the intervention more than their peers) are also a concern if child outcomes are to be measured. The fact that some settings (but particularly PVIs) were unable to monitor frequency of delivery (and therefore dosage) is an additional reason why a future trial may need to work with reception classes as they have explicit recording systems (such as class registers) and structured timetables.

The following sections explore the feasibility of collecting child outcome data with consideration of the burden on settings and staff, and the outcome measures that could be used. In doing so, the evaluation team acknowledges that while the Early Learning Goals set out in the Early Years Foundation Stage (EYFS) statutory framework assess children's communication and language development and PSED at age two to three and at the end of reception, these are high-level benchmarks and have not been standardised as an assessment tool. A more detailed assessment may allow evaluators to capture changes in specific skills or behaviours, which may provide more detail on how the intervention achieves impact (assuming it does this). For these reasons, a future efficacy trial would in likelihood need to collect data directly from children in early years settings.

Feasibility of collecting child outcome data

Child outcome data collection in settings would either require practitioners to deliver assessments or external assessors to visit settings to deliver tests—or possibly a mix of both. Direct data collection from children would also require that relevant information around data protection and withdrawal is shared with parents/caregivers of participating children. The current pilot study has demonstrated that all types of settings were able to communicate and share relevant documents with parents. Providing child data to evaluators may create some additional burden for setting staff, although school settings may have administrative staff who can assist with this process.

Assessments by setting practitioners

First, the issues for practitioners in settings delivering assessments are considered. While it is common for early years settings to monitor or track children's progress using a package such as the WellComm or EYFS statutory framework, assessments with individual children may take place over a number of weeks, which may not be appropriate for evaluation timelines. Moreover, different settings may use different monitoring frameworks, however, some local authorities mandate the use of a particular approach to monitor outcomes for children in their area. Timing of assessment and consistency of assessment tools could be scoped as part of area and setting selection. Evaluators may also wish to provide a briefing or training to practitioners and other setting staff to ensure that all data necessary for the evaluation is being collected consistently. Similarly, if practitioners were being asked to deliver an assessment that was new to them—for example, a specialist language test—they would need to participate in training on how to deliver this. Feedback from the pilot study showed that finding time for staff to attend training could be challenging for all types of

settings and training for a specialist test would most likely be delivered by an external expert either at a venue or online. This approach would increase cost and burden.

Assessments by external evaluators

In the case where assessments are delivered by external evaluators, practitioners would need to liaise with the appointed external assessors to arrange for them to come into settings to test children. This would also require finding a suitable space for this purpose—ideally a separate room or quiet space, such as a staffroom, but some settings, in particular PVI, may not have a suitable room. All setting types will likely have wide variation in children's attendance patterns, such as mornings only, half a week, 15 hours, 30 hours, and others, however, PVI settings may be more likely to offer wraparound care, for example, 8 a.m. to 6 p.m. so children may be attending at very different times during the day or week. This may necessitate assessments taking place across several days to test a sufficient number of children. Also, most PVI settings in the pilot struggled with staffing issues, which can mean that assessment visits are cancelled on days when staff-child ratios are particularly tight. For these reasons, accommodating external assessors may be particularly challenging for practitioners at PVI settings compared to those at schools. School settings may also be more familiar with participating in research studies. However, these challenges can be navigated with careful research design and settings of all types may experience any of these issues as well. Recent EEF efficacy trials have successfully worked with both PVI and maintained settings to assess impact of the Hanen Learning Language and Loving It intervention (Bhatti et al., 2024) and the Level 4 Group Triple P: Positive Parenting Programme (Dimova et al., 2021). It is also worth noting that all types of settings may have children join classes, groups, or rooms throughout the year.

Child outcome measures

Short term and long term outcomes for children that were identified in the theory of change include increased listening, collaboration, problem-solving and wider PSED skills, increased creativity and confidence offering ideas, and using the vocabulary and language of Tales Toolkit to create stories (including mark-making). As Tales Toolkit supports both communication and language and PSED, it would be good to explore both aspects in a future efficacy trial. However, feedback from practitioners on perceived changes in their practice and perceived outcomes for children suggest that Tales Toolkit has particular potential for impact around communication and language. In the PSED aspects of the intervention, language and communication played a key role in facilitating and modelling scenarios for children to develop problem-solving skills and to express and understand other people's thoughts and feelings. The evaluation team suggests focusing on communication and language for the primary outcome to be measured in an efficacy trial, and a secondary measure could look at impact on PSED.

Perceived communication and language outcomes described by practitioners included vocabulary, sentence building, and narrative or storytelling. There are a number of tests standardised for use with early years children in the U.K. that assess either overall language and communication skills (for example, WellComm, PLS-5, CELF do this quite comprehensively), language skills (Oxford LanguageScreen), or that assess specific language skills or areas such as vocabulary (BPVS3), sentence building (MacArthur-Bates Communicative Development Inventories—Words and Sentences) and narrative and storytelling (Bus Story Test from Renfrew Language Scales). Some assessments can be delivered by practitioners or completed by parents while others require professional expertise such as speech and language therapists or education psychologists. Also, the more comprehensive tests can take up to 40 minutes to an hour with breaks whereas a more specialised test may only take ten to 15 minutes. This is an important consideration as young children can find it difficult to focus and pay attention for a long time, although many assessments seek to make the tests as interesting and playful as they can. Finally, the sensitivity of standardised measures for early years can change greatly from one year group to the next, which can lead to floor or ceiling effects, so focusing on one year group might also help to ensure the most appropriate test is chosen. Many of these assessments have been used in other EEF trials which may be useful when comparing Tales Toolkit and its characteristics to other early years interventions.

Perceived PSED outcomes described by practitioners included problem-solving, sharing, turn-taking, and listening. It is worth noting that a few staff who had extended Tales Toolkit delivery at their setting to include children aged two to three did not observe so many of the PSED outcomes with this younger cohort and felt that they were a bit young to fully engage in aspects such as discussing emotions. This suggests that any cohort for future studies would need to be at least three years of age to reliably observe PSED impacts. There are standardised assessments for measuring overall PSED from three years (Strengths and Difficulties Questionnaire) or four years (Social Communication Questionnaire).

