EasyPeasy: Learning through play

Evaluation Report

July 2019

Independent evaluators:

Lyn Robinson-Smith, Victoria Menzies, Helen Cramman, Yuqian (Linda) Wang, Caroline Fairhurst, Sarah Hallett, Nadin Beckmann, Christine Merrell, Carole Torgerson, Sue Stothard, and Nadia Siddiqui.

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

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

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

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

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

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

0207 802 1653

jonathan.kay@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The project was independently evaluated by a team from Durham University and York Trials Unit, University of York:

Lyn Robinson-Smith, Victoria Menzies, Helen Cramman, Yuqian (Linda) Wang, Caroline Fairhurst, Sarah Hallett, Nadin Beckmann, Christine Merrell, Carole Torgerson, Sue Stothard, and Nadia Siddiqui.

The lead evaluators were Lyn Robinson-Smith, Victoria Menzies, and Helen Cramman.

Contact details:

Name: Victoria Menzies

Address: School of Education, Durham University, Leazes Road, Durham, DH1 1TA

Tel: 0191 334 4177

Email: victoria.menzies@durham.ac.uk

Executive summary

The project

EasyPeasy provides game ideas to the parents of preschool children to encourage play-based learning at home, with the aim of developing children's language development and self-regulation. Parents receive a text message directly from EasyPeasy which links to videos of example games that they can play with their children, plus tips and advice about learning through play. The games target skills within the Early Years Foundation Stage areas of learning. This project aimed to support children aged three to four in school nurseries. The intervention lasted 20 weeks, with parents receiving weekly texts linked to 65 games over this period. Use of EasyPeasy was supported by a member of nursery staff (the 'Pod Leader'), trained by EasyPeasy.

One hundred and two primary schools from nine Local Authorities across England participated in this efficacy trial from September 2017 to July 2018. The programme was evaluated using a randomised controlled trial, testing the impact of EasyPeasy on children's language development compared to 'business as usual' in control schools using the Clinical Evaluation of Language Fundamentals (CELF) assessment. Social, emotional, and behavioural outcomes were also assessed; 1,205 children were included in the evaluation. Interviews, case studies, focus groups, and a survey were conducted to explore how the programme was implemented and to obtain feedback from participants.

Key Conclusions	
1.	Children in schools receiving EasyPeasy did not make any additional months' progress in language development compared to children in control schools, as measured by a composite, summary language score. This finding has a moderate to high security rating.
2.	There were small increases in 'word structure' and 'concepts and following directions' language subscales (equivalent to one month's additional progress) compared to children in control schools, but no additional months' progress in 'sentence structure' or 'expressive vocabulary'.
3.	Mixed results were found for children's social, emotional, and behavioural outcomes. Children who received EasyPeasy made small increases in sociability, cognitive self-regulation, and emotional self-regulation compared to the control group. However, effects on externalising, internalising, and prosocial behaviour and behavioural self-regulation favoured the control group.
4.	Parents receiving EasyPeasy reported improvements in the home learning environment. This included large increases in 'modelling', 'responsivity', and 'variety of activities and interactions'. These results are less secure than the main findings due to the small number of parents assessed, and should be treated as exploratory.
5.	Engagement from parents for the continued use of EasyPeasy was considered to be a challenge for nurseries. The most effective ways of encouraging parent participation included integration of the games into the classroom, introducing parents to the games at 'Stay and Play' sessions, and parents sharing comments.

EEF security rating

The primary finding has a moderate to high security rating. This was an efficacy trial, which tested whether the intervention worked under everyday conditions in a large number of schools. It was a well-designed randomised controlled trial and few pupils who started the trial were not included in the final analysis: one school (1%) withdrew from the trial and did not allow post-testing to be conducted, and 6% of pupils randomised with pre-tests had missing data. There was some imbalance on language scores at baseline, which reduces the security rating. All results were confirmed by additional analyses.

Additional findings

Small, non-statistically-significant effects in favour of the intervention group were observed in the individual CELF subscale scores, each equating to a month or less of progress.

Small effect sizes in favour of the intervention group were found for the CSBQ subscales of 'sociability', 'emotional self-regulation', and 'cognitive self-regulation' (effect sizes 0.04, 0.06, and 0.14). The positive effect on cognitive self-regulation is consistent with previous evaluations of EasyPeasy, and here, it is based on a teacher (rather than parent) measure. However, it is also important to note that the effect here is substantially smaller and not statistically significant.

Conversely, small effect sizes in favour of the control group (without access to EasyPeasy) were found for the 'externalising', 'internalising', 'prosocial behaviour', and 'behavioural self-regulation' subscales (effect sizes -0.08, -0.08, -0.02, -0.02, respectively). None of the differences in CSBQ outcomes between groups were statistically significant.

A measure of the home learning environment (HOME), based on parent interviews and observations of the home, favoured the intervention group, and the difference was statistically significant (effect size 0.76, 95% CI 0.11 to 1.42, p = 0.02). This result may provide some support for the programme's theory of change. However, this measure was only assessed in a small and self-selecting sample of 50 parents across treatment and control groups (<5% of the total), so results are underpowered by design, and should be viewed as exploratory only.

Pod Leaders and parents viewed the potential of EasyPeasy positively, acknowledging the quality of the games. However, Pod Leaders were not confident of the real depth of the programme's impact as the level of active parental involvement in using EasyPeasy was lower than their expectations. A survey indicated that the numbers of parents accessing EasyPeasy reduced during the intervention period. By the middle of the programme, half of settings estimated that 25% of parents or fewer were accessing EasyPeasy. Pod Leaders considered the main barrier to delivery to be a lack of time to engage with EasyPeasy, for both themselves and parents. However, Pod Leaders also reported that when parents engaged with the EasyPeasy games, the parents valued the time they had spent with the child.

Cost

The cost per pupil per year (for the 20-week intervention) is estimated to be £24.48 averaged over three years. The only direct cost to nurseries is the purchase of the annual licence for EasyPeasy. Nurseries estimated that for the 20-week intervention, approximately 23 hours of staff time was required for setup and ongoing support and encouragement for parents within their nursery.

Impact

Table 1: Summary of impact on primary and secondary language outcomes

Outcome/ Group	Effect size (95% confidence interval)	Estimated months' progress	EEF security rating	No. of pupils	P value	EEF cost rating
CELF Core Language (Primary Outcome)	0.04 (-0.10, 0.18)	0	🔒🔒🔒🔓🔓	1,128	0.60	£ £ £ £ £
Sentence structure	0.00 (-0.15, 0.16)	0	N/A	1,124	0.96	£ £ £ £ £
Word structure	0.06 (-0.12, 0.24)	1	N/A	1,120	0.52	£ £ £ £ £
Expressive vocabulary	0.01 (-0.11, 0.12)	0	N/A	1,112	0.92	£ £ £ £ £
Concepts and following directions	0.05 (-0.13, 0.23)	1	N/A	1,017	0.58	£ £ £ £ £

Introduction

Intervention

The EasyPeasy intervention was delivered by EasyPeasy and evaluated by researchers at the School of Education at Durham University. The intervention involved parents[1] of preschool-aged children (age three to four years) receiving text messages which linked to online suggestions for simple games to play with their children (games were played offline and not on electronic devices). The intervention was designed to improve parent-child interactions and learning at home with their children, and, through this, to accelerate children's cognitive development, self-regulation, and language and communication skills.

Who: Recipients of the intervention

EasyPeasy was targeted at children aged three to four years in school nurseries who were moving into the Reception year at school in the September following the intervention. All children within the nursery who met this criterion were invited to participate in the trial.

What: Procedures

The EasyPeasy intervention involved parents receiving regular automated text messages containing a link to a webpage detailing ideas for games which parents could play with their children. The webpage containing a short video clip along with information on child development and explanations of each of the games. The games were to be played in the home environment by parents and children together. Parents also received automated reminder emails throughout the course of the intervention to encourage their participation. The 20-week intervention continued to run throughout school holidays.

Coordination of the intervention within each nursery was carried out by a 'Pod Leader', a member of staff selected by their nursery and trained by EasyPeasy to lead the use of the intervention with parents. A 'Pod' could vary in size from a subset of a single class to the whole of the nursery cohort. Most nurseries chose to offer EasyPeasy to their whole nursery cohort, however, not all parents chose to participate. From the parents who did choose to use EasyPeasy, not all opted to be part of the trial. Therefore, a Pod could contain many more parents than just those participating in the trial. Pod Leaders had access to an online dashboard to facilitate delivery of the 20-week programme. Pod Leaders were also expected to register parents into the programme online, to oversee parental activity, and to foster continued parental engagement, including discussions with parents about their experiences with the games.

Pod Leaders received weekly email reports throughout the 20-week intervention period summarising the engagement levels of parents from their Pod by three numbers: (1) 'play dates'[2] occurring within their Pod that week, (2) playdates so far in the Pod, and (3) games, facts, and tips delivered to the Pod.

What: Materials

Parents received weekly messages providing access to 65 games via a text to their mobile phone during the 20-week intervention period. Games included Yummy Strawberries, Selfies, Stepping Stones, Imaginary, and Draw Round Child. Each game took only a few minutes to play and was targeted towards the Early Years Foundation Stage (EYFS) area(s) as well as learning goal(s). For example, the Stepping Stones game linked to the 'communication and language' area while the learning goal was about listening and attention. In this game, pieces of paper were cut into mathematical shapes—triangles, squares, or circles. These shapes were placed in lines as if they were stepping stones on a river. Parents then asked their child, using mathematical language, to cross the river. Some of the games in the intervention

[1] Throughout this report, the term 'parents' is used to refer to parents and main caregivers. EasyPeasy was offered to parents, main caregivers, and other caregivers with responsibility for eligible children at participating nurseries.
[2] The definition of a 'playdate' was a parent clicking on the text message link and accessing the information page about the game. No further information was provided to Pod Leaders as to whether the parent then went on to play the game with their child. Due to the lack of additional information available, the number of playdates was used as a proxy for whether parents played the games with their children and hence as a measure of the level of parental engagement. Playdate information was provided by Pod Leaders at the end of the intervention.

were tailored to different levels with the aim of differentiating the activity according to ability levels. Parents could comment online and these comments were available for other parents in the Pod to read as well as the Pod Leader.

Who: Intervention providers

The intervention was delivered by EasyPeasy, with parents then playing the games provided with their children. The implementation of the intervention was facilitated by Pod Leaders in each participating nursery school, trained by EasyPeasy, to support and encourage the ongoing engagement of parents. The one-hour online webinar training presentation delivered by EasyPeasy explained the rationale for the project, what research had been done to date, how to use the EasyPeasy dashboard, provided examples of games, and who to contact during the course of the project.

Where: Location

The games in the EasyPeasy intervention were designed to be played at home between the parent and child.

Nurseries participating in the trial were located in nine Local Authorities (LAs) across England: Bedford, Camden, Coventry, Doncaster, Durham, Islington, Luton, Knowsley, and Oldham.

How much: Duration and dosage

The EasyPeasy intervention provided access to 65 different games to parents over a 20-week period via weekly text messages during the spring and summer terms of 2018.

The 20-week intervention was implemented in two batches to account for differences in the timing of the pre-intervention testing periods; 44 nurseries were randomised in January 2018 with intervention nurseries completing their 20-week programme in June 2018, while the remaining 58 nurseries were randomised in February 2018 ending their 20-week programme in July 2018.

Tailoring:

EasyPeasy was a prescribed intervention and optimum treatment fidelity was encouraged. However, nurseries were permitted to make minor adaptations to the way that parents were supported and encouraged to participate—for example, in the frequency and method of reminders to parents from the nursery (outside the EasyPeasy interface) and the level of support from the nursery to parents to help with understanding and implementing the games.

How well (planned): Strategies to maximise effective implementation

Before the start of the intervention, the Pod Leaders in the intervention group were required to attend a one-hour online webinar training presentation led by EasyPeasy. The webinar introduced the Pod Leader dashboard, provided an introduction to the games and commenting feature, went through the practicalities of adding parents to a Pod, and advised on how to maximise parental engagement. Pod Leaders were also able to call or email EasyPeasy for support during the intervention period with questions around the delivery of EasyPeasy. Pod leaders received a weekly email from EasyPeasy telling them what parent engagement had been like during the previous week. The level of support provided by settings to parents varied greatly. Some nurseries provided regular support sessions for parents, whilst others did not provide any additional support once parents were signed-up to use EasyPeasy.

Theory of Change

The assumption within the theory of change is that the EasyPeasy intervention will change child self-regulation which will lead to accelerated development in language and communication and improved school readiness (see diagram in Appendix C). The expectation is that this will be achieved through the nursery teachers engaging with the parents regarding EasyPeasy and the parents engaging with their children through the EasyPeasy games. As described with examples in the What: Materials section above, each game is targeted towards both Early Years Foundation Stage (EYFS) area(s) (for example, communication and language skills) as well as having learning goals (for example, listening, attention, and the ability to focus on following instructions). As well as improved self-regulation and language

and communication development from playing the games, the expectation is that there will also be an improved home learning environment due to greater parent-child interaction. The expected impact was that this will lead to an improvement in children's readiness to learn along with improved parental engagement with the school.

Background evidence

It is well documented that children from disadvantaged backgrounds show lower attainment levels on entry to school than those with greater socioeconomic advantages, especially in language and communication skills, and this gap begins to emerge before the start of school (Merrell et at., 2014; Tymms et al., 2014). These skills have been a good indicator of later education attainment (Snow et al., 1998).

The number of words spoken to children by the age of four can vary by as many as 30 million (Hart and Risley, 1995), and children who experience a cognitively stimulating home environment are at an advantage in the learning process. The quality of the early home learning environment is related to the availability of quality educational resources (Melhuish et al., 2008), and evidence highlights an association between disadvantaged households and lower quality educational resources (Foster et al., 2005). Therefore, an early intervention at this stage, specifically in tailoring the activities parents do with their children, can have a particularly positive impact (Lugo-Gill, 2008; Tamis-LeMonda, 2009). These stimulating activities may also develop the child's ability and motivation towards learning more generally (Melhuish, 2010).

In line with this background, which points to the importance of language development and a supporting home learning environment in early years, the impact of EasyPeasy on child self-regulation and behaviour, parent self-efficacy, and parent stress has previously been evaluated as an intervention through two small-scale trials in two Local Authorities—Bournemouth and Newham (Jelley, Sylva and Karemaker, 2016; Jelley and Sylva, 2018). The former study was conducted with children aged between two years and four months and six years old in children's centres, with randomisation carried out within the settings; the intervention lasted 18 weeks. In the latter study, participants were children aged three to four attending children's centres, and in this study, randomisation was carried out at the setting level. The findings were that EasyPeasy had a positive effect on parent-reported child cognitive self-regulation (effect size 0.35) and on parents' self-efficacy regarding their sense of control (discipline and boundaries; effect size 0.26). However, both studies reported limitations in the findings such as social desirability (self-report measures completed by parents wanting to show child in best light), low retention between recruitment and post-test, small sample size, and no objective independent child self-regulation measure. Further robust evidence is required to strengthen contemporary education policy in this area. Consequently, the Education Endowment foundation (EEF) commissioned Durham University to conduct an independent evaluation of the EasyPeasy intervention in a large-scale efficacy trial.

Evaluation objectives

The evaluation investigated the impact and implementation of the intervention. The research questions for the impact evaluation were:

1. What was the impact of the EasyPeasy intervention on the language and communication development of children aged three to four years old (primary outcome)?[3]

2. What was the impact of the EasyPeasy intervention on the self-regulation of children aged three to four years old (secondary outcome)?

3. What was the impact of the EasyPeasy intervention on the social-emotional development of children aged three to four years old (secondary outcome)?

4. How effective was the EasyPeasy intervention at improving the home learning environment of children aged three to four years old (secondary outcome)?

[3] The CELF Core Language Standard Score is the primary outcome, with the individual subscales (Sentence Structure, Word Structure, Expressive Vocabulary, Concept, and Following Directions) as secondary outcomes.

The processes involved for nurseries and parents to implement the intervention, perceptions of relevant stakeholders, and the extent to which the intervention was delivered to the original design were explored through the implementation and process evaluation. The research questions included:

1. To what extent are the settings and parents engaging with and delivering the intervention?
2. How is the intervention being disseminated to parents through settings and is the method effective?
   (a) Are nominated staff engaging with EasyPeasy and using it as a mechanism to foster communication and engagement between parents?
   (b) Are parents engaging with EasyPeasy and implementing the suggested activities within the home?
3. Can all parents access and engage with the EasyPeasy intervention—for example, from all socio-economic backgrounds, in EAL families, in families with SEND children?
4. What are the different stakeholder viewpoints on the intervention—
   (a) those of setting practitioners?
   (b) parents?
5. How effective and appropriately pitched are the activities—
   (a) for parents to deliver?
   (b) for children to receive?
6. What is the acceptability of the intervention to parents and does this differ depending on the socioeconomic status of the parents (based on their child's eligibility for Early Years Pupil Premium), their child's Special Education Needs diagnosis, or English as an Additional Language status?
7. What are the key success factors required for the EasyPeasy intervention to work well?
8. What are the barriers to successful delivery of the intervention—
   (a) for parents?
   (b) for setting practitioners?
9. What areas of the programme could be further developed following completion of the project?

Ethics and trial registration

Ethical approval for the evaluation was received from Durham University's School of Education Ethics Committee on 31 August 2017 and from the University of York's Department of Health Sciences Committee on 1 September 2017. Approval from both universities was sought in line with their respective ethics application processes. The evaluation was conducted in accordance with the British Educational Research Association's ethical guidelines and Durham University's ethics committee research guidelines.

Several additional ethical approvals were obtained from Durham during the course of the project: quality assurance of CELF assessment through audio recording, approved 9 November 2017; case studies, approved 4 April 2018; video recording for CELF training, approved 23 April 2018.

Nurseries opted in to participating in the trial by completing a Memorandum of Understanding (MoU), signed by the headteacher of the school and the nominated Pod Leader from each participating school nursery. An example of the MoU for all school nurseries other than those in Doncaster LA—and that signed by schools in Doncaster LA—can be found in the Additional Appendices.