The Ages and Stages Questionnaire includes subscales for communication and problem solving but has only been validated with British children aged two years to two years and six months, whereas it has been standardised from one to 66 months for children in the U.S.A. As with the language and communication assessments, these assessments have been used in other EEF trials, which may be useful when comparing the Tales Toolkit and its characteristics to other early years interventions.

Summary

While selecting outcome measures requires consultation with the Tales Toolkit delivery team and perhaps some changes to the theory of change, findings from the pilot study identified some key considerations and criteria. It may be more feasible to focus on children in reception because they are more likely to engage in later aspects of the programme, such as making their own story books, are less likely to have wide variation in child attendance patterns and may be eligible for a wider number of standardised tests. Reception is also more structured in terms of timetables and has statutory obligations around recording attendance (in registers), which would ensure more robust tracking of the cohort. In the pilot, PVI settings were more likely to experience staffing issues, especially turnover, offer wraparound care, which may impact on variation in children's attendance patterns, and have limited additional space and resources, which could make accommodating externally delivered assessments more challenging for this group, so evaluators may need to consider extra support or time for assessments with this group. However, recent EEF efficacy trials have successfully worked with both PVI and maintained settings so these challenges could be navigated with careful research design.

Given that the attainment gap is already evident in the early years, it is agreed that this age group should be the focus of interventions to close this gap. Final decisions around year group, measures, and settings should be rooted in capturing where the programme provides most value in terms of closing the attainment gap, what is most useful to the early years sector, but also what is feasible in terms of additional interventions and evaluation activities. Although 57% of children in early years provision in England in 2024 were in PVI settings⁴³ and recent research suggests that the proportion of disadvantaged children in PVI settings has been increasing (Stewart and Reader, 2021), decisions around future trials also need to weigh the practicalities of including PVIs in interventions and evaluation given the difficulties that part of the sector is currently experiencing due to the expansion of funded provision.

⁴³ Department for Education (2024) 'Education Provision: Children Under 5 Years of Age', in 'Table Tool: Create Your Own Tables': <https://explore-education-statistics.service.gov.uk/data-tables/fast-track/1d509c7f-5172-4088-ad01-08dc9a834b55>

Conclusion

Table 12: Summary of pilot findings

Research question	Finding
<p>1. Does the programme lead to a change in practice (as per the short term outcomes in the ToC)?</p> <p>1.1 Is there evidence of collaborative storytelling, quality interactions, and child-led learning?</p>	<p>Most practitioners surveyed indicated (1) that Tales Toolkit had led to changes in their practice and more quality interactions between children and staff, (2) increased confidence and knowledge to support children's language development, collaborative thinking, and play, and (3) that the sessions became more child-led and storytelling more collaborative over time. While some observations led by evaluators confirmed these findings, other observations were less child-led. While it is not possible to firmly attribute this, evaluators believed it could be due to limited practitioner confidence due to, in some cases, not having fully completed the training.</p>
<p>1.2 Is there evidence of practitioners using the resources in their classrooms/other learning environments (e.g., outside)? How do they use them and does this vary by type of setting (for example, PVI vs. maintained settings)?</p>	<p>Most practitioners surveyed reported that Tales Toolkit resources were available for children to access outside of sessions and that the approach has embedded into the wider setting, demonstrated by examples such as children using the problem-solving language outside of the sessions. Where resources were not made widely available, reasons included concern that they would get lost or damaged.</p> <p>Few settings had been able to provide the resources in the outside learning environments due to inclement weather (and in some cases a concern that the resources would get ruined).</p>
<p>2. Do practitioners perceive the intervention has led (or will lead) to positive outcomes for children (e.g., their language and communication skills, PSED, creativity and writing), especially for children EYPP, EAL and SEND)?</p>	<p>Most staff surveyed identified perceived positive outcomes for children in the EYFS prime areas of CLL and PSED including creativity, storytelling, increased vocabulary, speaking more, increased problem-solving skills, and increased turn-taking and sharing. For children with SEND, outcomes were reported for language and communication and PSED especially, and for children with EAL, for language and communication. There were more variations in the mark-making outcomes, depending on the age and development stage of children. However, not all staff delivering session had completed all the training and not all had completed the mark-making sessions when expected.</p>
<p>3. Do practitioners have confidence and motivation to deliver the intervention?</p>	<p>Evidence from surveys and interviews showed that most staff were confident and motivated to deliver the intervention. Where staff were less confident this was often where they had not participated in all the training or had fewer opportunities to deliver it. Survey evidence showed practitioners self-reported increases in confidence for supporting children with collaborative thinking and play, creating thinking and problem-solving skills, and writing or mark-making, as well supporting children with SEND or receiving EYPP.</p>
<p>4. Are there any unintended consequences or negative effects? (e.g., pushing out other programmes, activities widening the attainment gap between disadvantaged and non-disadvantaged pupils, perceived burden)</p>	<p>Most staff interviewed or surveyed did not identify negative effects of delivering Tales Toolkit. Most staff surveyed agreed there had been unintended consequences of delivery, but the examples provided were positive or focused on expected outcomes, such as more boys engaging and children articulating problems and/or emotions.</p>
<p>5. How do pupils engage with the different activities/aspects of the programme?</p>	<p>Evidence from case studies and surveys indicated that, generally, children are engaged with Tales Toolkit sessions, particularly enjoying taking items out of the bag and coming up with ideas. However, some staff reported that it was difficult to keep all children engaged as part of larger groups. Most staff reported adapting delivery to help children</p>