Parents signed a participation agreement to take part in the trial and to agree to the collection of children's demographic information, participation in baseline and outcome assessments, collection of school destination data, and linking the child's data with the National Pupil Database (NPD) (see Additional Appendices). Parents participating in the home visits completed an additional participation agreement to participate in this part of the trial (see Additional Appendices).

Additional permission from participants was sought as part of the process evaluation. Relevant information sheets and participation agreement forms are provided in the Additional Appendices for:

- Pod Leaders participating in case-study interviews;
- parents participating in case-study focus groups;
- the Pod Leader practitioner survey for the nurseries in the intervention group; and

- the developer—regarding an interview post-intervention (at the end of the trial).

Upon agreement of the trial protocol, the trial was registered with ISRCTN (www.controlled-trials.com) in February 2018 (no. ISRCTN22325174).

Data protection

The General Data Protection Regulation (GDPR) came into effect in May 2018, in the middle of the evaluation project. Privacy notices for all participating Pod Leaders, parents, and the developer were distributed in May 2018 or at the time of participation in additional activities after May 2018. Privacy notices are provided in the Additional Appendices for:

- Pod Leaders;
- parents;
- parents participating in home visits;
- Pod Leaders participating in case-study interviews;
- parents participating in case-study interviews; and
- the Developer.

'Legitimate interest' is the legal basis for processing data for this evaluation. Although right to erasure is not one of the rights under the lawful basis of legitimate interest, ethically, all participants were given this right as outlined in the participation agreements in the Additional Appendices. All parents and Pod Leaders were given the option of withdrawing from the project at any time. All of the interviews carried out as part of the project were recorded digitally, transcribed, and then analysed according to the ten research questions presented in the Process Evaluation. Interviewees were assured that the content of the interview would remain confidential and would be treated in accordance with the Data Protection Act 1998 and GDPR.

All sensitive data collected for Special Education Needs and Disability (SEND) status necessary for the research purposes of the evaluation was processed under Article 9 (j) of GDPR. No sensitive data is shared outside Durham University except to the University of York in a pseudonymised format for analysis.

Project team

Delivery team

EasyPeasy was developed and delivered by a team from EasyPeasy including:

Jen Lexmond, Founder and CEO;

Jane Bradbury, Account Manager; and

Nicola Doherty, Research Manager.

Evaluation team

The independent evaluation was led by a team of researchers from Durham University in collaboration with the York Trials Unit at the University of York.

The principal investigator (PI) for the project was Dr Lyn Robinson-Smith. Lyn's role included designing the trial, writing the protocol, and contributing towards the writing of the final report. Christine Merrell became acting PI from September 2017 to December 2017. Victoria Menzies became acting PI of the evaluation from December 2017 to October 2018. Her role included management of the project and leading on the process evaluation. Dr Helen Cramman oversaw the impact evaluation for the project, led the team conducting the baseline assessments, and coordinated and contributed to the writing of the final report, and was PI for the project from October 2018.

Dr Yuqian (Linda) Wang was the main research associate on the project. Within her role, she conducted the case studies and HOME visits, contributed to training sessions for CELF assessors, conducted quality assurance of the CELF assessment process, and contributed to the writing of the final report.

Sarah Hallett coordinated pre- and post-intervention testing, oversaw the data collection, entry, cleaning, and coding of impact evaluation and school destination data, managed the finances and HR requirements for the evaluation, and contributed to the writing of the report. She was also first point of contact for nurseries and independent assessors.

Dr Nadin Beckmann devised and delivered training in the administration of the CELF assessment in the context of this trial and advised on appropriate quality assurance processes to ensure the reliability of data collected.

Professor Christine Merrell and Professor Carole Torgerson contributed to the design and conduct of the trial and provided expertise to interpret the findings from the outcomes measures and advised on the implementation of the evaluation. Dr Susan Stothard advised on appropriate outcome measures for the trial.

Caroline Fairhurst (York Trials Unit, University of York) contributed to the design and conduct of the trial, developed the SAP, undertook the statistical analysis for the impact evaluation, and contributed to the writing of the report.

A team of trained assessors (employed by Durham University and trained by Dr Nadin Beckmann) administered the CELF assessments. Trained research assistants (employed by Durham University and trained by Dr Yuqian Wang) completed the HOME assessments.

Methods

Trial design

This trial was a pragmatic two-armed cluster randomised controlled efficacy trial, conducted with random allocation at the nursery level in a 1:1 ratio to either the intervention group, which received the EasyPeasy (20-week) intervention, or a 'business as usual' control group. The reasons for adopting nursery-level randomisation were to avoid potential contamination between groups and to ensure equity across pupils in the same year groups. Table **2** summarises the trial design.

School nurseries were randomly allocated to receive the intervention in the 2017/2018 academic year (Jan/Feb 2018 to Jun/Jul 2018), or to a 'waitlist' control group, offered the intervention in 2018/2019 for use with their nursery (age 3+) cohort regardless of the results of the project. Schools did not receive a financial incentive for participating in the evaluation (schools in Doncaster paid £250 each as a contribution towards the LA's cost of participating). Therefore, all schools had access to the 20-week EasyPeasy programme with their nursery cohort in either 2017/2018 or 2018/2019. Intervention schools also had the opportunity to use EasyPeasy with their Reception cohort in 2017/2018, but these children did not form part of the trial.

Due to challenges in hiring and training a sufficient number of staff to complete the baseline CELF assessments, pre-intervention assessments were conducted in two batches with randomisation occurring at the end of each batch of testing for nurseries that had completed the assessments within that batch. The first batch of 44 school nurseries were assessed during the period 4–18 December 2017 and the second batch of 58 nurseries were assessed during the period 8–26 January 2018. Subsequently, randomisation and the intervention were started in separate batches to ensure that each nursery received the intervention for 20 weeks following the baseline assessments and randomisation. Post-intervention CELF assessments were also completed in batches during the period 11–29 June 2018 for batch 1 and 2–20 July 2018 for batch 2.

The original aim was to recruit 120 school nurseries and, allowing for attrition, for 108 school nurseries to complete the full trial with equal numbers of schools in the intervention and control groups. Although initial interest from schools in taking part was high (134 schools returned a signed MoU), the final number participating in pre-testing was lower due to: a lack of parent interest in some nurseries; competing priorities, such as Academy conversion and Ofsted, meaning that pre-testing was not possible; and nurseries not returning parental participation agreement forms and pre-test paperwork. The final number of nurseries participating in pre-testing was 102. While recruitment and pre-testing were ongoing, discussions took place between EasyPeasy, the EEF, and the evaluation team about the sample size that was required for the trial to continue. A revised power calculation was considered—with 102 nurseries and ten families per nursery (originally 13)—and this sample size was agreed, between the evaluation team and the EEF, to be adequate for the trial to continue.

Participant selection

The aim was to recruit eligible school nurseries that were: state-funded primary schools in England; with three-year-old pupils in their school population who had not previously been involved with EasyPeasy; and had an average ever-Free School Meal (FSM) percentage of >30% overall (note: %FSMEVER was changed to %EYPP by EasyPeasy during recruitment). Level of FSM/EYPP was used to ensure schools with the highest levels of deprivation were targeted. Due to challenges with school nurseries not being able to recruit enough parents, not returning information during the set-up phase, as well as some school nurseries not completing the pre-assessments, 15 schools with <30% EYPP were also included in the trial. The protocol originally stated that additional schools with <30% FSMEVER would be considered if recruitment of schools with preferred characteristics was problematic.

Table 2: Trial design

Trial type and number of arms		**Two-armed cluster randomised controlled efficacy trial (random allocation at nursery level).**
Unit of randomisation		School nursery.
Minimisation factor		Number of children with parental agreement to participate per nursery (2 levels; <14; ≥14).
Primary outcome	variable	Language and communication.
	measure (instrument, scale)	Clinical Evaluation of Language Fundamentals (CELF) Preschool 2 UK—Core Language Standard Score.
Secondary outcome	variable	Language and communication.
	measure (instrument, scale)	CELF Preschool 2 UK individual subscale scores: Sentence Structure, Word Structure, Expressive Vocabulary, Concept, and Following Directions.
Secondary outcome	variable	Self-regulation and social-emotional development.
	measure (instrument, scale)	Child Social Behaviour Questionnaire (CSBQ) subscales: Sociability, Externalising, Internalising, Prosocial, Behavioural Self-regulation, Cognitive Self-Regulation, and Emotional Self-regulation.
Secondary outcome	variable	Home environment.
	measure (instrument, scale)	Home Observation and Measurement of Environment (HOME) total and subscale scores: Learning materials, Language stimulation, Responsivity, Academic stimulation, Modelling, Variety of activities, and parental interaction.

Participating schools were only eligible to take part in the study if they agreed to all of the study requirements outlined in the MoU (see Additional Appendices) which described their commitment to the delivery of EasyPeasy including participation in the recruitment of a minimum number of families to the trial and administration of the measures (outlined in next sections). School nurseries were requested to gain a minimum number of ten parents/guardians signing agreements to participate whose children were three years old at the start of the intervention and due to start reception class in September 2018.

The school recruitment process for the trial was led by EasyPeasy and supported by the evaluation team at Durham University. School recruitment took place in three phases.

The first phase involved liaising with and recruiting LAs by the EasyPeasy team. Originally four LAs were to be recruited; however, due to receiving interest from many more, nine LAs were recruited to the trial (Bedford, Camden, Coventry, Doncaster, Durham, Knowsley, Luton, Islington, and Oldham). LAs each provided a £15,000 contribution towards the cost of EasyPeasy whilst the remaining cost was covered by the EEF. Camden and Islington and Bedford and Luton LAs split the cost equally between themselves, each paying £7,500.

The second phase of recruitment consisted of outreach activities led jointly by EasyPeasy and LA partners; 341 eligible schools within the nine LAs received information about the trial through a mail-out and were advised that their participation in the trial would be determined on a 'first come first served' basis and that a limited number of places were available. Schools indicated their willingness to participate in the trial by returning their signed MoU to EasyPeasy which then shared these with the evaluation team. Each school was responsible for nominating a staff member to manage their school's participation in the trial; this staff member also took on the role of 'Pod Leader'. If the school was randomly allocated to receive the intervention, the allocated staff member also took on the role of coordinating the use of EasyPeasy in the school nursery. Participating schools were not required to contribute to the cost except schools in Doncaster LA that were asked to contribute £250 each during recruitment towards the cost to the LA of participating. The impact on recruitment of Doncaster schools having to pay to participate is not known.

Once schools had returned signed MoUs, the third phase of recruitment was for the evaluation team at Durham University to send out welcome packs to school nurseries which included parent information sheets, parent participation agreement forms, nursery 'usual practice' questionnaires, pupil demographic information collection forms, and pupil baseline measures to be completed by nursery staff. Return of parent participation agreement forms along with the completed baseline and demographic information was a requirement for the schools to participate in the trial. All parents of children aged three to four years, who were due to start Reception in the September following the intervention, were invited to participate in the trial.

Outcome measures

The primary and secondary outcomes for the trial are summarised in Table 3.

Primary outcome

The primary outcome for the trial was the Core Language Standard Score from the Clinical Evaluation of Language Fundamentals Preschool 2 UK assessment ('CELF Preschool assessment'), measured prior to the 20-week intervention and immediately after. The assessment was developed by Wiig, Secord and Semel (2006) and is available through Pearson Clinical assessments.[4]

The assessment was selected as it provides a measure for expressive and receptive language skills in young children, which were skills that the EasyPeasy intervention aimed to improve. The assessment contains multiple-choice questions, with the correct picture answer along with other distractor pictures on a page in a display booklet set on a table in front of the test administrator and the child. The child is asked by the test administrator to follow instructions and to point to the answer in the display booklet that best describes the answer to the question that has been asked, for example, 'point to the cat after I have pointed to the monkey'. The assessment is split into several subtests. Each subtest includes various question items (ranging from 20–22) and test administrators were trained to follow the assessment's discontinuation rules to ensure that pupils were only asked questions appropriate to their ability.

For the purposes of this trial, the following norm-referenced subtests of the CELF Preschool assessment measure were assessed and used in the analysis:

- Sentence Structure (22 items; discontinue rule: five consecutive zero scores)—raw score from 0–22, with a higher score indicating a better outcome (scaled scores range 1–18);
- Word Structure (24 items; discontinue rule: eight consecutive zero scores)—raw score from 0–24, with a higher score indicating a better outcome (scaled scores range 1–19);
- Expressive Vocabulary (20 items; discontinue rule: seven consecutive zero scores)—raw score from 0–40, with a higher score indicating a better outcome (scaled scores range 1–19); and
- Concepts and Following Directions (22 items; discontinue rule: six consecutive zero scores)—raw score from 0–22, with a higher score indicating a better outcome (scaled scores range 1–14).

[4] http://www.pearsonclinical.co.uk/Psychology/ChildCognitionNeuropsychologyandLanguage/ChildLanguage/CELF-Preschool2UK/CELF-Preschool2UK.aspx

Table 3: Description of outcome measures and collection methods

	Outcome	Data collection method	Variable
Primary	Change in children's language and communication between pre and post data.	**CELF-Preschool 2 UK** developed by Wiig, Secord and Semel (2006).[5] This will be administered to all children in the study sample by trained assessors. Administration time of relevant scales: 20–25 minutes per child.	Core Language Standard Score, comprising the first three of the four subtests listed below in the secondary outcomes.
Secondary	Change in children's language and communication between pre and post data.	**CELF-Preschool 2 UK** This will be administered to all children in the study sample by trained assessors. Administration time of relevant scales: 20–25 minutes per child.	Individual subscale scores: • sentence structure • expressive vocabulary • word structure subtest • concepts and following directions.
	Changes in children's self-regulation and social-emotional development between pre and post data.	**Child Social Behaviour Questionnaire (CSBQ)** developed by Howard and Melhuish (2016). The questionnaire can be completed on paper by a class teacher/practitioner for all eligible children. Completion time: ≤5 minutes.	Assessment strands reported as subscales include: Sociability, Externalising, Internalising, Prosocial, Behavioural Self-regulation, Cognitive Self-Regulation, Emotional Self-regulation.
	Changes in home environment between pre and post data.	**Home Observation and Measurement of Environment (HOME)** (Caldwell and Bradley 1984). Planned to be administered to a sub-sample of 50 households split evenly between control and intervention groups. Administration time: 45–60 minutes by a trained researcher.	Assessment strands include: Learning materials, Language stimulation, Responsivity, Academic stimulation, Boundaries, Variety of indoor/outdoor activities.

For each subtest, a raw score (sum of items scores) was converted to a scaled score using the age appropriate tables in the CELF manual. Subtest scaled scores were available from three years and zero months to six years and 11 months and reported for each six-month interval in these tables. In addition, a Core Language Standard Score was derived by summing the scaled subscale scores of Sentence Structure, Word Structure, and Expressive Vocabulary, where there was a valid score for at least one of these subscales. This score serves as the primary outcome. The individual subtest scores serve as secondary endpoints. Table **4** provides a breakdown of the measure's subtests internal reliabilities (derived from U.S. data, n = 800) with Cronbach's alphas for the four subtests ranging from 0.78 to 0.84 across the ages three years and zero months to four years and 11 months.

Table 4: CELF Preschool assessment correlations between subtests and the Core Language Standard Score (Wiig et al., 2006)

	Sentence Structure	Word Structure	Expressive Vocabulary	Concepts & Following Directions
Sentence Structure				
Word Structure	0.53			
Expressive Vocabulary	0.58	0.62		
Concepts & Following Directions	0.61	0.58	0.61	
Core Language Standard Score	0.83	0.84	0.86	0.71

[5] See: http://www.pearsonclinical.co.uk/Psychology/ChildCognitionNeuropsychologyandLanguage/ChildLanguage/CELF-Preschool2UK/CELF-Preschool2UK.aspx

Independent test administrators (including researchers, teachers, teaching assistants, and independent early years staff) conducted the assessments pre- (before randomisation) and post-intervention. In total, 38 test administrators (n = 21 pre-testing, n = 20 post-testing) supported the trial, with three administrators involved with both pre- and post-testing. Each test administrator was required to attend mandatory training led by the evaluation team, was blinded to the nursery's random allocation for post-intervention testing, and underwent quality assurance (further information below).

The evaluation team liaised with nursery staff to arrange suitable dates for the test administrators to visit the nursery to assess the children. Two visits per nursery were initially scheduled on days when most children would be present. There were a small number of children for whom there was relevant parental agreement to participate but were not present at the nursery on the day(s) the test administrator visited (for example, due to sickness, holidays); 45% of nurseries (equating to 10% additional days of assessment time) received a third visit to ensure that the maximum number of children possible were assessed.

For pre-testing, test administrators were primarily recruited from the areas where nurseries were based through a supply teacher agency and training was held in two venues across the country. Due to difficulties with the availability and reliability of test administrators recruited through the agency, a different approach was taken for post-testing: post-testing was conducted by test administrators recruited by the evaluation team in Durham. Training was, therefore, also undertaken in Durham. The training at both pre- and post-testing periods included background to the CELF Preschool assessment, how to administer it, practice using the assessment, background to the project, reporting processes, data protection, safeguarding, quality assurance audio recording procedure, practice assessment requirements for quality assurance purposes, and who to contact with any questions or issues during delivery of the assessments.

All test administrators were audio recorded during the delivery of the assessments for quality assurance (QA) purposes, provided parents had given consent for the recording (n = 45 parents did not consent to audio recording; n = 88 parents from seven nurseries did not receive the opt-out form from the nursery and hence the children were not recorded). The evaluation team listened to the audio recordings of 10% of the assessments carried out by each test administrator to make a judgement of quality of delivery and marking and to ensure that assessments were conducted in line with the requirements detailed in their training. For the three test administrators who conducted both pre- and post-testing, their assessments were quality assured only at the pre-testing stage. This QA process was different to that described in the original protocol where test administrators would have been observed by members of the evaluation team while delivering the assessment. Audio recording was deemed more thorough given that all assessments (where consent was given) could be recorded reducing the risk of a child's responses being affected by an observer being present.

Secondary Outcomes

CELF-Preschool 2 UK norm-referenced subtest scaled scores were used as secondary outcomes from the trial.

Self-regulation and behaviour scores were gathered using the Child Self-regulation and Behaviour Questionnaire (CSBQ; Howard & Melhuish, 2016). The CSBQ is a 34-item questionnaire pertaining to children's everyday behaviours related to social and emotional development and self-regulation. CSBQ measures:

- self-regulation:
    - cognitive—items 5, 6, 8, 12, 18;
    - emotional—items 2, 10, 11 (reverse scored), 14 (reverse scored), 23 (reverse scored), 26 (reversed); and
    - behavioural—items 7 (reverse scored), 13, 15, 29 (reverse scored), 30, 31 (reverse scored);
- sociability—items 1, 4, 9, 16 (reverse scored), 22 (reverse scored), 27, 32;
- prosocial behaviour—items 15, 19, 24, 27, 30;
- externalising problems—items 3, 20, 23, 26, 28; and
- internalising problems—items 17, 21, 25, 33, 34.

For each item, the test administrator was asked to evaluate the child's frequency of target behaviours on a five-point scale (1 = 'not true' to 5 = 'very true'). The seven subscale scores are obtained by taking the average of the component item scores (first reversing any relevant score). Higher scores for all subscales indicate a more favourable outcome. No

total score is available. Cronbach's alphas for each subscale are as follows: sociability = 0.74, internalizing = 0.78, emotional self-regulation = 0.83, cognitive self-regulation = 0.87, externalizing = 0.88, prosocial = 0.89, and behavioural self-regulation = 0.89.

The assessment was completed by the children's class teacher for all children pre- and post-intervention for those with relevant parental agreement; therefore, the outcome (post-test) assessment was not conducted blind to group allocation.

The Home Observation Measure of the Environment (HOME) was chosen as it provides a descriptive profile which yields a systematic assessment of a child's home environment to measure (within a naturalistic context) the quality and quantity of stimulation and support available to the child. It does this by measuring the active receipt of inputs from objects, events, and transactions occurring within the home environment (Caldwell, 1994). Totsika and Sylva (2004) note the strongest advantage of HOME is the correlation it has to measures of cognitive development. The HOME inventory was included as a secondary outcome for this trial to investigate if, how, and why the intervention changed the home learning environment.

The 'Early Childhood HOME' is suitable for evaluating the appropriateness of the home environment for children aged three to six years. The completion of HOME involves a visit by a researcher to the child's home, and includes an interview with the main caregiver and observation of the home environment whilst the child is present and awake. It is made up of 55 items that are grouped in eight different subscales that are scored in a binary manner (yes/no). Six of the eight subscales are relevant to the learning aims of the evaluation. These are:

1. learning materials;
2. language stimulation (between child and caregiver);
3. responsivity (verbal interaction between child and caregiver);
4. academic stimulation;
5. modelling; and
6. variety of activities and parental interaction.

The HOME inventory was intended for use in a subsample of 50 households (25 in both the intervention and control group). All parents whose children had completed pre-testing (n = 1205) were invited to participate in the HOME interviews. Fifty-one parents replied agreeing to participate and all were interviewed pre-intervention (n = 26 intervention, n = 25 control) and 48 households post-intervention (n = 23 intervention, n = 25 control). The group was self-selected and therefore may not be representative of the total sample. Pre-intervention HOME visits for both the intervention and control groups were conducted in the immediate weeks following randomisation. Parents who had already provided agreement for their child to participate in attainment measures as part of the evaluation were invited to take part in the HOME visits. These visits were conducted by trained researchers. As an incentive for parents participating in the HOME visit, they were provided with a £50 Love2Shop gift voucher after the final visit.

Additional pupil-level data will be collected as part of the evaluation to enable the long-term tracking of the EasyPeasy intervention through the NPD and, where available, the Early Years Foundation Stage Profile (EYFSP) scores. Data from the EYFSP will be available in November 2019 and an addendum to the report will be published in spring 2020.

Sample size

Protocol

In the original protocol we aimed to recruit 120 school nurseries and 13 children per nursery, which would allow for some attrition (15% was expected). The sample size calculations included here are those included in the amended, and final, protocol.

Due to issues with recruitment (some nurseries being unable to recruit enough parents, or return required paperwork) and pre-testing causing some nurseries to withdraw, there was concern about not reaching the number of nurseries specified in the original protocol. Discussions were therefore held during the recruitment and pre-test period between the EEF, the EasyPeasy team, and the evaluation team about the sample size calculation assumptions and the expected effect size as well as what would be an acceptable number of nurseries. The sample size calculations included here

were completed by the EEF using parameters discussed with the evaluation team. The sample size calculation was subsequently reviewed by, and agreed with, the trial statistician. We made the following assumptions: equal allocation; two-sided alpha of 0.05; an intra-cluster correlation of 0.11; ten children per nursery whose parents provide agreement to participate; and the proportion of variance explained by covariates at the individual level (Level 1, pupil level) to be 0.25 and at the cluster level to be 0.16 (Level 2, setting level). Based on 102 nurseries (the number randomised and still in the trial at the time of discussions), we would have 80% power to show a difference of 0.224 of an effect size between the control and intervention groups for the primary outcome.

The sample size calculations were conducted using MDES Calculator for Two-Level Cluster Random Assignment Design (CRA2_2) Treatment at Level 2 by the PowerUp tool.[6]

CELF pre-test data was available for 1205 pupils across all 102 randomised nurseries (average 11.8 per nursery). With this sample size—and making the following assumptions: equal allocation, two-sided alpha of 0.05, an intra-cluster correlation of 0.11, and the proportion of variance explained by covariates at the individual level (Level 1, pupil level) to be 0.25 and at the cluster level to be 0.16 (Level 2, setting level)—the minimum detectable effect size for the primary outcome with 80% power is 0.216.

Analysis

Ultimately, 1128 pupils from 101 nurseries were included in the primary analysis (note: 102 nurseries completed pre-testing but one control nursery closed immediately following pre-testing in Dec 2017 and pupils moved to two intervention nurseries, thereby remaining in the trial). The observed ICC at the nursery level obtained from the primary analysis model was 0.18 (95% CI: 0.14 to 0.24). The proportion of variance explained by covariates at the pupil level was 0.52 and at the cluster level was 0.54 (Level 2, setting level). Therefore, the estimated minimum detectable effect size for the primary analysis was 0.197 (Table **5**).

Table 5: Power and sample size calculations (using the PowerUp tool) Model 3.1: MDES Calculator for Two-Level Cluster Random Assignment Design (CRA2_2)—Treatment at Level 2

Assumptions		Comments
Alpha Level (α)	0.05	Probability of a Type I error
Two-tailed or one-tailed test?	2	
Power (1-β)	0.80	Statistical power (1-probability of a Type II error)
Rho (ICC)	0.18	Proportion of variance in outcome that is between clusters
P	0.50	Proportion of Level 2 units randomised to treatment: J_T / (J_T + J_C)
R_1^2	0.52	Proportion of variance in Level 1 outcomes explained by Level 1 covariates
R_2^2	0.54	Proportion of variance in Level 2 outcome explained by Level 2 covariates
g*	2	Number of Level 2 covariates
n (Average Cluster Size)	10	Mean number of Level 1 units per Level 2 cluster (harmonic mean recommended)
J (Sample Size—no. of Clusters)	101	Number of Level 2 units
M (Multiplier)	2.83	Computed from T_1 and T_2
T_1 (Precision)	1.98	Determined from alpha level, given two-tailed or one-tailed test
T_2 (Power)	0.85	Determined from given power level
MDES	**0.197**	Minimum Detectable Effect Size

[6] http://web.missouri.edu/~dongn/PowerUp.htm

Randomisation

Randomisation was undertaken by the trial statistician who had no involvement in the recruitment of nurseries and was blinded to their identity (settings were identified by a unique trial identifier only). Nurseries were randomly allocated 1:1 using minimisation (via software MinimPy; Saghaei and Saghaei, 2011) based on eligible cohort size as measured by the total number of children eligible to participate per nursery. Nursery cohort size was dichotomised for the minimisation as <14/≥14 as 14 was the median number of pupils eligible to participate per nursery in the first batch of nurseries randomised.

Randomisation was conducted in two batches due to delays in collecting pre-test data from nurseries. The first took place on 19 December 2017 and included 44 nurseries that had completed (or mostly completed) the CELF pre-test. Where some CELF assessments still needed to be completed, these nurseries were not informed of their randomisation until after the assessments were completed (8/44 nurseries). The second batch of randomisation took place on 23 January 2018 and included the remaining 58 nurseries, of which 57 had completed the CELF assessment. One nursery still had some CELF assessments to be completed and it was informed of its allocation the following day after completion of the assessment.

Since the trial statistician undertook the randomisation and conducted the statistical analysis, the analysis was not undertaken blinded to allocation.

Statistical analysis

Analysis was conducted in Stata v15 using the principles of intention to treat, where data was available, including all nurseries and pupils in the groups to which they were randomised irrespective of whether or not they actually received the intervention.

Statistical significance has been assessed using two-sided tests at the 5% level. Estimates of effect with 95% confidence intervals (CIs) and p-values are provided.

A CONSORT diagram has been produced, and discussed, to show the flow of nurseries and pupils through the trial.

School and pupil characteristics and measures of prior attainment are summarised descriptively by randomised group, both as randomised and as included in the primary outcome analysis. No formal statistical comparisons were undertaken on baseline data (Senn, 1994). The unadjusted difference between groups on the pre-tests is reported as a Hedges' g effect size with 95% CI. Histograms of pre- and post-test Core Language Standard Score data are presented.

Primary outcome analysis

The primary objective of this study was to investigate the effectiveness of the intervention on the language and communication skills of the participating children, assessed via the Core Language Standard Score of the CELF Preschool assessment. Unadjusted raw and scaled outcome scores are summarised by trial arm. The correlations between pre- and post-intervention scores are provided, for both the individual subscale scores and the total Core Language Standard Score.

The difference in attainment between pupils in the intervention group and those in the control group was compared using a multilevel mixed-effect linear regression model at the pupil level with Core Language Standard Score as the response variable. Group allocation, baseline Core Language Standard Score, and the number of children with parental agreement to participate within the nursery (minimisation factor, in its continuous form) were included as fixed effects in the model.

Pupil-level fixed effects:

- baseline Core Language Standard Score (continuous).

Nursery-level fixed effects:

- number of children with parental agreement to participate (continuous).

Adjustment was made for clustering at the nursery level by including nursery as a random effect.

Model equation:

$$Y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 w_i + \beta_3 I_{Ai} + u_{ij}\gamma_i + \varepsilon_{ij}$$

Y_{ij} = response of the j-th member of cluster (nursery) i, *i=1,…,m, j=1,…,n_i*

m = number of clusters (nurseries)

n_i = size of cluster (nursery) i

x_{ij} = Baseline Core Language Standard Score for j-th member of cluster (nursery) i

w_i = number of eligible pupils for cluster (nursery) i

I_{Ai} = indicator variable for group allocation of cluster (nursery) i (0=Control, 1=Intervention)

β_0, β_1, β_2, β_3 = fixed effect parameters

u_{ij} = random effect for j-th member of cluster (nursery) i

γ_i = random effect parameter

ε_{ij} = residual error term for j-th member of cluster (nursery) i

Model assumptions were checked as follows: the normality of the standardised residuals were checked using a histogram and qq plot, and the homoscedasticity of the residuals assessed using a scatter plot of fitted values against the residuals.

Missing data

The amount of missing baseline and outcome data is summarised, and reasons for missing data were explored and are provided in the report, where available. A multilevel mixed-effect logistic regression model was run to assess for statistically significant predictors of missing primary outcome data at the pupil level, including pre-test CELF Core Language Standard Score, gender, ethnicity, EYPP status, EAL, SEND, and age. At the nursery level it included allocation, number of eligible pupils, LA, percentage of pupils eligible for EYPP, and number of pupils due to start school in 2018, and nursery as a random effect. Significant predictors and possible mechanisms for the missing data are discussed.

The impact of missing data on the primary analysis was additionally assessed using multiple imputation by chained equations, predicted by pre-test total Core Language Standard Score, nursery, allocation, number of eligible pupils, and any variables found to be significant in the 'drop-out' model described above.

A 'burn-in' of 10 was used and 20 imputed datasets were created. The primary analysis was then rerun within the imputed datasets and Rubin's rules were used to combine the multiply imputed estimates.

Compliance

The EasyPeasy programme was delivered across multiple levels:

- the EasyPeasy team delivered the intervention via website application, training, and text message to parents;
- the Pod Leader at the school attended webinar training and signed parents up to the programme, monitored and supported parent engagement throughout the programme through Pod management (using their web-based platform), and also potentially used the EasyPeasy games with the nursery children in their setting;

- the parents of children in the school nursery received the text messages and could then access the games and information provided on the web-app, play the games with their child, and interact with the Pod community (of other parents and the nursery staff on the web-app and face-to-face in the setting); and
- the child played the EasyPeasy games with their parent.

In this trial, aspects of compliance and fidelity arose from four levels as described below:

At the top level of the technical delivery, EasyPeasy provided a report detailing any technical failures with the web-app or the sending of text messages to parents throughout the programme. The number, frequency, and type of technical issues experienced is summarised.

At the Pod Leader level, EasyPeasy provided the evaluation team with details of which Pod Leaders attended the training. We also collected data through the Pod Leader practitioner survey on the amount of time Pod Leaders spent supporting delivery of EasyPeasy. This data is summarised in the cost section of the impact analysis.

At the parent level, EasyPeasy provided information on the level of parental drop-out (number of parents who asked to be removed from the EasyPeasy programme) across the project, which is reported to give an insight into how many parents did not engage at all with the programme. Jelley et al. (2016) reported that 3% of parents chose to opt out of receiving the intervention during the EasyPeasy pilot study. The number and proportion of parents who withdraw from the intervention are reported.

It was not possible to measure whether parents actually played the games with their children; however, this was explored during the parental focus groups which took place as part of the case studies in six settings (see Parental Engagement section of Implementation and Process Evaluation Analysis).

A Complier Average Causal Effect (CACE) analysis for the primary outcome was conducted to account for Pod Leader attendance at the live webinar training. This variable was measured at the nursery level (that is, 'Did the nursery's Pod Leader attend the live webinar training in person or not?'). A two-stage least-squares instrumental variable approach was used, with random group allocation as the instrumental variable (Dunn, 2005). Compliance was defined as a dichotomous variable: 0 (Pod Leader did not attend the live webinar training) and 1 (Pod Leader attended the live webinar training).

Additional compliance analysis

One nursery from the control group closed down post-randomisation and the ten pupils from that nursery attended one of two nurseries from the intervention group. As a sensitivity analysis, a second CACE analysis was conducted on the primary outcome considering these pupils in the nurseries they moved to after their original nursery closed.

Post-hoc sensitivity analysis: assessor quality

The quality of each assessor was graded by listening to 10% of audio recordings from each assessor. The grading scale was 'very good', 'good', 'acceptable', or 'not acceptable', based on agreed criteria for how each assessor adhered to the assessment scoring and administration instructions. Post-hoc sensitivity analyses were conducted in which the primary ITT analysis was repeated with an additional covariate for the grading of the assessor (at pre-test, and at post-test) and including a random effect for assessor (within which 'nursery' is nested since the same assessor tested all children within a nursery). Whilst reviewing the final report, the EEF also suggested post-hoc analyses excluding the results of pupils tested by an assessor who was graded as 'not acceptable', so this was added.

Secondary outcome analysis

The secondary outcomes of the individual CELF Preschool assessment scaled subscale scores and the seven CSBQ subscales were summarised and analysed exactly as described for the primary outcome of the Core Language Standard Score, adjusting for the associated pre-test subscale score instead of the Core Language Standard Score pre-test score. The raw CELF Preschool assessment scores are summarised, but were not analysed.

A similar approach was taken for the summary and analysis of the HOME total and subscale scores, except the analysis models were not adjusted for the associated pre-test HOME score since this was measured post-randomisation. Sensitivity analysis adjusting for pre-test HOME score was undertaken to assess the impact of this on the results.

The reliability of the HOME scale made up of the six subscale scores was investigated using Cronbach's alpha as a measure of internal consistency.

Intra-cluster correlations (ICCs)

The intra-cluster correlation coefficient (ICC) associated with nursery for the post-test outcomes has been extracted from each multilevel analysis model, with the 95% CI. The ICC associated with the nursery for the pre-test scores is also presented with a 95% CI. The pre and post-test ICCs have been calculated for the CELF (scaled scores) and CSBQ outcomes, but not for HOME due to the limited sample size and because households from across the whole sample took part in the HOME assessment so there are very few per nursery and so we would expect only a negligible ICC.

Subgroup analyses

Subgroup analyses for the primary outcome based on pupil eligibility for the Early Years Pupil Premium (EYPP), English as an Additional Language (EAL), and gender (separately) were undertaken via the inclusion of the variable and an interaction term between the variable and group allocation in the primary analysis model.

Effect size calculation

Effect sizes and their 95% CI were calculated by dividing the adjusted mean difference (and upper and lower 95% confidence limits) between the intervention and control group by the total variance (between and also within nursery variance) obtained from a multilevel model without covariates.

Implementation and process evaluation

The implementation and process evaluation research questions were developed by the evaluation team following the IDEAS workshop with the project delivery team. The intention of these questions was to enable the evaluation to track change in relation to the anticipated impact of EasyPeasy identified in the Theory of Change above. As detailed in Table **6**, the research questions also relate specifically to Humphrey's identified dimensions within a process evaluation (Humphrey et al., 2016).

Research question 9—What areas of the programme could be further developed following completion of the project?—is not directly aligned with an implementation dimension, instead promoting a review of the available evidence to make recommendations for future implementation of the programme.