	<p>with EAL or SEND to engage. Where settings made resources available to children outside of sessions, children sometimes played with these or initiated a Tales Toolkit session with a practitioner in free-flow or child-led play.</p>
<p>6. Is the intervention feasible for practitioners to engage in and implement?</p>	<p>Most staff surveyed indicated that they were able to engage with Tales Toolkit, implement it in their classroom, and that the sessions had fitted into their working day and routine. Evidence from case studies also suggested that for the most part Tales Toolkit was straightforward to implement and engage in and most staff found the resources easy to use. However, ten settings (33%) withdrew from programme delivery over the period of the evaluation, seven of which were PVI's. High attrition, especially among PVI's, was primarily explained by staffing issues. Combined with other factors such as class sizes, age, or needs of children (SEND or EAL) this could affect how, when, and how often practitioners delivered Tales Toolkit in their setting.</p>
<p>7. Can settings deliver with fidelity and within routine practice? Is there variation in delivery across settings (e.g., embedded into the curriculum, planning, practice, resources, and language or in separate sessions/activities)?</p>	<p>Most staff surveyed agreed that delivering and preparing Tales Toolkit was straightforward and fitted into the normal day. There were some issues with fidelity with some delivering sessions less than once a week (by the endline). Evaluators observed staff in all setting visited using the structure of 'character-setting-problem-solution' in sessions and the physical resources.</p> <p>Most staff surveyed said that Tales Toolkit had been embedded into their classroom and the observations and interviews confirmed this with practitioners delivering with fidelity and within routine practice, staff and children using the Tales Toolkit language outside of sessions, and the ethos embedded across the settings.</p> <p>However, not all staff who were delivering had completed all the training, so they may not have been able to deliver the intervention with complete fidelity (see RQ8).</p> <p>A few staff may have misunderstood the mark-making activities following Session 4 of the training (which may have been due to not all staff completing this session or not doing it when or as expected).</p> <p>Clearer specification of what constitutes compliance to the training is needed (and confirmation if any elements are optional). Most staff adapted delivery for children who were less able to engage, such as those with EAL and SEND or for different age groups. Tailoring to children's needs was permitted.</p>
<p>8. Does the delivery method of the training (context and mode) have an impact on engagement with the programme (practitioners)?</p>	<p>The asynchronous online mode of delivery was easily accessible to settings and could be revisited. Staff participating in the training found it useful and the discussion activities were especially informative.</p> <p>However, the asynchronous model may have also caused completion issues for some settings. Not all staff delivering the intervention had done all the training and some who had did not do it in the expected way (for example, in a group or with longer videos). This may have meant that some were unable to deliver all aspects of the programme with fidelity. The reasons for non-completion of the training were primarily to do with difficulties getting all staff together (in groups) and general issues with staffing (including staff turnover and the importance of maintaining ratios). There was some misunderstanding around the expectations for the training, for example, new staff only observing colleagues deliver, staff doing all the training on their own, staff only completing the shorter videos with no discussion activities (these were designed for staff doing training on their own). This may have had an impact on engagement with the programme.</p>
<p>9. What are the potential barriers and facilitators of the programme?</p>	<p>Key facilitators included buy-in from senior staff and support from colleagues. Discussion activities during the training were especially helpful for producing ideas, sharing learning, troubleshooting, and embedding the intervention.</p> <p>Key barriers to delivery centred on time pressures, children needing a lot of support around language or PSED, variation in child attendance patterns, and staffing issues (including turnover, shortages, and absences), including staffing challenges arising from</p>

	<p>expanding provision with changes to entitlements for free childcare hours happening during the pilot. Staffing issues were particularly acute among PVI settings.</p>
<p>10. What support is needed and how can a useful support model be developed?</p>	<p>Settings could access support via email or a call with the delivery team, and there was also a Facebook group. However, the delivery team reported that few settings contacted them and there was little activity on the Facebook group.</p> <p>When interviewed, a few staff had queries on whether they were delivering correctly and some wanted advice on delivering to groups with different needs, for example, different age groups or those with EAL or SEND.</p> <p>A more proactive support model could facilitate the delivery team to identify settings or practitioners' concerns, monitor compliance with the training, delivery, and fidelity and engagement, and enable more sharing of ideas or practice. Scoping work could help identify suitable or preferred formats for engagement.</p>
<p>11. Do the training and resources constitute a standard intervention that can be replicated and scaled up?</p>	<p>The asynchronous online mode of the training readily allows for scale-up and the delivery team reported it could provide sufficient physical resources with some preparation time. However, the issues with not all staff delivering who had completed all of the training or not delivering to all children at least once per week are a concern. Improved guidance and increased monitoring of training completion alongside specified, mandatory delivery elements could address this.</p> <p>Variations in children's attendance could make delivering to all children once per week challenging and also lead to some children participating in multiple sessions per week. Advice on delivery in this context could help and the delivery team could set a limit for maximum dose if needed.</p>
<p>12. What, if anything, should be updated or changed for a trial evaluation (e.g., ToC, materials, delivery method)?</p>	<p>The ToC model could be updated with more specific information on expectations for inputs (around completion of training), outputs (around frequency of delivery), and the causal mechanisms and assumptions.</p> <p>Additional detail could also be provided around outcomes in the EYFS prime areas of CLL and PSED. The team may also want to investigate including outcomes in other areas—such as fine-motor skills—as these were identified as perceived outcomes.</p> <p>An even more proactive model of support may be necessary to facilitate engagement and ensure fidelity in delivery. Providing more proactive support to the training facilitator role in each setting could also increase compliance with the training. Practitioners had some practical suggestions for the physical resources, such as laminating materials and providing more video examples of delivering Tales Toolkit with groups of different ages, sizes, or needs such as EAL or SEND.</p>
<p>13. Are there any key contextual factors that appear to facilitate or impede successful implementation at a larger scale, e.g., setting types?</p>	<p>The delivery and evaluation teams had difficulties engaging with settings and particularly struggled to keep PVIs engaged during the pilot (which was time-consuming). These issues occurred especially during recruitment to the pilot, the evaluation, and post-intervention when arranging payments (for staff cover and engagement in evaluation activities). This may have been due to the changes in funded provision being announced and implemented during the pilot.</p> <p>Most PVIs recruited to the project experienced staffing issues (limited resources and staff turnover), and seven of the 13 initially recruited withdrew from programme delivery during the pilot. These staffing issues can make it difficult for settings to engage with and deliver the intervention, but including PVIs is important: in England, over half of children in the early years sector attend PVIs. The delivery team could consult with PVI settings to identify how Tales Toolkit could work best in such settings.</p>