Table 6: Implementation and Process Evaluation research questions and dimensions

	Research Question	Dimensions of Implementation
1.	To what extent are the settings and parents engaging with, and delivering, the intervention?	Fidelity Dosage Reach
2.	How is the intervention being disseminated to parents through settings and is the method effective? (a) Are nominated staff engaging with EasyPeasy and using it as a mechanism to foster communication and engagement between parents? (b) Are parents engaging with EasyPeasy and implementing the suggested activities within the home?	Fidelity Adaptation Dosage
3.	Can all parents access and engage with the EasyPeasy intervention (e.g. from all socio-economic backgrounds, in EAL families, in families with SEND children)?	Fidelity Dosage Reach Adaptation
4.	What are the different stakeholder viewpoints on the intervention— (a) those of setting practitioners? (b) parents?	Participant responsiveness Programme differentiation
5.	How effective and appropriately pitched are the activities— (a) for parents to deliver? (b) for children to receive?	Quality
6.	What is the acceptability of the intervention to parents and does this differ depending on the socio-economic status of the parents (based on their child's eligibility for Early Years Pupil Premium), their child's Special Education Needs Diagnosis, or English as an Additional Language status?	Reach
7.	What are the key success factors required for the EasyPeasy intervention to work well?	Quality Monitoring of control/comparison conditions
8.	What are the barriers to successful delivery of the intervention— (a) for parents? (b) for setting practitioners?	Quality Monitoring of control/comparison conditions
9.	Based on the findings of the trial, what areas of the EasyPeasy programme could be further developed following completion of the project?	N/A

In developing the IPE methodology, the evaluation team took the research questions listed above and considered how best to collect evidence in relation to each. The process evaluation took place over the full intervention year, and evidence was collected using the methods described below.

Tracking Data

Tracking fidelity data was collected by EasyPeasy; this data formed some evidence of the extent to which settings, and parents in the intervention group, received and adhered to the intervention and their engagement with the programme. This data included:

- technical delivery—data collected from EasyPeasy about the number of web-app outages or technical failures in sending text messages to parents;
- Pod Leader webinar training attendance—data collected by EasyPeasy for the number of Pod Leaders joining the online webinar training session; and
- overall parental drop-out—data collected by EasyPeasy as to the number of parents actively requesting to be removed from the contact list for receiving any further messages from EasyPeasy.

Parent-level information on engagement or withdrawal from the intervention was not available from EasyPeasy. Information on the number of playdates per week within a Pod (reported by Pod Leaders in the Pod Leader practitioner

survey at the end of the intervention) has been used as a proxy for parent engagement with the intervention. The findings are reported in the Engaging Parents section of the Implementation and Process Evaluation analysis section.

Usual practice survey

Prior to randomisation, each school was asked to identify a key member of staff (known within the context of the project as the 'Pod Leader') to be responsible for coordinating the evaluation within their school and delivering the intervention should they be randomly allocated to receive it. These staff members were asked to complete a paper-based usual practice survey prior to randomisation and post-intervention (see Additional Appendices). This was to capture current provision of parental engagement and school readiness activities, including any associated costs. Quantitative data was analysed using descriptive statistical analyses. Qualitative data was used to provide context where appropriate for the quantitative findings.

Pod Leader practitioner survey

All Pod Leaders in the intervention group were requested to complete a survey (delivered online and paper-based) post-intervention (see Additional Appendices). This was to investigate their experiences of the training, support, and delivery of EasyPeasy, any noticeable changes to parental engagement (if any), and the overall impact (if any) that EasyPeasy had on children's learning and development, parental engagement, and general practice within their setting. Quantitative data was analysed using descriptive statistical analyses. Qualitative data was open coded and thematically analysed; this was then considered in relation to each of the research questions. Findings were triangulated with those from the other data from the implementation and process evaluation.

Case studies: Pod Leader interviews and parent focus groups

Case studies were conducted by the Research Associate for the project in 10% of intervention school nurseries (n = 6). These were grouped into longitudinal, cross-sectional, and best practices case studies; these are detailed further below in Table **7** including the frequency and type of contact. The purpose of the longitudinal case studies was to capture Pod Leaders' views over the course of the intervention, along with parents' views at the end of the intervention. Cross-sectional case studies took a snapshot of activity relating to EasyPeasy at a single point in time providing additional context and evidence that the longitudinal nurseries were similar to other nurseries in the sample. The longitudinal and cross-sectional nurseries were chosen at random by the independent trial statistician who was not involved in the recruitment of nurseries. Stata (v15) was used to randomly select eight nurseries from the 51 intervention nurseries; these were then randomly allocated to the two types of case study or to be nursery reserves for these. Two nurseries were selected for the longitudinal case studies with two reserve nurseries, and two nurseries for the cross sectional case studies along with two reserve nurseries. The characteristics of each participating nursery are detailed in Table **7**.

Two best-practice case-study school nurseries were identified towards the end of the intervention. The two best-practice nurseries were chosen and recruited to participate in the case studies by the EasyPeasy delivery team who viewed their delivery of the EasyPeasy programme to be successful using two criteria in their selection: (1) the enthusiasm of the Pod Leader and (2) high parental engagement (assessed by EasyPeasy from their data on the number of weekly playdates within Pods and anecdotal evidence from discussions between EasyPeasy and schools).

During the case studies, the evaluation team interviewed the Pod Leader; written agreement to participate was gained from each of these individuals (see Additional Appendices) and interviews were recorded for transcription purposes. Interviews lasted approximately 30 minutes and followed the relevant interview schedule (see Additional Appendices). The Pod Leader interviews provided an understanding of how nurseries deliver and monitor the intervention, any barriers, strategies for keeping parents engaged, and to gauge potential impact, if any, of an improved home learning environment via EasyPeasy on children's learning within the nursery. Understanding the Pod Leaders use of the weekly report made available by EasyPeasy was also discussed during these interviews. The interview provided (a) an opportunity to gauge how, if at all, the Pod Leader encouraged parents perceived as being inactive to re-engage, (b) how they acknowledged, if at all, parents perceived as being very active, and (c) to assess how Pod engagement might be improved, if applicable. Case-study transcripts were analysed in NVIVO[7] software using open coding and thematic

[7] https://www.qsrinternational.com/nvivo/what-is-nvivo

analysis. The findings were then considered in relation to each of the research questions. Findings were triangulated with those from the other data from the implementation and process evaluation.

Table 7: Timeline and selection methods for the case studies

School Nursery	Case study allocation	Timepoint of Pod Leader interviews	Pod Leader background	Selection method
A	Longitudinal	Telephone interview—start; face-to-face interview—end.	Full time nursery teacher	Nurseries selected at random by statistician
B	Longitudinal		Full time reception teacher	
C	Cross-sectional	Face-to-face interview—mid-way.	Full time nursery teacher	
D	Cross-sectional		Part time nursery teacher	
E	Best practice	Face-to-face interview—end.	Assistant Headteacher, Early Years Leader	Nurseries identified by EasyPeasy
F	Best practice		Deputy Head Early Years Leader	

During visits to case-study nurseries, the evaluation team also conducted a focus group with parents who engaged (at any level) with the EasyPeasy intervention (discussed further below). The frequency, time-points, and participation of each focus group is listed in Table **8**. Parents were recruited to focus groups using an information sheet (see Additional Appendices) developed by the evaluation team and distributed by staff at each setting. Written parental agreement to participate was gained prior to the focus group (see Additional Appendices). Each focus group took place within the nursery setting, was recorded for transcription purposes, guided using a semi-structured interview schedule (see Additional Appendices), and lasted approximately 45 minutes. Each parent received a £15 Love2Shop voucher for taking part.

The focus groups explored acceptability and delivery of EasyPeasy and how it was used and adapted in individual circumstances, including any barriers that were faced. The focus groups also explored if the programme had an impact on parents' understanding of child development and parents' attitudes and engagement with their child and their child's learning. The focus group aimed to gauge any changes to the level of parent-school interactions as a result of implementing the EasyPeasy programme.

Table 8: Case study parent focus group participation

School Nursery	Case study allocation	Time point	Number of participants	Date of focus group
A	Longitudinal	End	5	06/07/18
B	Longitudinal	End	0	The researcher attended the nursery twice to conduct the focus groups, but no parents attended on either occasion.
C	Cross-sectional	Mid-way	5	27/04/18
D	Cross-sectional	Mid-way	5	03/05/18
E	Best practice	End	4	01/07/18
F	Best practice	End	6	09/07/18

Developer Interview

A face-to-face interview with the CEO of EasyPeasy was conducted by the PI of the project and Research Associate. The interview took place post-intervention during August 2018 and was recorded for transcription purposes. This interview aimed to provide a deeper understanding of the delivery of EasyPeasy from the developer's perspective, its perceptions of the programme's impact, barriers during the implementation, cost, and to gain a picture of future developments planned.

Costs

Information on the costs of implementing EasyPeasy were collected via the Pod Leader practitioner survey collected immediately following the intervention period during case-study interviews with the Pod Leaders and during the interview with the delivery team. Costs for the purchase of EasyPeasy have been sourced from the EasyPeasy website and from

discussion with the developer for the costs in the third year of the per-pupil estimate. EasyPeasy publish the cost of purchasing the product on their website;[8] prices are dependent upon school size. We have made three assumptions in our calculation of cost per pupil:

- the cost estimate does not include any discounts available to schools that have participated in the trial (a school that has participated in the present trial receives a 10% discount for the 2018/2019 academic year; in subsequent years, schools will pay standard prices);
- the school nursery is of an average size of 31 children (calculated from the average number of pupils in a setting that nurseries participating in the trial reported would be moving on to Reception in the following year); and
- the annual licence fee does not increase from the value in 2018/2019.

Timeline

Table **9** below shows a timeline of activities related to the evaluation and intervention delivery.

Table 9: Timeline of the EasyPeasy randomised controlled efficacy trial

Date	Activity
1 Jun–31 Aug 2018	Recruitment of Local Authorities.
31 Aug–31 Dec 2018	Recruitment events to facilitate recruitment of schools; 134 school nurseries recruited with MoUs signed. Recruitment of parents within school nurseries; 102 school nurseries with signed MoUs returned completed welcome packs; • collection of parental participation agreement forms per nursery class (minimum 10 per nursery); • collection of CSBQs completed by nursery staff; • collection of Usual Practice Surveys by nursery staff; and • collection of pupil demographic information. Test administrator recruitment through a supply teacher agency to conduct CELF pre-assessments. Recruitment of additional test administrators to conduct CELF pre-assessments.
30 Nov–1 Dec 2017	Training conducted for agency supply teachers and additional test administrators.
4–18 Dec 2017	Pre-intervention CELF assessments conducted by test administrators for Batch 1 nurseries (n = 44).
8 Jan 2018	Randomisation—Batch 1 nurseries informed of their allocation.
9–10 Jan 2018	Training conducted for agency supply teachers and additional test administrators.
8–26 Jan 2018	Pre-intervention CELF assessments conducted by test administrators for Batch 2 nurseries (n = 58).
26 Jan 2018	Randomisation—Batch 2 nurseries informed of their allocation.
10 Aug 2018	Detailed statistical analysis plan produced and submitted to the EEF; amended SAP submitted 17 Aug 2018.
11 Dec 2017–31 Jan 2018	51 HOME visits (25 control group and 26 intervention group) conducted post-randomisation and pre-intervention.
4–12 Jan 2018	Pod Leader training.
13 Jan 2018–1 Jun 2018	Intervention (20 weeks): - Batch 1 nurseries—Jan–May 2018; - Batch 2 nurseries—Feb–June 2018. Quality assurance of pre-intervention CELF assessment test administrators using audio recordings.
28 Feb 18–2 Mar 2018	Longitudinal case study—two school nurseries in intervention group participated in a telephone interview.
27 Apr 2018–3 May 2018	Cross-sectional case study—two school nurseries in the intervention group participated in an in-school face-to-face interview mid-way through the programme.

[8] https://easypeasyapp.com/pricing/

1 Apr 2018–3 Jun 2018	Test administrator recruitment and training to conduct post-intervention tests; training of administrators took place on 4–5 June 2018.
11–29 Jun 2018	Post-Intervention CELF assessments conducted by test administrators for Batch 1 nurseries (35). Quality assurance of test administrators—using an audio-recorded practice session conducted before assessments were carried out in schools.
2–20 Jul 2018	Post-intervention CELF assessments conducted by test administrators for Batch 2 nurseries (65). Quality assurance of test administrators—using an audio-recorded practice session conducted before assessments were carried out in schools.
11 Jun 2018–22 Jul 2018	Best practice case study—two school nurseries in intervention group participated in face-to-face interviews at the end of the trial. Longitudinal case study—Pod Leaders in two school nurseries in the intervention group participated in a follow up face-to-face interview; parents in one of these schools participated in a focus group. No parents in the second school attended the focus group on two dates. 48 HOME visits (25 control group and 23 intervention group).
17 Jul 2018	Revised protocol produced and submitted to the EEF.
6 Aug 2018	Interview with the developers.
16 Jul 2018–30 Nov 2018	Collection of children's school destination data.
15 Jul 2018–30 Nov 2018	Quality assurance of post-intervention CELF assessment test administrators using audio recordings. Collate data, analysis, and report writing.
30 Nov 2018	Draft report submitted to the EEF.
15 Dec 2018	Collate school destination data for NPD submission.
31 Apr 2020	Early Years Foundation Stage Profile (EYFSP) report (planned).

Impact evaluation

Participant flow including losses and exclusions

School and pupil recruitment

A total of 341 schools were approached to participate in the evaluation. Recruitment of schools was stopped once the target sample of 120 schools, plus approximately 10%, had been signed up (134 schools—Figure 1). These 134 schools were assessed for eligibility. All 134 schools were potentially eligible for the study but 32 withdrew their interest or were withdrawn by the evaluation team prior to randomisation for the following reasons:

- too few parent participation agreement forms returned (we originally expected 13 per nursery, but as this was difficult for nurseries to recruit that many parents, we agreed on a minimum of ten and the protocol was updated), n = 18;
- missed the deadline to return paperwork, n = 8;
- nursery too busy or issues with staffing levels, n = 5; and
- nursery closed, n = 1.

Therefore, 102 nurseries participated in the trial. Within these 102 nurseries, a total of 1488 pupils were eligible (per nursery: median 14.6, SD 4.5, median 13, range 5 to 36).

Pre- and post-testing with the primary outcome

Of the 1488 pupils, 1205 (81.0%) across all 102 nurseries were assessed using the CELF Preschool assessment at pre-test prior to their nursery being informed of their random allocation (per nursery: median 11.8, SD 3.3, median 12, range 4 to 25—Figure 1). Therefore, 283 pupils, from 88 of the nurseries, were not assessed using CELF for the following reasons:

- absent on day of testing, n = 253;
- refused to participate, n = 18;
- had left the nursery, n = 8;
- missed off the assessor's list in error, n = 2;
- and partial parental agreement to participate covering CSBQ only received, n = 2.

Up to three visits were carried out in nurseries to assess pupils who were unavailable on the previous assessment visit or where there had not been time to assess all pupils. Due to the nature of nursery attendance, not all pupils attend nursery five days a week. Common patterns of attendance are 2.5 days per week (Mon–Wed or Wed–Fri), five mornings a week, five afternoons a week, or five full days a week. Wherever possible, multiple visits to nurseries were arranged for a different time of the week to try to assess pupils with different attendance patterns. At randomisation—with the 1205 pupils and under the same assumptions as were used in the protocol—the MDES for the primary outcome analysis is estimated at 0.216 (Table 10).

Post-testing was only supposed to take place with children that had been pre-tested; therefore, post-testing with the CELF Preschool assessment was expected in 1205 pupils. There was a mix-up at some nurseries and a number of children were post-tested that were not pre-tested (19 pupils across 12 nurseries). Post-testing with the CELF Preschool assessment was undertaken with a total of 1149 pupils across 101 nurseries (intervention group, n = 51; control group, n = 50; pupils per nursery: median 11.4, SD 3.3, median 11, range 4 to 25). Of the 1205 pupils that were pre-tested, 1130 (93.8%) were post-tested. Reasons for missing post-test data (n = 75) among the 1205 children that were pre-tested were:

- absent on day of testing, n = 45
- nursery withdrew from trial, n = 15;
- had left the nursery, n = 9;

- refused to participate, n = 4; and
- missed off the assessor's list in error, n = 2.

Pre- and post-testing with the secondary outcomes

The CSBQ secondary outcome was completed by the teacher and so there was a higher response rate for this than the CELF primary outcome. At pre-test, valid CSBQ data was available for 1433 of the 1488 eligible pupils (96.3%). As with the CELF, post-testing was only supposed to take place with children that had been pre-tested; however, 24 pupils across three nurseries had CSBQ data at post-test but not at pre-test. Of the 1433 pupils with pre-test CSBQ data, 955 (66.6%) had post-test data.

The HOME outcome was implemented in 51 pupils' households at pre-test and 48 out of the 51 households at post-test.

Nursery attrition

Two nurseries withdrew during the course of the trial (2.0%), both in the control group: one due to the nursery closing (children from this nursery merged with two participating intervention schools and so post-testing could still be undertaken with these pupils) and one due to being too busy/staffing issues (no post-testing was carried out within this nursery).