<p>14. What will the cost of the intervention be to settings? Are the costs considered affordable? What is the willingness of settings to pay for the programme and commit the time, beyond the pilot? How does the cost and affordability vary by type of setting?</p>	<p>Tales Toolkit was considered by staff to be straightforward and was not time-consuming to deliver. Few identified any additional costs although a couple of settings described purchasing extra sets or materials to create props. Several setting managers, with hindsight, felt that if they had paid for Tales Toolkit, the investment would be worth it for the benefits generated. A few settings said they would not be able to afford continuing the subscription to Tales Toolkit after the funded pilot period ended. Given the current economic climate and severe budget restrictions, managers in settings of all types did not think they would find the full cost of Tales Toolkit affordable if not funded.</p> <p>If the delivery team adopts an even more proactive model of support, this could have implications for delivery team resources and costs.</p>
<p>15. Are changes in key outcomes (e.g., communication and language) measurable?</p>	<p>Practitioners reported perceived outcomes for children in both CLL and PSED. These included increased vocabulary, storytelling, creativity, mark-making, use of problem-solving language and behaviour, and talking about how they and others feel. There are validated standardised assessments which could be used to assess these areas with U.K. child cohorts (see RQ17). Practitioners observed slightly different impacts with the different year groups—pre-school or reception—because of the different stages of development they were at, so a future evaluation could focus on one age group to capture consistent changes for that development stage.</p>
<p>16. Is it feasible to involve practitioners in research activities—in particular, support in the collection of primary outcome data/depth interviews (e.g., due to time-burden)?</p>	<p>Staff in all types of early years settings experienced time pressures, and staff-child ratios meant it was difficult for staff to leave classrooms or groups. Ten settings (33%) withdrew from programme delivery over the period of the evaluation mostly due to staffing issues; seven of these were PVIs, which indicates specific challenges with engaging this sector. Similarly, the main barrier to research participation was staffing issues: PVI settings in particular struggled with staff turnover. However, these issues can be navigated with careful research design. Although there were challenges with the response rate to the endline survey, levels of participation in the baseline survey and case studies in the evaluation were good overall, and recent EEF efficacy studies have engaged with PVI settings successfully. Changes to entitlements for free childcare in September 2025 will likely see pressure on settings increase with increased demand.</p>
<p>17. Would it be feasible to collect child outcome data? What outcome measure(s) might be used?</p>	<p>Information collected in early years to monitor children’s developmental progress is often not a standardised assessment and may not capture specific outcomes relevant to Tales Toolkit. Previous EEF efficacy trials in early years settings have demonstrated that it is feasible for settings to accommodate external assessors although care should be taken to reduce disruption and burden on settings. PVI settings in particular may not have a separate quiet space for conducting assessments. There were also variations in attendance patterns for children so multiple visits would likely be required to test sufficient numbers.</p> <p>As noted (in RQ 15), most staff reported perceived outcomes around CLL and PSED. There are a number of assessments standardised for use with early years children in the U.K.</p> <p>Practitioners reported differences in perceived outcomes depending on children’s stage of development so focusing a future trial on one age group would help increase the consistency of findings.</p>

Formative findings

Overall, evidence suggests that the training and delivery materials were high quality, well-structured, easy to use, and feasible to implement. Practitioners were able to deliver the Tales Toolkit sessions using the ‘character, setting, problem, and solution’ structure and most children were engaged. However, there were some issues for compliance and fidelity centring on the delivery of training in settings and the embedding and delivery of Tales Toolkit exemplified by significant variations in practice within settings. While the intervention is designed to be flexible and adapted to the needs of

individual settings and children, some practice was not consistent with the delivery model. Some relatively small changes to the intervention could facilitate fidelity and compliance.

Evidence from the surveys and case studies showed that only around half of staff had completed all five training sessions. This would be a serious compliance issue for a future efficacy trial and is assumed to have an impact on staff's ability to deliver Tales Toolkit with fidelity. The delivery team advises that practitioners can start delivering Tales Toolkit sessions after completing the second training session but it is assumed that full completion of the training is required to deliver all aspects of the programme fully. Also, some settings did not follow the recommended schedule for delivering the training across three to five months.

Evidence from the survey and case studies also identified a lot of variation in how settings and staff were delivering Tales Toolkit, including settings where sessions were delivered less than once a week. This was not compliant with the model, which required settings to deliver sessions to all children in their class or group at least once per week. A further challenge was where children's attendance varied and staff delivered multiple sessions throughout the week: this meant that some children were receiving multiple 'doses' per week so may be benefitting more than others. Engagement with other elements of delivery such as mark-making and making resources available outside of Tales Toolkit sessions varied greatly among settings. Some of these issues affecting delivery fidelity may be a consequence of staff not completing all training sessions in the suggested timeline, but also suggests that for future delivery, the delivery team could consider clearer communication around the required elements and minimum/maximum doses per child, a fixed training session schedule for practitioners, approaches to proactively monitor delivery at settings, or whether they wish to identify core elements of delivery that are compulsory and which elements are optional.

There are a number of ways that this could be addressed by the delivery team through:

- making it clearer in their materials that it is a requirement that settings complete all the training and set out the permissible timings for the sessions across an academic year;
- monitoring settings' engagement with the training, either via the website or by contacting settings to confirm compliance;
- facilitating compliance through providing support for training facilitators at settings, such as training for facilitators or calls at the start of delivery to discuss setting plans for training delivery and offer advice; this could help ensure that settings understand training requirements as well as help resolve barriers to training;
- offering top-up introduction sessions throughout the year for staff who have joined a setting or class/group later in the year or providing advice on how to incorporate staff who join later in the training delivery; and
- considering and specifying which aspects of the training are mandatory for compliance.

On this latter point, the fourth and fifth training sessions covered mark-making and an opportunity for staff at a setting to reflect on their practice and delivery of Tales Toolkit. Feedback suggested that the incorporation of mark-making depended on the age or readiness of children. In contrast, the first three training sessions covered high quality interactions and explained how to use the Tales Toolkit resources. The delivery team could consider making the first three sessions mandatory.

The evaluation found that few staff or settings had contacted the delivery team for support or help. The issues identified above suggest that a more proactive model of support may be beneficial to support settings' understanding of the intervention. This could facilitate higher levels of compliance with the training and fidelity in delivery. Some scoping work could help identify suitable formats or preferred modes of engagement, such as social media platforms or messaging services such as WhatsApp and explore this for different setting types, especially as the delivery team experienced challenges contacting PVI settings. The delivery team may wish to consult with PVI settings to explore how Tales Toolkit could work best in such settings.