Figure 1: Participant flow diagram for the primary outcome analysis (CELF Preschool assessment)

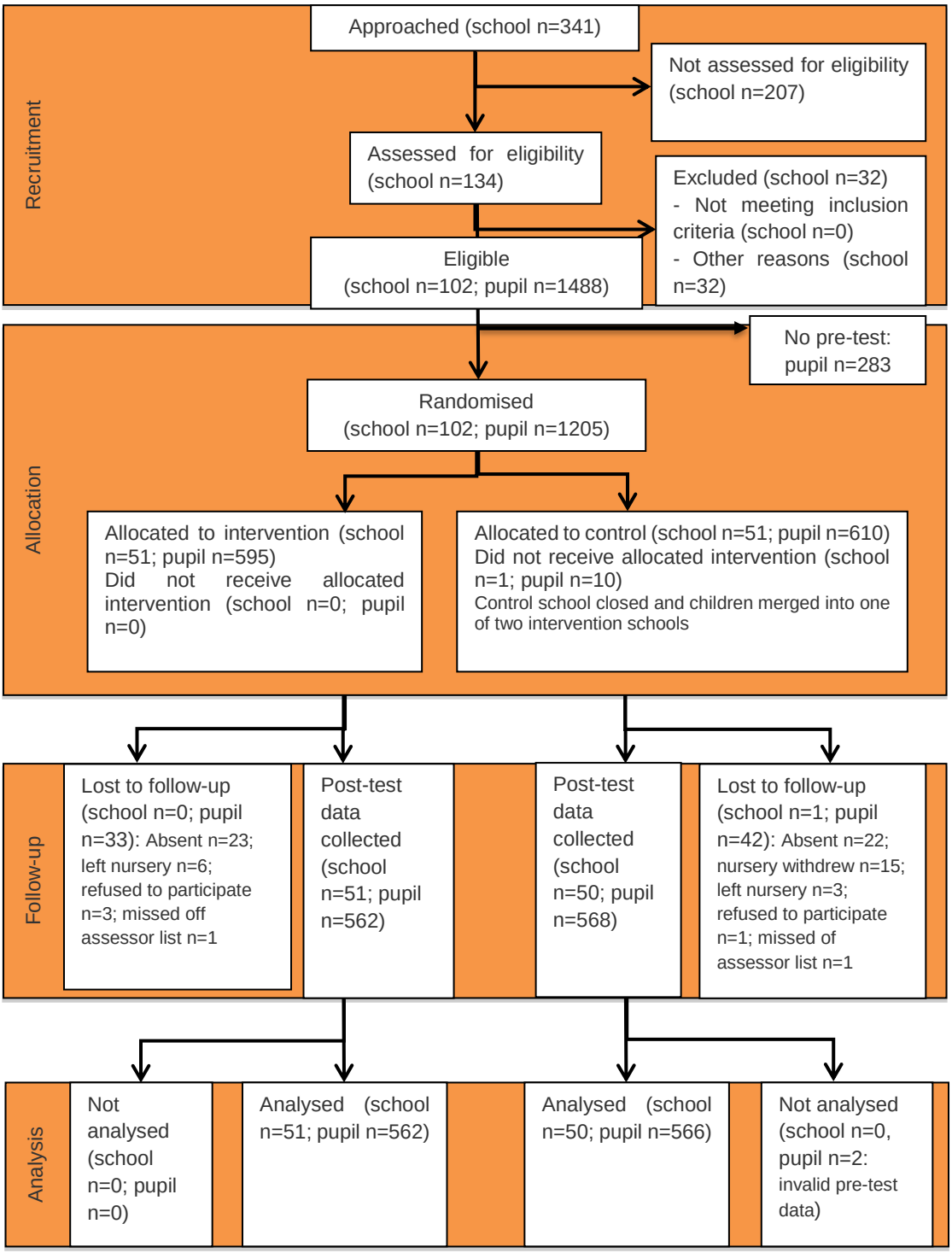


Table 10: Minimum detectable effect size for primary outcome analysis at different stages

		Protocol	Randomisation	Analysis
MDES		0.224	0.216	0.197
Pre-test/post-test correlations	level 1 (pupil)	0.25	0.25	0.52
	level 2 (school)	0.16	0.16	0.54
Intra-cluster correlations (ICCs)	level 2 (school)	0.11	0.11	0.18
Alpha		0.05	0.05	0.05
Power		0.8	0.8	0.8
One-sided or two-sided?		2	2	2
Average cluster size		10	11.8	10
Number of schools	intervention	51	51	51
	control	51	51	50
	total	102	102	101
Number of pupils	intervention	510	595	562
	control	510	610	566
	total	1020	1205	1128

Attrition

Overall, 1128 of the 1205 pupils were pre-tested using the CELF assessment—overall 93.6%; intervention, n = 562 (94.5%); control, n = 566 (92.8%)—had valid pre- and post-intervention CELF test data and so were included in the primary outcome analysis model (Figure **1**).

Pupil and school characteristics

Characteristics and measures of prior attainment for the 102 randomised nurseries and 1205 pupils with pre-test CELF data are presented in Table **11**, and are broadly similar between the intervention group and the control group. No formal hypothesis testing was performed on baseline data, so comparisons were made by eye only.

Most nurseries were located within the Doncaster, Knowsley, or Oldham Local Authorities in both groups. The mean percentage of pupils per school eligible for EYPP was 47.2% in both the intervention (SD 16.7) and control groups (SD 16.9).

Intervention group nurseries tended to have a slightly higher number of pupils starting school in 2018 (mean 34.2, SD 16.4) compared to control group nurseries (mean 29.5, SD 17.6).

There was no missing Local Authority data. There were six missing data points for the Pupil Premium eligibility percentage and eight missing data points for the number of pupils starting school at each nursery in 2018.

Pupils were aged, on average, around 45 months (3.75 years). Half the pupils were male (49.8% in the intervention group, and 48.4% in the control group), and nearly two-thirds were white (intervention group 66.6%; control group 65.7%). A considerably higher percentage of randomised pupils in the control group were eligible for EYPP, were EAL,

and had SEND, than in the intervention group. These differences remained when considering the groups as included in the primary analysis (Table **13**).

The mean (SD) pre-test Core Language Standard Score was slightly higher in the intervention group than the control group (intervention group 25.1 (9.0); control group 24.4 (8.3); effect size 0.08, 95% CI -0.03 to 0.19). A higher CELF score indicates a better outcome. Two fewer pupils have a scaled score (n = 1203) than had available CELF data (n = 1205) since one was too young for a scaled score to be computed for them, and one had missing data for all component subtests (data only valid for the Concepts subtest). The overall distribution of scores at pre- and post-test are presented in a histogram in Figure **2** and appear to be reasonably normally distributed (visual inspection only, no statistical testing performed).

The ICC (95% CI) associated with nursery for the pre-test CELF Core Language Standard Score and its four scaled subtest scores (in the order they appear in) are: 0.26 (95% CI 0.19 to 0.34); 0.22 (95% CI 0.15 to 0.28); 0.29 (95% CI 0.22 to 0.37); 0.17 (95% CI 0.11 to 0.23); and 0.27 (95% CI 0.20 to 0.34).

The ICC (95% CI) associated with nursery for the seven pre-test CSBQ subscale scores (in the order they appear in Table 11) are: 0.11 (95% CI 0.06 to 0.16); 0.10 (95% CI 0.06 to 0.15); 0.26 (95% CI 0.19 to 0.34); 0.17 (95% CI 0.11 to 0.23); 0.14 (95% CI 0.08 to 0.19); 0.15 (95% CI 0.10 to 0.21); and 0.13 (95% CI 0.08 to 0.18).

There was some imbalance in the HOME scores at pre-test between the intervention and control groups (Table **11**). For the total score, and for the subtests of Learning materials, Modelling, and Variety of activities and parental interaction, the average scores are higher in the intervention group than the control group (effect size >0.10). For the subtest of Responsivity, the difference favours the control group (effect size -0.23). These differences could have occurred by chance, perhaps due to the small sample size.

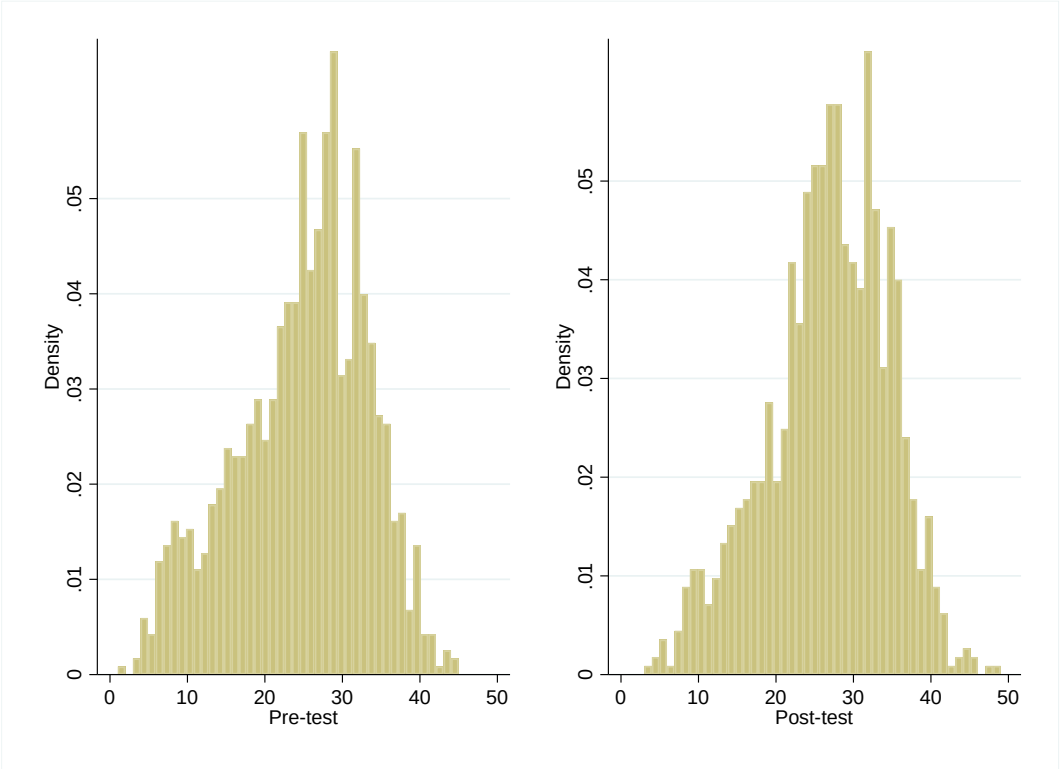
Table 11: Baseline comparison as randomised for the primary outcome (nursery n = 102; pupil n = 1205)

Nursery-level (categorical)	Intervention group		Control group	
	n/N (missing)	Count (%)[a]	n/N (missing)	Count (%)[a]
Local Authority Bedford Camden Coventry Doncaster Durham Islington Knowsley Luton Oldham	51/51 (0)	4 (7.8) 2 (3.9) 1 (2.0) 12 (23.5) 6 (11.8) 4 (7.8) 10 (19.6) 0 (0.0) 12 (23.5)	51/51 (0)	1 (2.0) 1 (2.0) 1 (2.0) 8 (15.7) 6 (11.8) 6 (11.8) 10 (19.6) 3 (5.9) 15 (29.4)
Nursery-level (continuous)	n (missing)	Mean (SD)	n (missing)	Mean (SD)
Percentage of pupils eligible for EYPP	47 (4)	47.2 (16.7)	49 (2)	47.2 (16.9)
Number of pupils starting school in 2018	46 (5)	34.2 (16.4)	48 (3)	29.5 (17.6)
Pupil-level (categorical)	n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Gender, male	589/595 (6)	293 (49.8)	601/610 (9)	291 (48.4)	
Ethnicity White Black/Caribbean Asian Middle Eastern Mixed Other Unknown	595/595 (0)	396 (66.6) 26 (4.4) 92 (15.5) 9 (1.5) 38 (6.4) 3 (0.5) 31 (5.2)	610/610 (0)	401 (65.7) 31 (5.1) 111 (18.2) 6 (1.0) 35 (5.7) 4 (0.7) 22 (3.6)	
Eligible for EYPP	512/595 (83)	74 (14.5)	489/610 (121)	138 (28.2)	
EAL	562/595 (33)	143 (25.4)	589/610 (21)	180 (30.6)	
SEND	560/595 (35)	20 (3.6)	577/610 (33)	41 (7.1)	
Pupil-level (continuous)	n (missing)	Mean (SD)	n (missing)	Mean (SD)	Effect Size
Age, months	595 (0)	45.9 (4.0)	610 (0)	45.4 (3.6)	N/A
CELF-Preschool 2 UK (scaled scores)					
Core Language Standard Score[b]	595 (0)	25.1 (9.0)	608 (2)	24.4 (8.3)	0.08 (-0.03, 0.19)
Sentence Structure	593 (2)	8.0 (3.3)	607 (3)	7.7 (3.1)	0.07 (0-.04, 0.18)
Word Structure	587 (8)	8.7 (3.6)	607 (3)	8.4 (3.4)	0.09 (0-.03, 0.20)
Expressive Vocabulary	586 (9)	8.8 (3.4)	605 (5)	8.4 (3.4)	0.11 (0.00, 0.23)
Concepts and Following Directions	550 (45)	7.3 (3.5)	541 (69)	6.9 (3.3)	0.11 (0.00, 0.23)
Pupil-level (continuous)	n (missing)	Mean (SD)	n (missing)	Mean (SD)	Effect Size
CELF-Preschool 2 UK (raw scores)					
Sentence Structure[b]	593 (2)	10.0 (4.9)	608 (2)	9.6 (4.7)	0.09 (-0.03, 0.20)

Word Structure	587 (8)	10.3 (5.6)	608 (2)	9.6 (5.3)	0.14 (0.02, 0.25)
Expressive Vocabulary	586 (9)	13.8 (7.3)	606 (4)	12.8 (7.0)	0.15 (0.03, 0.26)
Concepts and Following Directions	550 (45)	8.0 (4.9)	542 (68)	7.3 (4.5)	0.16 (0.04, 0.27)
Child Self-regulation and Behaviour Questionnaire					
Sociability	588 (7)	3.6 (0.9)	572 (38)	3.6 (0.9)	0.00 (-0.11, 0.12)
Externalising	588 (7)	1.7 (0.8)	572 (38)	1.7 (0.8)	0.02 (-0.09, 0.14)
Internalising	588 (7)	1.6 (0.6)	572 (38)	1.6 (0.7)	-0.05 (-0.17, 0.06)
Prosocial behaviour	588 (7)	3.6 (0.8)	572 (38)	3.5 (0.9)	0.11 (-0.01, 0.22)
Behavioural self-regulation	588 (7)	3.7 (0.9)	572 (38)	3.7 (0.9)	0.00 (-0.11, 0.12)
Cognitive self-regulation	588 (7)	3.3 (0.8)	572 (38)	3.3 (0.9)	0.08 (-0.04, 0.19)
Emotional self-regulation	588 (7)	3.9 (0.7)	572 (38)	3.9 (0.8)	-0.07 (-0.19, 0.04)
Home Observation Measure of the Environment (n = 51)					
Total	26 (0)	34.5 (6.7)	25 (0)	33.3 (6.7)	0.17 (-0.38, 0.71)
Learning materials	26 (0)	7.8 (1.7)	25 (0)	6.9 (2.1)	0.48 (-0.07, 1.03)
Language stimulation	26 (0)	5.5 (1.6)	25 (0)	5.6 (1.3)	-0.04 (-0.58, 0.50)
Responsivity	26 (0)	5.5 (2.3)	25 (0)	6.0 (2.3)	-0.23 (-0.77, 0.32)
Academic stimulation	26 (0)	3.9 (1.0)	25 (0)	3.9 (1.0)	-0.04 (-0.58, 0.51)
Modelling	26 (0)	4.7 (1.3)	25 (0)	4.4 (1.2)	0.30 (-0.25, 0.84)
Variety of activities and parental interaction	26 (0)	7.0 (1.9)	25 (0)	6.6 (1.6)	0.27 (-0.28, 0.81)

[a] Percentages out of valid (non-missing) cases for categorical variables: EYPP—Early Years Pupil Premium; EAL—English as an Additional Language; SEND—Special Education Needs and Disability.
[b] Two fewer pupils have a Core Language Standard Score (n = 1203) than had any available CELF data (n = 1205) since one was too young for a scaled score to be computed for them, and one had missing data for all component subtests (data only valid for the Concepts subtest).

Figure 2: Distribution of CELF Core Language Standard Score scores (pre- and post-test)



Outcomes and analysis

Primary intention-to-treat (ITT) analysis

A total of 1149 (intervention group, n = 573; control group, n = 576) pupils completed the post-test CELF assessment. At post-test, the mean (SD) in the intervention group was 27.3 (7.8) and in the control group was 26.3 (8.2) (Table 12; Figure 2). Pre- *and* post-test Core Language Standard Score data was available for 1128 pupils across 101 nurseries (intervention group, n = 562, control group, n = 566) and this was included in the primary analysis model. The characteristics of these pupils are presented in Table 13 (the HOME outcome is not included in this table since it was only conducted in a subset of pupils). The age, proportion of males, and the distribution of ethnicities of pupils in the analysis in the intervention and control groups were similar; in all, pupils were aged about 45 months, around half were male, and two-thirds were white. A higher proportion of control pupils in the analysis were eligible for EYPP, had ELA, and were SEND compared to the intervention group.

The number of children included in the primary analysis model from each of the 101 nurseries ranged from 4 to 19 (median 11) in the intervention group, and 6 to 25 (median 11) in the control group. The mean (SD) pre-test Core Language Standard Score for this subset of participants was 25.4 (8.8) in the intervention group, and 24.6 (8.3) in the control group (Hedges' g effect size between the groups 0.10, 95% CI -0.02 to 0.226). The mean (SD) post-test Core Language Standard Score for this subset of participants was 27.3 (7.8) in the intervention group, and 26.3 (8.2) in the control group.

The unadjusted mean difference is 1.02 (95% CI 0.10 to 1.94) in favour of the intervention group (Table 14). The adjusted mean difference is 0.30 (95% CI -0.81 to 1.40, p = 0.60) in favour of the intervention group. Therefore, there was no statistically significant evidence of a difference in Core Language Standard Score between the intervention and control group. The estimated effect size is very small at 0.04 (95% CI -0.10 to 0.18), and does not equates to any additional months' progress (Table 12).

The total variance used to calculate the effect size, obtained from a model which did not adjust for pre-test score, was 63.8 (Table 14))—the sum of 51.8 (random variation between pupils, within-cluster variance) and 12.0 (heterogeneity between nurseries, between-cluster variance).

Table 12: Primary and secondary analysis for post-test CELF Preschool assessment outcome measure (scaled scores)

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)[a]	Mean (95% CI)	n (missing) [a]	Mean (95% CI)	n in model (intervention; control)	Hedges g (95% CI)	p-value
Core Language Standard Score	562 (33)	27.3 (26.7, 28.0)	568 (42)	26.3 (25.6, 27.0)	1128 (562; 566)	0.04 (-0.10, 0.18)	0.60
Sentence Structure	562 (33)	8.9 (8.7, 9.2)	567 (43)	8.8 (8.5, 9.1)	1124 (560; 564)	0.00 (-0.15, 0.16)	0.96
Word Structure	562 (33)	9.2 (8.9, 9.4)	566 (44)	8.7 (8.4, 9.0)	1120 (557; 563)	0.06 (-0.12, 0.24)	0.52
Expressive Vocabulary	561 (34)	9.2 (9.0, 9.5)	563 (47)	8.9 (8.7, 9.2)	1112 (554; 558)	0.01 (-0.11, 0.12)	0.92
Concepts and Following Directions	553 (42)	8.0 (7.8, 8.3)	558 (52)	7.5 (7.2, 7.8)	1017 (518; 499)	0.05 (-0.13, 0.23)	0.58

[a] Denominator is number of pupils pre-tested for CELF (n = 1205: intervention group, n = 595; control group, n = 610).

Table 13: Baseline comparison as included in the primary and secondary analysis (nursery n = 101; pupil n = 1128)

Nursery-level (categorical)	Intervention group		Control group	
	n/N (missing)	Count (%)[a]	n/N (missing)	Count (%)[a]
Local Authority Bedford Camden Coventry Doncaster Durham Islington Knowsley Luton Oldham	51/51 (0)	4 (7.8) 2 (3.9) 1 (2.0) 12 (23.5) 6 (11.8) 4 (7.8) 10 (19.6) 0 (0.0) 12 (23.5)	50/50 (0)	1 (2.0) 1 (2.0) 1 (2.0) 8 (16.0) 6 (12.0) 6 (12.0) 9 (18.0) 3 (6.0) 15 (30.0)
Nursery-level (continuous)	n (missing)	Mean (SD)	n (missing)	Mean (SD)
Percentage of pupils eligible for EYPP	47 (4)	47.2 (16.7)	48 (2)	47.0 (17.0)
Number of pupils starting school in 2018	46 (5)	34.2 (16.4)	47 (3)	29.5 (17.8)

Pupil-level (categorical)	n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Gender, male	556/562 (6)	283 (50.9)	557/566 (9)	274 (48.4)	
Ethnicity White Black/Caribbean Asian Middle Eastern Mixed Other Unknown	562/562 (0)	379 (67.4) 23 (4.1) 81 (14.4) 9 (1.6) 37 (6.6) 3 (0.5) 30 (5.3)	566/566 (0)	372 (65.7) 29 (5.1) 103 (18.2) 6 (1.1) 34 (6.0) 4 (0.7) 18 (3.2)	
Eligible for EYPP	484/562 (78)	67 (13.8)	457/566 (109)	118 (25.8)	
EAL	530/562 (32)	129 (24.3)	548/566 (18)	167 (30.5)	
SEND	528/562 (34)	19 (3.6)	536/566 (30)	34 (6.3)	
Pupil-level (continuous)	n (missing)	Mean (SD)	n (missing)	Mean (SD)	Effect Size
Age, months	562 (0)	45.9 (4.1)	566 (0)	45.5 (3.6)	N/A
CELF-Preschool 2 UK (scaled scores)					
Core Language Standard Score	562 (0)	25.4 (8.8)	566 (0)	24.6 (8.3)	0.10 (-0.02, 0.22)
Sentence Structure	560 (2)	8.0 (3.3)	565 (1)	7.8 (3.0)	0.06 (-0.05, 0.18)
Word Structure	557 (5)	8.8 (3.6)	565 (1)	8.4 (3.3)	0.10 (-0.01, 0.22)
Expressive Vocabulary	555 (7)	8.9 (3.4)	563 (3)	8.4 (3.3)	0.13 (0.01, 0.25)
Concepts and Following Directions	522 (40)	7.5 (3.5)	505 (61)	7.0 (3.3)	0.15 (0.03, 0.28)
CELF-Preschool 2 UK (raw scores)					
Sentence Structure	560 (2)	10.1 (4.9)	565 (1)	9.8 (4.7)	0.08 (-0.04, 0.20)
Word Structure	557 (5)	10.5 (5.6)	565 (1)	9.7 (5.2)	0.15 (0.03, 0.27)
Expressive Vocabulary	555 (7)	14.0 (7.2)	563 (3)	12.8 (7.0)	0.16 (0.05, 0.28)
Concepts and Following Directions	522 (40)	8.1 (4.9)	505 (61)	7.3 (4.5)	0.17 (0.05, 0.29)
Child Self-regulation and Behaviour Questionnaire					
Sociability	555 (7)	3.6 (0.9)	529 (37)	3.6 (0.9)	-0.02 (-0.14, 0.10)
Externalising	555 (7)	1.7 (0.8)	529 (37)	1.7 (0.8)	0.02 (-0.010, 0.14)
Internalising	555 (7)	1.6 (0.6)	529 (37)	1.6 (0.7)	-0.04 (-0.16, 0.08)
Prosocial behaviour	555 (7)	3.6 (0.8)	529 (37)	3.5 (0.9)	0.09 (-0.03, 0.21)
Behavioural self-regulation	555 (7)	3.7 (0.9)	529 (37)	3.7 (1.0)	0.00 (-0.12, 0.12)
Cognitive self-regulation	555 (7)	3.4 (0.8)	529 (37)	3.3 (0.9)	0.04 (-0.08, 0.16)
Emotional self-regulation	555 (7)	3.9 (0.7)	529 (37)	3.9 (0.8)	-0.07 (-0.19, 0.05)

[a] Percentages out of valid (non-missing) cases for categorical variables: EYPP—Early Years Pupil Premium; EAL—English as an Additional Language; SEND—Special Education Needs and Disability.

Table 14: Effect size estimation (CELF Preschool assessment)

Outcome	Unadjusted differences in means (95% CI)	Adjusted differences in means (95% CI)	Intervention group		Control group		Pooled variance	ICC
			n (missing)[a]	Variance of outcome	n (missing) [a]	Variance of outcome		
Core Language Standard Score	1.02 (0.10, 1.94)	0.30 (-0.81, 1.40)	562 (33)	60.9	568 (42)	67.0	63.8	0.18 (0.14, 0.24)
Sentence Structure	0.12 (-0.23, 0.47)	0.01 (-0.45, 0.47)	562 (33)	8.6	567 (43)	9.7	9.1	0.12 (0.08, 0.18)
Word Structure	0.46 (0.08, 0.84)	0.19 (-0.39, 0.78)	562 (33)	10.4	566 (44)	11.4	10.9	0.22 (0.16, 0.29)
Expressive Vocabulary	0.33 (-0.01, 0.67)	0.02 (-0.31, 0.34)	561 (34)	8.7	563 (47)	8.7	8.7	0.08 (0.04, 0.14)
Concepts and Following Directions	0.53 (0.17, 0.89)	0.16 (-0.40, 0.71)	553 (42)	9.1	558 (52)	10.2	9.6	0.21 (0.16, 0.27)

[a] Denominator is number of pupils pre-tested for CELF (n = 1205: intervention group, n = 595; control group, n = 610).

The observed ICC at the nursery level obtained from the primary analysis model was 0.18 (95% CI 0.14 to 0.24). The proportion of variance explained by covariates at the individual level (Level 1, pupil level) was 0.52 and at the cluster level was 0.54 (Level 2, setting level). Therefore, the estimated minimum detectable effect size for the primary analysis was 0.20 (Table **10Error! Reference source not found.**).

The model fit was assessed via visual inspection of graphical plots, and no evidence of vitiation of the underlying assumptions was observed (Appendix D); hence, sensitivity analyses considering transformations of the outcome and/or covariate data to improve model fit were not required.

The correlation between pre- and post-intervention scores for the scaled and raw CELF scores are presented in Table **15**. The correlation for the primary outcome of the scaled total score (the Core Language Standard Score) is quite strong at 0.72. Correlations for scaled individual subscale scores range from 0.50 to 0.73.

Table 15: Correlation between pre- and post-intervention administrations of CELF Preschool assessment outcome

Scaled or raw	CELF outcome	Correlation between pre- and post-intervention administrations of outcome
Scaled scores	Core Language Standard Score	0.72
	Sentence Structure	0.50
	Word Structure	0.55
	Expressive Vocabulary	0.73
	Concepts and Following Directions	0.51
Raw scores[a]	Sentence Structure	0.53
	Word Structure	0.57
	Expressive Vocabulary	0.75
	Concepts and Following Directions	0.54

[a] No total score supported from 'raw scores' for this outcome.

Missing data

In the adjusted mixed-effect logistic regression with presence or absence of post-test Core Language Standard Score data as the outcome variable, two of the included covariates were observed to be statistically significant predictors, assessed at the 5% level. These were: gender (males less likely to have missing data); and number of pupils at the nursery due to attend school in 2018 (pupils attending larger schools less likely to have missing data). These variables were included in the multiple imputation. To investigate the impact of missing data, the primary analysis was repeated using predicted values obtained via multiple imputation by chained equations. The adjusted mean difference in Core Language Standard Score between the two groups following multiple imputation was 0.31 (95% CI 0.80 to 1.42); this

difference is not statistically significant at the 5% significance level (p = 0.58). The estimated effect size for the intervention effect is very small at 0.04 of a standard deviation (95% CI 0.10 to 0.18), which does not relate to any months' additional progress, and is virtually identical to results of the primary analysis. The total variance used to calculate the effect size was 65.1—the sum of 52.4 (random variation between pupils, within-cluster variance) and 12.7 (heterogeneity between nurseries, between-cluster variance).

Compliance

The intervention delivery team reported compliance data back to the evaluation team. They reported that there were no web-app outages or technical failures with the sending of text messages to parents during the trial period. The failure rate relating to parents not receiving any text messages (that is, all the attempts were marked as undelivered) was approximately 3.6%, and the rate of parental drop-out was 3.7%.

Of the 51 nurseries randomised to receive the EasyPeasy intervention, the Pod Leader from 35 (68.6%) attended the training via live webinar with the ability to ask questions in webinar session. The remaining 16 (31.4%) all watched a recording of the training, and received a subsequent follow-up telephone call and/or emails to answer any questions and check the nursery was set up to deliver the intervention.

The CACE estimate of the effect of the Pod Leader attending the training in person on the pupils' Core Language Standard Score was a predicted adjusted increase of 0.61 points (95% CI -0.99 to 2.21, p = 0.45; estimated effect size 0.08, 95% CI -0.12 to 0.28). This is larger than the treatment effect estimated from the primary intention-to-treat model indicating that children performed better when the nursery complied with the intervention, but is still not statistically significant.

One nursery in the control group closed down shortly after randomisation, and its ten participating pupils moved to one of two intervention nurseries. Seven of the pupils moved to a nursery where the Pod Leader did not attend the training in person; however, the remaining three did. In the initial CACE analysis, these ten pupils were analysed by association with the nursery they were in at enrolment (the control nursery, which subsequently closed down). A second CACE analysis was conducted which recoded the nursery for these ten pupils to the nursery they moved to after the initial nursery closed (that is, in this analysis, three of the pupils were assumed to have complied with the intervention). Results were very similar to the first CACE analysis (CACE estimate 0.61, 95% CI -0.96 to 2.18, p = 0.44; estimated effect size 0.08, 95% CI -0.12 to 0.27).

Post-hoc sensitivity analysis—assessor quality

Pre-test

Twenty-one independent assessors conducted the CELF pre-intervention tests with the pupils across the 102 nurseries (1205 pre-tests; average of 57 per assessor, SD 36.1, median 48, range 12 to 124). The quality of each assessor was graded (as measured by listening to 10% of audio recordings from each assessor); 12 (57%) were graded 'very good', a third (n = 7) as 'good', and two (10%) as 'acceptable' based on agreed criteria for how each assessor adhered to the assessment scoring and administration protocol. Grading scale and criteria were determined by a trained psychologist and carried out by a researcher in the evaluation team who was first quality-assurance checked by the psychologist prior to grading all the assessors. Of the 1128 pupils in the primary analysis, 584 (52%) were assessed at pre-test by the 12 'very good' assessors, 478 (42%) by the seven 'good' assessors, and 66 (6%) by the two 'acceptable' assessors. A post-hoc sensitivity analysis was conducted in which the primary ITT analysis was repeated with an additional covariate for the grading of the assessor at pre-test and including a random effect for assessor. The adjusted mean difference is 0.33 (95% CI -0.49 to 1.15, p = 0.43) in favour of the intervention group. Therefore, there was no statistically significant evidence of a difference in Core Language Standard Score between the intervention and control group in this analysis. The estimated effect size is positive but small at 0.04 (95% CI -0.06 to 0.14), and does not equate to any additional months' progress.

Post-test

Twenty independent assessors (three of whom were involved in pre-testing) conducted the CELF post-intervention tests with the pupils across the 101 nurseries (1149 post-tests; average of 57 per assessor, SD 28.8, median 50, range 13 to 120). Half the assessors were graded 'very good' (n = 10), six as 'good' (30%), one as 'acceptable' (5%), and three as 'unacceptable' (15%). The definition of 'unacceptable' for quality assurance, in this case, was if assessors deviated from the assessment scoring and administration protocol, for example, by asking children a question more than once in different ways, marking a question as incorrect and then explaining the correct answer to the child (which could affect the child's confidence to answer following questions), or inconsistency in marking. Of the 1128 pupils in the primary analysis that were post-tested, 619 (55%) were assessed by the 10 'very good' assessors, 330 (29%) by the six 'good' assessors, 12 (1%) by the one 'acceptable' assessor, and 167 (15%) by the three 'not acceptable' assessors.

A post-hoc sensitivity analysis was conducted in which the primary ITT analysis was repeated with an additional covariate for the grading of the assessor at post-test and including a random effect for assessor. The adjusted mean difference is -0.05 (95% CI -0.83 to 0.73, p = 0.90) in favour of the control group. Therefore, there was no statistically significant evidence of a difference in Core Language Standard Score between the intervention and control group in this analysis. The estimated effect size is negative and very small at -0.01 (95% CI -0.10 to 0.09), and does not equate to any additional months' progress.

An additional post-hoc sensitivity analysis was conducted excluding pupils from the primary analysis whose post-tests were conducted by an assessor who was graded as 'unacceptable'. This removed 167 pupils across 18 nurseries. The adjusted mean difference is -0.12 (95% CI -0.99 to 0.75, p = 0.79) in favour of the control group. Therefore, there was no statistically significant evidence of a difference in Core Language Standard Score between the intervention and control group in this analysis. The estimated effect size and 95% CI is very similar to the analysis described in the previous paragraph (0.01, 95% CI -0.12 to 0.09), and does not equate to any additional months' progress.

Subgroup analyses

Summary statistics for the Core Language Standard Score are presented in Table **16** by EYPP status, EAL status, and gender. These summaries indicate that, in general, EYPP and EAL children tended to perform worse than their peers and female pupils tended to perform slightly better than males. In adjusted regression analyses that included interaction effects, the effect of the intervention on post-test Core Language Standard Score was not seen to be substantially altered by having EYPP status (interaction effect p = 0.58), having EAL (interaction effect p = 0.99) or by gender (interaction effect p = 0.12).

Table 16: Summary of post-intervention CELF Core Language Standard Score by randomised group and EYPP status, EAL status, and gender

	Raw means			
	Intervention group		Control group	
Outcome	n	Mean (SD)	n	Mean (SD)
EYPP: Yes No	67 428	26.9 (7.9) 27.3 (7.8)	118 348	24.3 (7.7) 27.2 (8.0)
EAL: Yes No	132 409	21.7 (8.5) 29.0 (6.8)	171 387	21.3 (7.3) 28.4 (7.6)
Gender: Male Female	286 281	26.7 (7.5) 27.9 (8.0)	276 291	25.2 (8.1) 27.2 (8.2)

Secondary outcome analysis

The raw CELF Preschool assessment scores are summarised in Appendix E, but were not analysed.

The secondary outcomes of the individual CELF Preschool scaled subscale scores and the seven CSBQ subscales are summarised; these were analysed exactly as described for the primary outcome, adjusting for the associated pre-test score instead of the Core Language Standard Score pre-test score. Findings from these analyses are presented in Table 12, Table 14, Table 17, and Table 18. No statistically significant differences were observed between the two groups.

Between 1017 and 1124 pupils are included in the analysis models for the CELF scaled subscale scores. The adjusted differences in means for all subscale scores all marginally favour the intervention group, but the effect sizes between the two groups are very small (≤0.06), equating to a month or less of additional progress.

Pre-test CSBQ data (for any subscale) was available for 1433 (96.3%) pupils—intervention group, n = 746 (98.6%); control group, n = 687 (94.0%). Post-test CSBQ data (for any subscale) was available for 979 (65.8%) pupils—intervention group, n = 529 (69.9%); control group, n = 450 (61.6%). CSBQ responses were completed by the class teacher and nurseries were contacted multiple times with reminders to complete and return forms at both pre- and post-testing. However, a number of nurseries did not return the forms at post-testing. For any one subscale, the number of pupils included in the analysis model was either 953 (intervention group, n = 528; control group, n = 425) or 955 (intervention group, n = 529; control group, n = 426); approximately 64% of the randomised population. A higher score indicates a better outcome. There were no statistically significant differences between the two groups in any of the CSBQ subscales.

There were some moderate to large imbalances in pre-test HOME scores between the intervention and control groups. A similar approach was taken for the analysis of the HOME total and subscale scores, except the analysis models were not adjusted for the associated pre-test HOME score since this was measured post-randomisation. Results are given in Table 19 and Table 20. All 48 households for which HOME data was collected at post-test were included in all analysis models.

A higher score indicates a better outcome. At post-test, the unadjusted mean total score was higher in the intervention group (38.0) than in the control group (33.9). In the analysis adjusting for the minimisation factor of number of eligible pupils per school only, the adjusted difference in means favoured the intervention group (5.22, 95% CI 0.72 to 9.71), and was statistically significant (effect size 0.76, 95% CI 0.11 to 1.42, p = 0.02). Scores across all subtests also favoured the intervention group, and statistically significant differences were also observed in the Modelling and Variety of Activities and Parental Interaction subtests (p ≤ 0.05; Table 19). Across all subtest scores, the effect sizes ranged from 0.42 to 0.72.