Interpretation

Evidence to support evidence of promise

There was evidence of promise across the short-term outcomes based on staff-reported perceived impacts across all setting types. Short term outcomes for practitioners identified in the theory of change model included increased understanding of, and importance assigned to, quality interactions with children, collaborative storytelling and child-led learning, as well as increased confidence and motivation with delivering these, specifically with embedding Tales Toolkit delivery and problem-solving language in everyday practice. Most staff reported increased understanding and knowledge of supporting children's language development and most felt confident and motivated to deliver Tales Toolkit. Additionally, most staff said their confidence had increased in supporting children with collaborative thinking and play with others, creative thinking, and problem solving. Staff confidence in delivering Tales Toolkit increased over time but, notably, confidence levels were lower where staff had completed less of the training or had fewer opportunities to deliver sessions.

Most staff indicated that participating in Tales Toolkit had led to changes or improvements in their regular practice and more quality interactions between staff and children. Self-reported changes in practice around supporting language included increased modelling of a wide variety of words and having extended back and forth interactions with children. Self-reported changes in practice around collaborative storytelling and creativity included increased instances of supporting children to extend their ideas or narrative through sustained discussion and of supporting children who found it difficult to join in pretend play or storytelling. Staff-reported changes centred on PSED included increased instances of teaching children how to resolve conflicts and of staff talking through how they and others have resolved difficulties.

Most staff reported that Tales Toolkit had been embedded more widely at their setting through using the problem-solving language of Tales Toolkit outside of the sessions and making resources available to children outside of sessions. In a few settings, staff also delivered Tales Toolkit in free play when asked by a child. However, there were a few settings where resources were locked away after sessions to avoid damage or where Tales Toolkit delivery only happened in free play.

Most staff reported that children were engaging with Tales Toolkit and evaluators observed good levels of engagement. However, staff with large classes, such as 24 to 30 children, found it easier to engage children in smaller groups of ten to 12. Most reported perceived positive language outcomes in children, such as increased vocabulary and becoming more vocal. Staff also reported perceived positive outcomes around creativity and storytelling, with older children demonstrating better understanding of story structure and sequencing. For children with SEND, outcomes were reported for language and communication and PSED especially; for children with EAL, particular improvements centred on language and communication. There was less evidence of perceived positive impacts on writing, which some practitioners attributed to children's age or stage of development. Mark-making was covered in training Session 4, however, the fact that around one-third of staff surveyed at endline had not completed this session may have also impacted their ability to influence outcomes in this area. Staff also reported perceived positive outcomes around PSED, such as children using problem-solving language and approaches from Tales Toolkit outside of sessions, and increased turn-taking and sharing.

Feasibility

As discussed, while staff found the Tales Toolkit training to be high quality, well-structured, and easy to engage with, there were issues around non-completion of the full training and training not always delivered to the intended schedule and in the intended way. Feedback from staff indicated that the group discussion activities were particularly valuable for embedding practice in settings and facilitated peer support and shared learning among staff, but not all sessions could implement this aspect of training. Where settings only delivered the short versions of the training videos without discussion activities, this may place their staff at a disadvantage. The most common reason for staff not completing all training was that settings struggled to gather all staff together for training sessions due to difficulties releasing staff from direct work with children in order to maintain staff-child ratios. Additionally, 'usual practice' staff meetings were not always long enough to accommodate the training in one go (as they may have had to include other agenda items). Incorporating the Tales Toolkit training with relatively short notice was also difficult for some school-based settings where staff meetings, professional development, and INSET days are planned a long time in advance.

Although staff cover was offered as part of the pilot, take-up was limited. This was possibly due to there being no additional costs accrued as training was delivered during staff meetings or the funding being offered after the setting had already paid for cover needed. The cost of cover may also have been viewed as an unanticipated cost as settings plan budgets in advance, but the payment was offered towards the end of the academic year. However, the delivery team also reported difficulty engaging with settings to confirm whether cover had been needed to ascertain whether payment was due.

There was quite a lot of variation in the frequency of Tales Toolkit session delivery. Evidence suggests that most settings and practitioners did so at least once per week, but a minority did not achieve this. In some instances, settings had reduced the frequency of delivery to less than once per week over the course of the pilot rather than embedding it. A key factor determining frequency was the attendance pattern of children: where there was a wide variety of child attendance patterns, settings would deliver Tales Toolkit sessions several times per week in order to ensure that all children received at least one session per week. However, this could mean that children who attended longer hours participated in multiple sessions per week whereas others received one only, which raises an issue for consistency of dose for a future trial. Settings of all types had children on different attendance patterns and enrolling throughout the year but PVI settings were more likely to be offering 'wraparound' cover—such as 8 a.m. to 6 p.m.—so had a potentially wider range of variance across the day.

Settings of all types reported time pressures but PVI settings in particular struggled with staffing issues. Of the ten settings that withdrew over the course of the pilot, seven of these were PVI settings and in all cases their reason for withdrawing was staff issues including staff turnover, sickness, and a manager leaving. The high rate of setting withdrawal from the pilot (33%), of which 70% were PVI settings, suggests that the latter may be more likely to struggle to engage with the intervention at all and more likely to drop out of a future efficacy trial. However, it should still be recognised that there is wide variation within PVI settings: staff at one PVI setting, for example, observed that they were unusual as they were not experiencing any staffing issues. Scoping work with PVI settings to clarify obstacles and identify appropriate solutions may help the delivery team increase feasibility for this part of the sector.

Readiness for trial

Tales Toolkit is a high quality, well-structured intervention that is easy to deliver and the asynchronous online format of training delivery and downloadable resources means that scale-up would be relatively simple. However, the evaluation identified issues regarding how some settings engaged with the training, resources, and support, which may have impacted delivery and the extent to which the intervention was implemented in a standardised way. This would need to be addressed as part of scaling before a trial. The delivery team has stated that it would be able to provide sufficient numbers of physical resources for scale-up given sufficient advance notice. Alongside this, the team will need to consider the issues with compliance centring on full training completion (during the expected timeframe and in the expected way) and some settings delivering less (or more) than recommended (at least once per week). Suggested options for addressing these issues are detailed above, including a more proactive model of support and monitoring, and clearer specifications on mandatory training (including the timing, content, and what is delivered in group situations). The theory of change would benefit from the addition of an 'activities' section setting out the structure and elements of delivery to inform evaluation of fidelity and compliance in a future efficacy trial. It would also be useful for the delivery team to consider further guidance around delivery in contexts where there is greater variation in child attendance patterns.