Sensitivity analyses additionally adjusting for pre-test HOME score were undertaken to assess the impact on the results (Table 21). Adjusted differences in mean favoured the intervention group. The statistically significant difference in total score and Variety of Activities and Parental Interaction subtest remained. The effect size for Learning Materials reduced substantially when adjustment was made for the pre-test score from 0.42 (Table 19) to 0.06 (Table 21). At pre-test there was a large imbalance in this subscale score between the intervention and control groups of nearly half a standard deviation (0.48, Table 11), which seems to have been driving the large observed difference in post-test since this difference disappears once the pre-test score is controlled for.

The reliability of the HOME scale, made up of the six subscale scores, was investigated using Cronbach's alpha as a measure of internal consistency. The scale reliability coefficient at pre-test was 0.75, and at post-test was 0.83, suggesting that the items have relatively high internal consistency.

Table 17: Secondary analysis, post-test CSBQ

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)[a]	Mean (95% CI)	n (missing) [a]	Mean (95% CI)	n in model (intervention; control)	Hedges' g (95% CI)	p-value
Sociability	529 (217)	3.9 (3.9, 4.0)	426 (261)	3.9 (3.8, 4.0)	955 (529; 426)	0.04 (-0.13, 0.21)	0.65
Externalising	529 (217)	1.6 (1.5, 1.6)	426 (261)	1.6 (1.6, 1.7)	955 (529; 426)	-0.08 (-0.28, 0.12)	0.42
Internalising	528 (218)	1.4 (1.3, 1.4)	425 (262)	1.5 (1.4, 1.6)	953 (528; 425)	-0.08 (-0.29, 0.13)	0.44
Prosocial behaviour	528 (218)	3.9 (3.9, 4.0)	425 (262)	3.9 (3.8, 4.0)	953 (528; 425)	-0.02 (-0.23, 0.20)	0.86
Behavioural self-regulation	529 (217)	4.0 (3.9, 4.0)	426 (261)	4.0 (3.9, 4.0)	955 (529; 426)	-0.02 (-0.22, 0.19)	0.85
Cognitive self-regulation	529 (217)	3.8 (3.7, 3.9)	426 (261)	3.6 (3.6, 3.7)	955 (529; 426)	0.14 (-0.06, 0.33)	0.17
Emotional self-regulation	529 (217)	4.0 (3.9, 4.1)	426 (261)	4.0 (3.9, 4.1)	955 (529; 426)	0.06 (-0.18, 0.29)	0.63

[a] Denominator is number of pupils pre-tested for CSBQ (n = 1433; intervention group, n = 746; control group, n = 687).

Table 18: Effect size estimation (CSBQ)

Outcome	Unadjusted differences in means	Adjusted differences in means	Intervention group		Control group		Pooled variance	
			n (missing)[a]	Variance of outcome	n (missing)[a]	Variance of outcome		ICC
Sociability	0.05 (-0.05, 0.15)	0.03 (-0.10, 0.17)	529 (217)	0.60	426 (261)	0.70	0.64	0.16 (0.09, 0.26)
Externalising	-0.05 (-0.15, 0.05)	-0.06 (-0.21, 0.09)	529 (217)	0.46	426 (261)	0.72	0.57	0.20 (0.11, 0.34)
Internalising	-0.10 (-0.18, -0.02)	-0.05 (-0.18, 0.08)	528 (218)	0.32	425 (262)	0.46	0.38	0.20 (0.13, 0.31)
Prosocial behaviour	0.03 (-0.08, 0.14)	-0.02 (-0.20, 0.17)	528 (218)	0.77	425 (262)	0.80	0.75	0.32 (0.16, 0.53)
Behavioural self-regulation	0.01 (-0.10, 0.12)	-0.02 (-0.19, 0.16)	529 (217)	0.70	426 (261)	0.78	0.72	0.29 (0.16, 0.48)
Cognitive self-regulation	0.16 (0.05, 0.27)	0.12 (-0.05, 0.28)	529 (217)	0.62	426 (261)	0.86	0.73	0.21 (0.14, 0.30)
Emotional self-regulation	0.02 (-0.08, 0.12)	0.05 (-0.14, 0.23)	529 (217)	0.57	426 (261)	0.74	0.63	0.31 (0.18, 0.49)

[a] Denominator is number of pupils pre-tested for CSBQ (n = 1433; intervention group, n = 746; control group, n = 687).

Table 19: Secondary analysis, post-test HOME

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model (intervention; control)	Hedges g (95% CI)	p-value
Total	23 (3)	38.0 (36.0, 39.9)	25 (0)	33.9 (31.0, 36.8)	48 (23; 25)	0.76 (0.11, 1.42)	0.02
Learning materials	23 (3)	7.2 (6.7, 7.7)	25 (0)	6.4 (5.4, 7.4)	48 (23; 25)	0.42 (-0.10, 0.95)	0.11
Language stimulation	23 (3)	6.5 (6.1, 6.8)	25 (0)	6.2 (5.7, 6.6)	48 (23; 25)	0.50 (-0.15, 1.14)	0.13
Responsivity	23 (3)	6.9 (6.2, 7.6)	25 (0)	6.2 (5.5, 6.8)	48 (23; 25)	0.59 (-0.09, 1.26)	0.09
Academic stimulation	23 (3)	4.5 (4.0, 4.9)	25 (0)	4.0 (3.5, 4.5)	48 (23; 25)	0.45 (-0.23, 1.12)	0.19
Modelling	23 (3)	4.9 (4.4, 5.4)	25 (0)	4.1 (3.7, 4.6)	48 (23; 25)	0.66 (-0.01, 1.33)	0.05
Variety of activities and parental interaction	23 (3)	8.0 (7.5, 8.5)	25 (0)	7.1 (6.3, 7.8)	48 (23; 25)	0.72 (0.07, 1.38)	0.03

Table 20: Effect size estimation (HOME)

Outcome	Unadjusted differences in means	Adjusted differences in means	Intervention group		Control group		Pooled variance
			n (missing)	Variance of outcome	n (missing)	Variance of outcome	
Total	4.04 (0.64, 7.44)	5.22 (0.72, 9.71)	23 (3)	20.9	25 (0)	50.1	46.6
Learning materials	0.77 (-0.28, 1.82)	0.79 (-0.18, 1.76)	23 (3)	1.4	25 (0)	5.3	3.5
Language stimulation	0.32 (-0.22, 0.86)	0.48 (-0.14, 1.09)	23 (3)	0.6	25 (0)	1.1	0.9
Responsivity	0.75 (-0.14, 1.64)	0.96 (-0.14, 2.05)	23 (3)	2.4	25 (0)	2.6	2.7
Academic stimulation	0.48 (-0.18, 1.14)	0.55 (-0.28, 1.37)	23 (3)	1.1	25 (0)	1.6	1.5
Modelling	0.79 (0.13, 1.45)	0.82 (-0.01, 1.65)	23 (3)	1.4	25 (0)	1.3	1.5
Variety of activities and parental interaction	0.92 (0.06, 1.78)	1.20 (0.12, 2.28)	23 (3)	1.3	25 (0)	3.2	2.7

Table 21: Effect size estimation (HOME—adjusted sensitivity analyses)

Outcome	Adjusted differences in means	Pooled variance	n in model (intervention; control)	Hedges' g (95% CI)	p-value
Total	4.22 (0.84, 7.60)	46.6	48 (23; 25)	0.62 (0.12, 1.11)	0.01
Learning materials	0.11 (-0.51, 0.73)	3.5	48 (23; 25)	0.06 (-0.27, 0.39)	0.73
Language stimulation	0.48 (-0.12, 1.08)	0.9	48 (23; 25)	0.50 (-0.13, 1.13)	0.12
Responsivity	1.07 (0.03, 2.12)	2.7	48 (23; 25)	0.66 (0.02, 1.30)	0.04
Academic stimulation	0.48 (-0.11, 1.08)	1.5	48 (23; 25)	0.39 (-0.09, 0.88)	0.11
Modelling	0.66 (-0.12, 1.43)	1.5	48 (23; 25)	0.53 (-0.10, 1.15)	0.10
Variety of activities and parental interaction	1.03 (0.11, 1.95)	2.7	48 (23; 25)	0.62 (0.07, 1.18)	0.03

Cost

The cost per pupil of delivering EasyPeasy in schools over a three-year period is shown in Table 22 and Table 23 below. The method for calculating the cost per pupil is stated above in the method section of the report and is in line with EEF guidelines.

Direct, marginal costs

The cost per pupil for the 20-week EasyPeasy intervention is estimated to be £24.48 (averaged over a three-year period).

To implement the EasyPeasy intervention, the only direct cost to a school is the annual licence fee. A single webinar training session for nursery leads is part of the package which schools purchase through their annual licence fee. The total cost for a school will depend upon the number of pupils in the school. The cost calculation is based on the average number of three- to four-year-old pupils (31) that schools reported would be moving on into Reception in the academic year following the trial. Calculations are carried out using the costs for the 2018/2019 academic year.

Table 22: Cost of delivering EasyPeasy

Item	Type of cost	Cost (exc. VAT)	Total cost over 3 years (exc. VAT)	Total cost per pupil per year over 3 years
Annual licence fee	Cost per school (dependent upon school size)	£759	£2277	£73.45
Total			£2277	£73.45

Table 23: Cumulative costs per pupil per year of EasyPeasy

	Year 1	Year 2	Year 3
EasyPeasy	£24.48	£48.97	£73.45

The cost per pupil per year is £24.48 averaged over three years (in line with the EEF guidance)[9] to allow for comparison between EEF projects.

[9] **https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_guidance_to_evaluators_on_cost_evaluation.pdf**

Additional resources required to deliver the intervention

The EasyPeasy programme is expected to run on-site at schools. Therefore, there are no associated venue costs or travel and subsistence costs. CPD is via an online webinar, which is part of the annual licence fee.

A few Pod Leaders raised the issue of additional monetary costs arising from implementing the programme. These included printing costs of up to £20, and costs of up to £50 for additional materials used in Stay and Play sessions, however, this was not consistent across all settings and was not a requirement for implementing EasyPeasy.

Access to a computer/tablet with internet connection was a prerequisite for schools to participate as this was required to access training and for use of the online dashboard to facilitate and encourage parents' use of EasyPeasy.

Staff time

The Pod Leader practitioner survey results gave details of time costs at different points during the intervention.

Initial set-up period:

- staff meetings and recruitment sessions with parents (normally one to two hours in total);
- webinar training (up to one hour); and
- entering parents' details in the online dashboard: time needed depended on the size of the Pod; according to the Pod Leaders who responded this question—
    - it was less than half an hour for 39.2% of the Pod Leaders;
    - half an hour to an hour for 53.6% of the Pod Leaders; and
    - between one and two hours for 7% of the Pod Leaders (the developer indicated that schools can use CSV files to upload parent names and phone numbers, potentially bringing this time down).

Implementation period (sending reminders to parents, reading emails, checking new games, etc.):

- 72% of Pod Leaders (as reported in the practitioner survey) felt that they needed approximately one hour per week to run the programme.

This finding was in line with the expectations of the developer of EasyPeasy. The developer suggested three hours of staff time were required during the set-up period, one hour for promoting EasyPeasy to parents, one and a half hours for training, and half an hour for entering parent details. She expected that Pod Leaders would need around two hours per week during the implementation period.

In summary, the input from school staff is estimated to be at least one hour per week when running the programme and approximately three hours for the initial set-up period. In total, therefore, an intervention of 20 weeks in a year requires approximately 23 hours of staff time.

There is no finding to suggest that schools might need to engage additional staff in order to implement the programme. However, current staff were conscious of a lack of time for combining EasyPeasy with their normal activity in the setting.

Cost of intervention for schools who took part in the intervention

The cost of the EasyPeasy annual license fee was waived for schools participating in the trial with the EEF covering 61% of the cost of the intervention and participating Local Authorities paying the remaining 39% for their settings to be part of the trial. Five LAs paid £15,000 each with two sharing the cost equally between them (paying £7,500 each). Settings in Doncaster paid their LA £250 to participate in the programme. These costs are not included in calculations to avoid double counting of costs as the £250 was a contribution from schools towards the cost for the LA participating in the trial.

Local Authorities were permitted to offer EasyPeasy to all schools within the Local Authority for the £15,000 cost. There was no obligation on schools to participate in the trial and the number of participating schools ranged from two to 26 across the LAs. Schools were then permitted to offer the use of EasyPeasy to all parents within a setting. The developers

were unable to specify how many settings and parents within each LA signed up to use EasyPeasy and, as such, the calculations in Table **24** below are based on the total number of pupils in each setting participating in the trial eligible to move on to Reception in 2018/2019 (gathered from the pre-intervention usual practice survey), with the average number used for the eight schools that did not provide this information.

Table 24: Calculation of the actual cost per pupil of participating in the trial

Cost for participation in the trial	Total no. of pupils	Total cost per pupil
Intervention cost paid by the EEF = £143,583 Intervention cost paid by Local Authorities = £90,000 Total intervention cost = £233,534	3,239	£233,534 ÷ 3239 (children) = £72.10

Implementation and process evaluation

Here we discuss the key themes emerging from the process evaluation that relate to the perceptions of Pod Leaders and parents and the resulting conclusions that can be drawn regarding barriers and facilitators to success. These have been categorised as implementation and fidelity, quality, perceived outcomes, and control group activity. Table **6** above provided an overview of the full set of implementation and process evaluation research questions. Under each category below is a table summarising the research questions relevant to that category. Evidence is drawn from all data collection sources including interviews, focus groups, and surveys.

Implementation and fidelity

Research questions	Dimensions
To what extent are the settings and parents engaging with and delivering the intervention?	Fidelity Dosage Reach
How is the intervention being disseminated to parents through settings and is the method effective? a) Are nominated staff engaging with EasyPeasy and using it as a mechanism to foster communication and engagement between parents? b) Are parents engaging with EasyPeasy and implementing the suggested activities within the home?	Fidelity Adaptation Dosage
Can all parents access and engage with the EasyPeasy intervention—for example, from all socioeconomic backgrounds, in EAL families, in families with SEND children?	Fidelity Dosage Reach Adaptation

Here we discuss the key themes emerging from the process evaluation that relate to the implementation and fidelity of the intervention. These have been categorised as training and support, staff roles, and engaging parents.

Training and support

According to data from the developers, 69% (n = 35/51) of Pod Leaders attended the one-hour live webinar training. During the training, Pod Leaders had the opportunity to ask questions; however, in the webinar observed by the evaluator, this was not an option used by most attendees. For those unable to attend the webinar training, EasyPeasy provided a video recording of the training, following this up with a phone call and emails. A follow-up phone call was not part of the standard training process for Pod Leaders who attended the webinar, however, all Pod Leaders had the option of contacting EasyPeasy at any point during the intervention to seek support. Overall feedback on the training indicated that 59% (n = 23/39) of Pod Leaders felt it to be 'very useful', 38% (n = 15/39) felt it was 'quite useful', while one Pod Leader found it to be 'not useful at all'. The developer stated that it had received positive feedback from Pod Leaders on the webinar-based training session and planned to continue with this practice.

Following the webinar training session, around half of Pod Leaders (n = 25/43, 58%) felt 'completely confident' to deliver EasyPeasy, with the remaining sample stating they felt 'quite confident' (n = 17/43, 40%), although one staff member reporting being 'not very confident'. The findings from the Pod Leader practitioner survey indicated that 76% (n = 34/45) of Pod Leaders cascaded their training knowledge to other staff in their setting, including teaching assistants and EYFS team members. Following the training, 37% (n = 15/41) of Pod Leaders required further support from the EasyPeasy delivery team: topics included navigating aspects of the online Pod or password resets. Some Pod Leaders felt the training could have included further information and ideas on how to approach and engage parents.

Staff roles

Interview data indicated that Pod Leaders perceived the purpose of EasyPeasy differently. Some Pod Leaders discussed how they perceived EasyPeasy to be used by parents within the home, and others perceived it as a vehicle to help the school to engage parents and/or to link home and classroom activities.

The perceived purpose of EasyPeasy ultimately impacted on how Pod Leaders viewed their role in association with the programme and how it was rolled out within the setting. Some Pod Leaders adopted a 'facilitator' role whereby they assisted parents with using EasyPeasy and answering their questions, whereas others adopted a 'leader' role proactively combining EasyPeasy games with classroom activities.

Pod Leaders who adopted a proactive leader role integrated EasyPeasy into everyday nursery practice (without parents) where appropriate:

> *'Sometimes what I do … I know the text message will go out and I will check what game it is for that week and so, for instance, band practice … we have instruments outside and we can involve it. Sometimes … I have looked up some of the games and some seem more applicable to nursery and others to home and so as we progress through the week then that is when I might start to use some more of the games rather than the ones that have been sent out already.'*
>
> *Pod Leader, longitudinal case-study Pod Leader interview, interviewed at the end of the intervention.*

The role that Pod Leaders adopted often determined their methods to promote and engage the use of EasyPeasy among parents.

Engaging parents

Findings from the Pod Leader practitioner survey—completed by n = 45/51 (88.2%) post-intervention—indicated that there were two predominant approaches at the beginning of the intervention to engage parents. These included (1) communicating with parents at drop-off and pick-up times (n = 36/45; 80%) and (2) sending letters, emails, and/or texts to parents (n = 39/45, 87%). During case-study interviews, some Pod Leaders also discussed how they used open days and/or parents meetings as an opportunity to promote EasyPeasy to parents.

Throughout the course of the intervention, more than half (n = 23/45, 51%) of Pod Leaders acknowledged that they undertook additional approaches to keep parents engaged, or to re-engage those who were not active. These took the form of Stay and Play sessions, visits to homes, advice on how to differentiate the games, and the setting of games as homework.