Feedback from settings suggests that Tales Toolkit would be considered quite a high-cost intervention (given the current economic position), but some managers interviewed commented that it would be worth the investment in light of the benefits they perceived for children and staff. They suggested that the benefits should be highlighted more strongly in promotional information. Settings participating in the pilot were offered a reduced subscription rate for continued access to the online resources but this could still be more than they could afford in the current economic climate. Case study settings of all types commented that budgets were tight, so full information on costs versus benefits was needed to inform investment decisions.

As discussed in the Feasibility section, practitioners perceived specific positive outcomes around language and also PSED, such as increased vocabulary, storytelling, problem-solving, or turn-taking. There are a range of tools suited to capturing these outcomes in children and child data collection is feasible to measure these outcomes.

Limitations of the evaluation

As this was a pilot, the evaluation did not include an impact evaluation to quantitatively measure the effect of Tales Toolkit on children's development in language, creativity, or PSED: practitioners' perceptions were used to identify evidence of promise and areas where the intervention was working well. This is a subjective judgement that is not as accurate and reliable as a relevant standardised assessment. It should also be noted that while evaluators observed delivery of Tales Toolkit during visits, there might have been an observer effect where evaluators' presence might affect staff and child behaviour. It was also possible that, despite briefing from the evaluator, staff in observations may have felt they were being judged on performance rather than being impartially observed for understanding of the activity.

As this was a pilot study, the sample includes a relatively small number of settings and it was not intended that the sample would be representative of the whole early years provider population in England. However, the study's attrition rate was high (and higher than anticipated) with ten of the 30 settings (33%) initially recruited withdrawing from programme delivery over the pilot. Unfortunately, seven of these ten were PVI, which meant that the final sample included only three PVI. This does mean that findings for this segment of the provider market are very limited. The high rate of withdrawal and lower response rate to the endline survey than to the baseline survey could also lead to bias in the survey findings, as settings and staff who were less engaged with the programme may be less likely to have participated in the survey (although staff from three of the ten settings that withdrew did take part in the endline survey).

The plan for the pilot was that the achieved sample would include settings in a range of levels of deprivation, but primarily in areas of higher deprivation (in Deciles 1 to 4) based on IDACI scores. However, seven settings were initially in deprived areas; three withdrew during the intervention, leaving only four (all in Decile 4). Therefore, the evaluation is limited in the extent to which it could explore perceived outcomes for children in areas of high deprivation and also has implications for understanding how delivery and engagement would work in arguably more challenging contexts and with lower levels of language development. This is also different from what would usually happen in scaled-up EEF trials where areas of high deprivation might be oversampled.

Attrition between pre- and post-treatment means that data presented from the endline survey and longitudinal cohort should be treated with caution due to the small sample size. The longitudinal cohort approach, while enabling analysis of change over time, is only for a small proportion of staff. Because of this, only the results showing the largest changes in the average scores are reported in the main report. The increase in the average scores seen in a number of dimensions is positive. However, no statistical analysis of significance was carried out, so the findings should be treated cautiously.

The evaluation was not able to ascertain exactly why attrition was so high as we were only able to interview managers in four settings that had withdrawn. But, in the recruitment phase, it became clear that many settings were having problems carrying out the training (as discussed in the Feasibility section). Anecdotally, lead practitioners reported to us that, due to staffing shortages and budget constraints (which meant they could not always afford to pay for supply), they had not been able to start in the autumn term as they had anticipated. In addition, a number of key staff had left the setting or changed roles in the autumn term, which meant they could not always carry on with the pilot. In a few cases, serious illness meant the setting could no longer participate. As reported, some staff in maintained school-based settings explained that they had not been able to fit the training into their plans for CPD over the coming academic year as the plans were already set in the previous summer term. The evaluation team's findings were identical to those of the delivery team, who also found it difficult to contact settings, for example, to invite them to webinars to discuss progress or even arrange the payments for their staff time.

As described in the protocol, the pilot was designed to include PVI which, although they make up a large proportion of early years provision, are rarely included in evaluations of this kind. In agreement with EEF, PVI were recruited to the trial so that data could be collected on the extent to which these types of settings can engage with interventions such as Tales Toolkit. Due to a lack of data from PVI, this study provides a relatively unique exploration and description of the feasibility of delivering this intervention to children in PVI.⁴⁴ However, this pilot was designed before the announcement of the expansion of funded childcare (in April 2023). The first expansion came in September 2023, just

⁴⁴ The method was not designed to enable an analysis of statistical significance.

as the evaluation activities began. The early years sector was already experiencing issues with staff recruitment and retention before the changes. Unfortunately, this evaluation has been affected by the wider context in the early years sector. The high attrition in survey responses between baseline (N = 156) and endline (N = 56) was higher than what could have been imagined at design stage. These findings across the dimensions must be treated with caution, particularly those looking at perceived child outcomes. However, the results do offer some evidence for how Tales Toolkit could feasibly be trialled in the future, with careful consideration to elements such as the early years workforce, what year group is considered most appropriate, and settings having adequate time to plan and prepare implementation.

Future research and publications

The design of a future efficacy trial would need to take into account some of the challenges faced by early years settings. Settings of all types in the pilot evaluation reported time pressures and there were particular challenges around staffing for most of the 13 PVI settings recruited, seven of which withdrew from programme delivery. It is important to include PVI settings in a future trial as over half of children in early years provision in England are in PVI settings. However, the high rate of attrition among PVI settings would require a larger sample size at trial to compensate and retain sufficient statistical power for the analysis and ensure sufficient numbers of PVI settings are included. Although findings from the qualitative work did generally concur with those of the survey data, a larger sample size would enable a more comprehensive analysis.

In addition, the need to maintain staff-child ratios means that particular thought and care needs to be taken around evaluation activities to ensure that these do not become too burdensome for settings, especially as the planned expansion of free childcare to 30 hours in September 2025 may create further pressure. However, levels of staff engagement with the evaluation activities and intervention were good. Variations in child attendance patterns need to be accounted for in both intervention delivery and if direct assessment of children in settings is used.

Practitioners identified positive outcomes around language and communication and PSED. Reported perceived outcomes varied by age group so focusing on one year group in a future efficacy trial would facilitate identifying consistent outcomes. Due to these variations, any future trial might do best to focus on children in reception or in the final year of nursery. As some of the perceived PSED outcomes were also involved in language and communication, such as turn-taking, a design with a communication and language measure as the primary outcome and a PSED outcome as the secondary measure would be well-placed to detect an effect of the intervention. While the Early Learning Goals set out in the Early Years Foundation Stage (EYFS) statutory framework assess children's communication and language skills and PSED at age two to three years and at the end of reception, these are high-level benchmarks and have not been standardised as an assessment tool so could not be recommended as a measure in a trial evaluation. Similarly, some settings may monitor or track children's progress using a package such as the WellComm or the EYFS statutory framework, but business-as-usual assessment may take place over weeks, which may not fit with evaluation timelines, and there is a risk of unintentional bias from the practitioner.

The evaluation team recommends the use of a standardised test designed for the relevant age group and focused on the outcome areas identified. As discussed, there are a number of tests standardised for use with early years children. These assessments have been used in other EEF trials, which may be useful when comparing Tales Toolkit and its characteristics to other early years interventions.

Direct assessment of children in settings by an independent external assessor is expensive and creates some burden for the settings in accommodating the testing process. PVI settings in particular may struggle with finding a quiet (ideally separate) space for assessments and where there is a lot of variation in children's attendance patterns this may require multiple visits, which would need to be accounted for in the efficacy trial design and budget. In the pilot, PVIs were more likely to be offering wraparound care, such as 8 a.m. to 6 p.m. which could increase variation in attendance patterns further. However, a well-designed trial that is sensitive to the context of early years settings should be able to identify relevant outcomes in a robust way. Previously, a number of EEF efficacy studies have been delivered with early years

children in pre-school and reception year including in maintained standalone settings, maintained school-based, and PVI settings, such as the Hanen Learning Language and Loving It trial⁴⁵ (pre-school) and Flexible Phonics⁴⁶ (reception).

Further research questions that could be investigated in a future trial include whether there are differences in outcome by setting type, such as PVI compared with maintained setting, or by attendance patterns, such as 30 hours, 15 hours, or less. The number of deprived settings in the current pilot, as identified by IDACI score, was quite low so a larger trial should aim to explore differences in outcome for more and less advantaged children as measured by EYPP or free school meals. A further, short pilot could explore whether a revised design for Tales Toolkit that specifies mandatory elements leads to similar perceived outcomes as the full delivery and could possibly compare the two models.

⁴⁵ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/learning-language-and-loving-it-accelerator-fund>

⁴⁶ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/flexible-phonics>

References

- Bhatti, S., Lilley, C., Morgan, H., Phillips, M., Phillips, N., Poquillon, G. and Ruddick-Trentmann, J. (2024) 'Hanan Learning Language and Loving It (LLLLI)', London: Education Endowment Foundation. <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/learning-language-and-loving-it-efficacy?utm>
- Bruner, J. (2020) 'Play, Thought, and Language', *Peabody Journal of Education*, 60 (3). <https://doi.org/10.1080/01619568309538407>
- Cumiskey, L. (2024) 'EYFS Results: Disadvantage Gap Widens for Reception Pupils', *Schools Week (28 November)*. <https://schoolsweek.co.uk/eyfs-results-disadvantage-gap-widens-among-reception-pupils/>
- Dawson, A., Garner, O. and Nancarrow, A. (2023) 'Coaching Early Conversation Interaction and Language (CECIL) Nottinghamshire Sustainability Evaluation Implementation and Process Evaluation', Institute for Employment Studies. https://www.employment-studies.co.uk/system/files/resources/files/CECIL_Notts%201.5%20report_589.pdf (accessed 17/01/25).
- Dawson, A., Huxley, C. and Garner, O. (2022) 'Coaching Early Conversation, Interaction and Language (CECIL) Evaluation: Implementation and Process Evaluation', Institute for Employment Studies. https://www.suttontrust.com/wp-content/uploads/2022/03/CECIL_IPEreport-FINAL-02.03.22.pdf
- Dawson, A., Williams, C. and Akehurst, G. (2021) 'OVO Foundation Report: Education Inequality in the Early Years', Institute for Employment Studies. <https://www.employment-studies.co.uk/resource/ovo-foundation-report-education-inequality-early-years>
- DfE (2023) 'Development Matters: Non-Statutory Curriculum Guidance for the Early Years Foundation Stage', London: Department for Education (revision of 2020 edn). <https://www.gov.uk/government/publications/development-matters--2/development-matters>
- DfE (2021) 'Statutory Framework for the Early Years Foundation Stage', London: Department for Education. <https://www.gov.uk/government/publications/early-years-foundation-stage-framework--2>
- DfE (2024) 'Early Years Foundation Stage Profile Results Academic Year 2023/2024', London: Department for Education. <https://explore-education-statistics.service.gov.uk/find-statistics/early-years-foundation-stage-profile-results/2023-24>
- Dimova, S., Culora, A., Gilder, L., Leenders, E., de Silva, A. and Brown, E. R. (2021) 'Level 4 Group Triple P: Positive Parenting Programme', London: Education Endowment Foundation. <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/level-4-group-triple-p-positive-parenting-program>
- EEF (2018a) 'The Attainment Gap 2017', London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Annual_Reports/EEF_Attainment_Gap_Report_2018.pdf
- EEF (2018b) 'Preparing for Literacy: Improving Communication, Language and Literacy in the Early Years, Guidance Report', London: Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/eef-guidance-reports/literacy-early-years/Preparing_Literacy_Guidance_2018.pdf?v=1682680700
- EEF (2021a) 'Early Years Toolkit | Communication and Language Approaches', London: Education Endowment Foundation. <https://educationendowmentfoundation.org.uk/education-evidence/early-years-toolkit/communication-and-language-approaches>
- EEF (2021b) 'Early Years Toolkit | Play-based Learning', London: Education Endowment Foundation. <https://educationendowmentfoundation.org.uk/education-evidence/early-years-toolkit/play-based-learning>
- EEF (2022) 'Implementation and Process Evaluation Guidance for EEF Evaluations', London: Education Endowment Foundation. <https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/EEF-IPE-Guidance-August-2022.pdf?v=1660741017>
- EEF Blog (March 2022) 'The ShREC Approach – 4 Evidence-Informed Strategies to Promote High-Quality Interaction with Your Children'. <https://educationendowmentfoundation.org.uk/news/eef-blog-the-shrec-approach-4-evidence-informed-strategies-to-promote-high-quality-interactions-with-young-children>

- EEF (2023) 'Over 700 Early Years Settings, Including Childminders, to Take Part in New EEF-Funded Projects', press release 29 March. <https://educationendowmentfoundation.org.uk/news/early-years-settings-including-childminders-to-take-part-in-new-eef-funded-projects>
- González, S. and Bonal, X. (2021) 'COVID-19 School Closures and Cumulative Disadvantage: Assessing the Learning Gap in Formal, Informal and Non-Formal Education', *European Journal of Education*, 56, pp. 607–622. <https://doi.org/10.1111/ejed.12476>
- Hennick, M. and Kaiser, B. (2022) 'Sample Sizes for Saturation in Qualitative Research: A Systematic Review of Empirical Tests', *Social Science and Medicine*, 292. <https://doi.org/10.1016/j.socscimed.2021.114523>
- House of Commons Education Committee (2019) 'Tackling Disadvantage in the Early Years', House of Commons, ninth report of session 2017–19, 7 February 2019. <https://publications.parliament.uk/pa/cm201719/cmselect/cmeduc/1006/1006.pdf>
- Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K. (2016) 'Implementation and Process Evaluation (IPE) for Interventions in Education Settings: An Introductory Handbook', London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Guidance_Final.pdf
- Hussain F., Morris, S., Basi, T. and Bristow, T. (2020) 'Manor Park Talks – Pilot Report', London: Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/documents/pages/projects/MPT_Pilot_Report_final.pdf?v=1687165542
- Hutchinson, J., Reader, M. and Akhal, A. (2020) 'Education in England Annual Report 2020', Education and Policy Institute (EPI), The Fair Education Alliance and Unbound.
- Jones Bartoli, A. J. (2018). 'Using Storytelling to Promote Literacy, Communication and Socio-Emotional Development in the Early Years', Goldsmiths University of London. <http://talestoolkit.com/wp-content/uploads/2018/11/Goldsmiths-research-2018.pdf>
- Kerr, M. E. and Franklin, J. (2021) 'The Economic Cost of Early Vulnerable Language Skills', *Pro Bono Economics*. <https://www.probonoeconomics.com/Handlers/Download.ashx?IDMF=59554aa1-a138-4c35-9104-9e527ba1391a>
- Law, J., Theakston, A., Gascoigne, M., Dockrell, J., McKean, C. and Charlton, J. (2017) 'Early Language Development: Needs, Provision, and Intervention for Preschool Children from Socioeconomically Disadvantaged Backgrounds', London: Education Endowment Foundation. <https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/early-language>
- Law, J., Charlton, J. and Asmussen, K. (2017) 'Language as a Child Wellbeing Indicator', Early Intervention Foundation. <https://www.eif.org.uk/report/language-as-a-child-wellbeing-indicator>
- Lindorff, A., Sylva, K., Ereky-Stevens, K. and Joseph A. (2022) 'Coaching Early Conversation, Interaction and Language (CECIL) Impact Evaluation', Sutton Trust, Esmée Fairbairn Foundation and the Lindsell Foundation.
- Ritchie, J. and Lewis, J. (2003) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, London: Sage.
- Robinson, R., Prokopoff, I., Cain, J. and Pokorny, J. (1991) 'AEIOU Framework'. <https://help.ethnohub.com/guide/aeiou-framework>
- Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R. and Bell, D. (2002) 'Researching Effective Pedagogy in the Early Years', London: Department for Education and Skills. <https://dera.ioe.ac.uk/id/eprint/4650/1/RR356.pdf>
- Speight, S., Takala, H., Read, H., Robinson, Y. and Fletcher, A. (2022) 'The REAL Programme Evaluation Report', London: Education Endowment Foundation. <https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/REAL-Programme-Evaluation-FINAL.pdf?v=1737038599>
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (eds.) (2010) *Early Childhood Matters: Evidence from the Effective Preschool and Primary Education Project*, Oxford: Routledge.
- Sylva, K., Bruner, J. and Jolly, A. (2017) *Play: It's Role in Development and Evolution*, Penguin Books.

- Stewart, K. and Reader, M. (2020) 'The Conservative Governments' Record on Early Childhood from May 2015 to Pre-COVID 2020: Policies, Spending and Outcomes', London: Education Endowment Foundation
<https://sticerd.lse.ac.uk/dps/case/spdo/spdorp08.pdf>
- Stewart, K. and Waldfogel, J. (2017) 'Closing Gaps Early', The Sutton Trust.
<https://www.lse.ac.uk/business/consulting/assets/documents/closing-gaps-early.pdf>
- Tucket, S., Robinson, D., Hunt, E. and Babbini, N. (2024) 'Social Mobility and Vulnerable Learners Report: Annual Report 2024', Education Policy Institute. <https://epi.org.uk/publications-and-research/annual-report-2024/>
- Tucket, S., Hunt, E., Robinson, D. and Cruikshanks, R. (2022) 'Covid-19 and Disadvantage Gaps in England 2021', Education Policy Institute. <https://epi.org.uk/publications-and-research/covid-19-and-disadvantage-gaps-in-england-2021/>
- Vygotsky, I., S. (1967) 'Play and Its Role in the Mental Development of the Child', *Soviet Psychology*, 5, p. 6–18.
<https://doi.org/10.2753/RPO1061-040505036>

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

To view this licence, visit <https://nationalarchives.gov.uk/doc/open-government-licence/version/3> or email: psi@nationalarchives.gsi.gov.uk

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

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](https://twitter.com/EducEndowFoundn)

 [Facebook.com/EducEndowFoundn](https://www.facebook.com/EducEndowFoundn)