The two Pod Leaders in the best practice case-study schools adopted a proactive 'leader' role and incorporated the EasyPeasy games into the nursery environment to encourage parental engagement. In one of these schools, this took the form of:

1. holding a weekly half-hour structured 'Stay and Play' session within the nursery, which enabled parents to understand the games, encouraged children to show the games to their parents, and meant that the games were played in the nursery first;
2. using the board on the entrance wall to display parents' comments about their use of EasyPeasy and about individual games, with the aim of attracting other parents' attention;
3. preparing the resources/materials required in the games on behalf of the parents, such as big rolls of paper and coloured paper, even though they could easily be found at home; and
4. rewarding children with special green balloons stickers that were exclusive to the EasyPeasy activities.

From the focus group which took place in this school, parents were very positive about their comments being displayed on the board and they appreciated reading other parents' comments. However, the Pod Leader at this setting felt that these approaches had not increased participation as much as they would have desired.

> *'We still weren't increasing the numbers that were accessing the app on a regular basis.'*

> *'For the amount that we had to put in, there was little gain.'*

The second best practice case-study school had been using an online messaging app for two years prior to the intervention for communicating with parents, and they considered this to be the most successful way they had tried for communicating with parents about EasyPeasy. The school messaging app had additional functions in comparison with

the comments section of the EasyPeasy Pod Leader dashboard. These functions included the ability to circulate reminders to parents, give in-app rewards to children (which could be viewed by parents), and message parents directly and know when this had been read.

Pod Leaders commented during interviews that the weekly update email from EasyPeasy did not assist significantly with monitoring ongoing engagement with parents. The EasyPeasy weekly update email provided information about the three most active and three least active parents within their Pod that week, number of playdates within the Pod that week, and a summary of the total number of playdates during the intervention. However, Pod Leaders felt that the lack of demographic data on individual parental engagement in the weekly report email made it difficult to target particular parents.[10] It was also commented upon by Pod Leaders that the use of parents' names in the weekly report rather than their children's names was problematic as parents' and children's surnames were not always the same. Although Pod Leaders commented that the reports contained too little detail, the emails were generally well read; the Pod Leader practitioner survey indicated that 80% (n = 36/45) Pod Leaders read the email each week, whilst the remaining 20% read it every other week or monthly. The developer reported, during the post-intervention interview, that the format of the weekly report was changed during the course of the intervention to make it more appealing for Pod Leaders to read. The developer also reports that each weekly email also provides parental engagement tips

EasyPeasy did not provide information to the evaluation team for the number of playdates per school throughout the intervention, which could have better gauged the level of parental engagement. Findings from the Pod Leader practitioner survey indicated that, in general, the numbers of parents accessing EasyPeasy reduced during the intervention period (Figure **3**). Even for schools with high parent engagement at the start of the intervention period (75% access rates at the beginning), the percentage of parents engaging with EasyPeasy fell to below 20% by the end of the 20-week intervention. For example, the Pod Leader at case-study school D reported that 65% of parents engaged in the first week (based on the number of playdates reported in the weekly update email the Pod Leader had received from EasyPeasy), but that this dropped to 20% straight away, and to even lower levels later. During the 20-week intervention, a total of 2465 parents registered to use EasyPeasy in participating Local Authorities (which will have included all parents of nursery children participating in the trial and nursery and reception pupils not participating in the trial); 3.6% (n = 91/2465) parents actively requested to be removed from lists receiving any further messages from EasyPeasy. EasyPeasy were unable to report whether these 91 parents were part of the trial or at what point in the 20-week programme they withdrew. Although none of the Pod Leaders in the case-study schools had parents asking to stop receiving messages, the numbers of play dates in the automated weekly reports to Pod Leaders were lower than expected, even within the two best practice settings. The Pod Leader at case-study school F, one of the best practice settings, reported poor engagement from parents to all nursery activities and this was also the case for this intervention; they reported that parents 'don't engage, no matter how much you encourage them, and so it has been really tricky'.

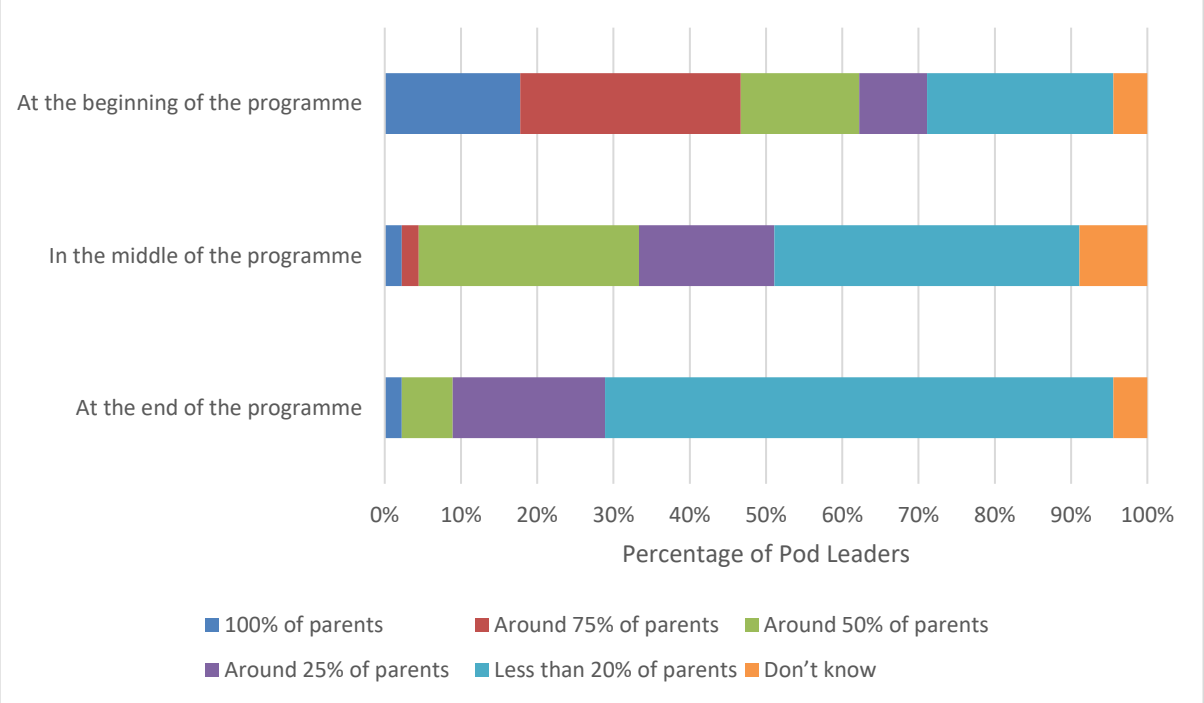
Pod Leaders reported in the case-study interviews that parents' engagement with the intervention did not differ according to socioeconomic background, and did not depend on whether the child had English as an Additional Language (EAL) or Special Educational Needs and Disability (SEND) status.

The EasyPeasy team provided data about technical failures during the trial. There were no web-app outages or technical failures with the sending of text messages to parents which could have impacted on their engagement. Between 1 May 2018 and 31 July 2018, 3.6% of parents were unable to receive messages on their phones. It is not known whether the 3.6% were parents who were participating in the trial.

[10] The developer reports that since the trial they have now changed the information available and now provide Pod Leaders with a breakdown of game activity at a parent level via their Pod Leader dashboard.

Figure 3: Pod Leaders' perception of parent engagement with EasyPeasy

Measured from responses to the Pod Leader practitioner survey administered to nurseries at the end of the intervention period (n = 45). Pod Leaders were asked approximately what percentage of the parents in their Pod accessed EasyPeasy at the beginning, middle, and end of the programme?



Quality

Here we discuss the key themes from the process evaluation that relate to the stakeholders' perceptions of the EasyPeasy intervention. These have been categorised as applicability of activities, parent delivery, and adaptability.

Research questions	Dimensions
What are the different stakeholder viewpoints on the intervention— a) for setting practitioners? b) for parents?	Participant responsiveness; programme differentiation
How effective and appropriately-pitched are the activities— a) for parents to deliver? b) for children to receive?	Quality
What is the acceptability of the intervention to parents and does this differ depending on the socioeconomic status of the parents (based on their child's eligibility for Early Years Pupil Premium), their child's Special Education Needs Diagnosis, or English as an Additional Language status?	Reach

Applicability of activities

Two thirds of Pod Leaders (30/45, 67%) felt that most games were appropriate for three- to four-year-olds while the remaining third (n = 13/45, 29%) felt that the games were too easy. During case-study visits, some Pod Leaders acknowledged that whilst the EasyPeasy games focused on different areas of learning that benefited school readiness, generally it was felt the games were too easy for nursery children.

> *'I would probably use it with my new children, so my younger children who are just turned three, and I would use it for them and I would also use it for my target children who are scoring quite low in language … not just in language but overall, I suppose, even personal, social and emotional … because each game doesn't just focus*

on language and one week it might be maths and [the next week it might be different], so it does focus on different areas of learning.'

Pod Leader, longitudinal case-study Pod Leader interview, interviewed at the end of the intervention.

During focus groups, some parents also stated the games were better pitched for preschool children as opposed to those in reception (an age range which EasyPeasy is also targeted at):[11]

'I would say [the games are best suited to] preschool when you are trying to get them to learn, or to take turns, or to wait for instructions, or to listen, really.'

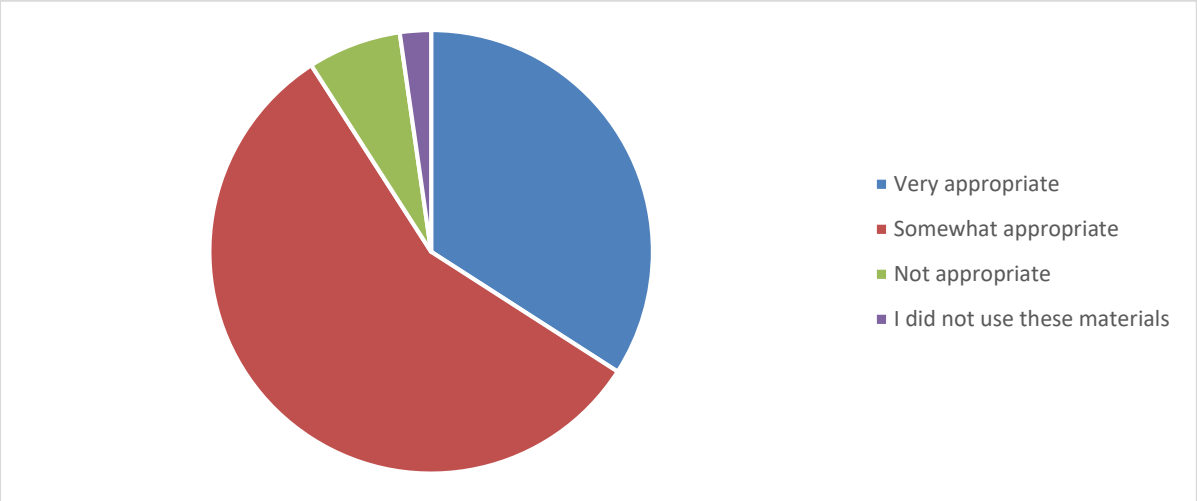
Parent, longitudinal case-study parent focus group, interviewed at the end of the intervention.

Parent delivery

During focus groups, parents discussed the practicalities of the EasyPeasy activities. Parents had positive experiences accessing and engaging with EasyPeasy through their mobile phones. During the focus group at Setting E, parents commented on the convenience that this brought noting that you can 'take the phone anywhere and you can do it anywhere'. Parents reported that the engagement with EasyPeasy was easy as the instructions were designed in a 'step-by-step' way to make all activities easier to understand and the resources needed for games can be found at home or bought easily. As detailed in Figure **4**, this was echoed by 80% of Pod Leaders who indicated on the practitioner survey that the instructions included in parents' welcome packs were pitched at the right level for parents.

Figure 4: Pod Leaders' view of the appropriateness of the instructions included in EasyPeasy welcome packs for parents

Measured from responses to the Pod Leader practitioner survey administered to nurseries at the end of the intervention period (n = 44).



Parents also commented that the games did not take a long time to play (two to three minutes) which they saw as a positive and inspired them and their children to invent more games of a similar kind.

'We've made things … I've made boats and things because one of the games was about making objects and so we make things more than probably we would have before.'

Parent, best practice case-study parent focus group, interviewed at the end of the intervention.

[11] The developer reports that, subsequently, the trial content has been divided into two programmes: one that is recommended to be used with nursery-aged children and another to be used with reception-aged children.

'There [is] stuff on [EasyPeasy] that takes nothing, and you have the stuff around the house, but you wouldn't think to do that yourself.'

Parent, best practice case-study parent focus group, interviewed at the end of the intervention.

Parents in case-study school D reported that they would use EasyPeasy at home if they had no plans or if the weather was not suitable for going out.

EAL parents participating in the parent focus groups reported that the games themselves were simple to understand, that the instructions were easy to follow, and that the oral English in the example video clips provided by EasyPeasy were also easy to comprehend. One family in the EAL category at case-study School B reported in the parent focus groups that the games could be played in their native language, and they found them to be useful outside of the English context too.

One SEND parent felt that the games helped children participate better in wider home life, for example, one of the games that required children to sort washing into colours and pairs, even though it took time for the child to understand.

'The laundry … you sort your washing into colours and pairs and that is why we do it with the socks 'cos we can sort it together. So, I will go to her: "You get all of [name] socks and you have to find the matching ones", and she has five different pairs of navy ones with different colour stripes on, and so she has to find each one and then there are daddy's socks, which are all the same … she then cheats, but do you get what I mean. It does take her a while, but she does like doing it, to be honest.'

Parent, cross-sectional case-study parent focus group, interviewed at the mid-point of the intervention.

Adaptability

The main element of the intervention that parents and nurseries perceived to be adaptable was the level of involvement of the nurseries in supporting parents to use the EasyPeasy games. As described earlier, this stemmed from how the Pod Leaders perceived their role, either as a facilitator to help parents use EasyPeasy and answer their questions, or alternatively as a leader proactively combining EasyPeasy games with classroom activities. Some nurseries reported providing significant structured support to parents to carry out the activities with their child in the nursery setting first (via Stay and Play sessions), whilst other settings expected parents to carry out the activities in isolation at home.

'So every morning and the beginning of every afternoon session on a Wednesday, parents are encouraged to come in and stay for half an hour and learn with their children and we model and demonstrate some things and we sit with the parents and we talk to them and help them to help their children … What we have done is we've also had the EasyPeasy game of the week on the whiteboard, modelling it with children with the parents there and we say, "Come and try this with your child as this is the game of the week and this is how you will do it." Some of the activities we then have a pack and say, "Go and try it at home." One example is when there was one where you draw around your child and so we had big rolls of paper and we just handed those out and say, "This is how you would do it, and there is your paper and you can try this at home."

Pod Leader, best practice case-study Pod Leader interview, interviewed at the end of the intervention.

Nurseries also adapted the intervention to use the EasyPeasy games within their own setting, in classroom-based activities with or without parental involvement.

Perceived impact of the intervention

Here we discuss key themes emerging from the process evaluation that relate to the perceived outcomes of the EasyPeasy intervention. These are categorised as children's outcomes, impact on parents, and impact on nursery staff.

Research question	Dimensions
What are the different stakeholder viewpoints on the intervention?	Participant responsiveness; programme differentiation

Children's outcomes

Most of the Pod Leaders from the case-study schools held a positive view about the potential of EasyPeasy, acknowledging the quality of the games. Several Pod Leaders expressed a view that EasyPeasy would work best at the beginning of the spring term, targeting those who struggled with confidence or with following instructions.

One parent stated:

> *'[Child] is really quiet and he doesn't like to speak out, and if somebody in the street approaches him, he will shy away and he wouldn't speak and so because [in the game] for example, the stepping stones, he has to shout it out, it's helping him show his voice and shout things out.'*
>
> *Parent, cross-sectional case-study parent focus group, interviewed at the mid-point of the intervention.*

Pod Leaders noted that the programme would be beneficial for Personal, Social and Emotional Development (PSED) in early years as they expected it to help with the transition to school and with skills such as confidence, following rules, and boundaries. However, they were not confident of the real depth of the programme's impact as the level of active parental involvement in using EasyPeasy recorded in the weekly reports did not match the expectations of the Pod Leaders (discussed earlier), even for the two best practice schools.

Pod Leaders were asked to report on the impact that EasyPeasy had on a range of areas (Figure **5**). Considering children's school readiness skills, 45% (n = 20/44) of Pod Leaders reported that EasyPeasy had had a mostly positive impact. However, 23% (n = 10/44) did not know whether it had had an impact.

Impact on parents

One of the initial aims of the intervention was to improve parent-child communication. Results from the Pod Leader practitioner survey indicated that the area in which the highest percentage of Pod Leaders reported a positive impact was in parents' communication and relationship with their children, where 59% (n = 26/44) reported a mostly positive impact. However, some parents participating in the focus groups reported that there had been no impact from using EasyPeasy on increasing or improving communication between themselves and other participating parents.

Regardless, parents in the focus group valued the opportunity EasyPeasy gave them to:

- do things with their child;
- compare what other children could do with what their child could do;
- find directions in how they might help their child to do better;
- gain awareness of some of the skills required for when their child moves in to Reception;
- understand the importance of communication at home, understanding 'how much they knew and how much they can share with each other'; and
- interact at home to give their child confidence to speak out, and to give them more patience to listen.

Around half of the Pod Leaders (n = 23/45, 51%) received mostly positive feedback from parents regarding the intervention. Of the remaining Pod Leaders, n = 15/45 (34%) received mixed (positive and negative) feedback and n = 6/45 (14%) received no feedback from parents. No Pod Leaders received mostly negative feedback from parents. Pod Leaders reported that parents often left positive comments in the dashboard after playing the games, including with

comments on children's self-regulatory skills. Pod Leaders also reported that when parents engaged with the EasyPeasy games, the parents valued the time they had spent with the child, and that, as a result, it was possible that the parents would change their stereotypical impressions of the use of information technology as a 'babysitting tool':

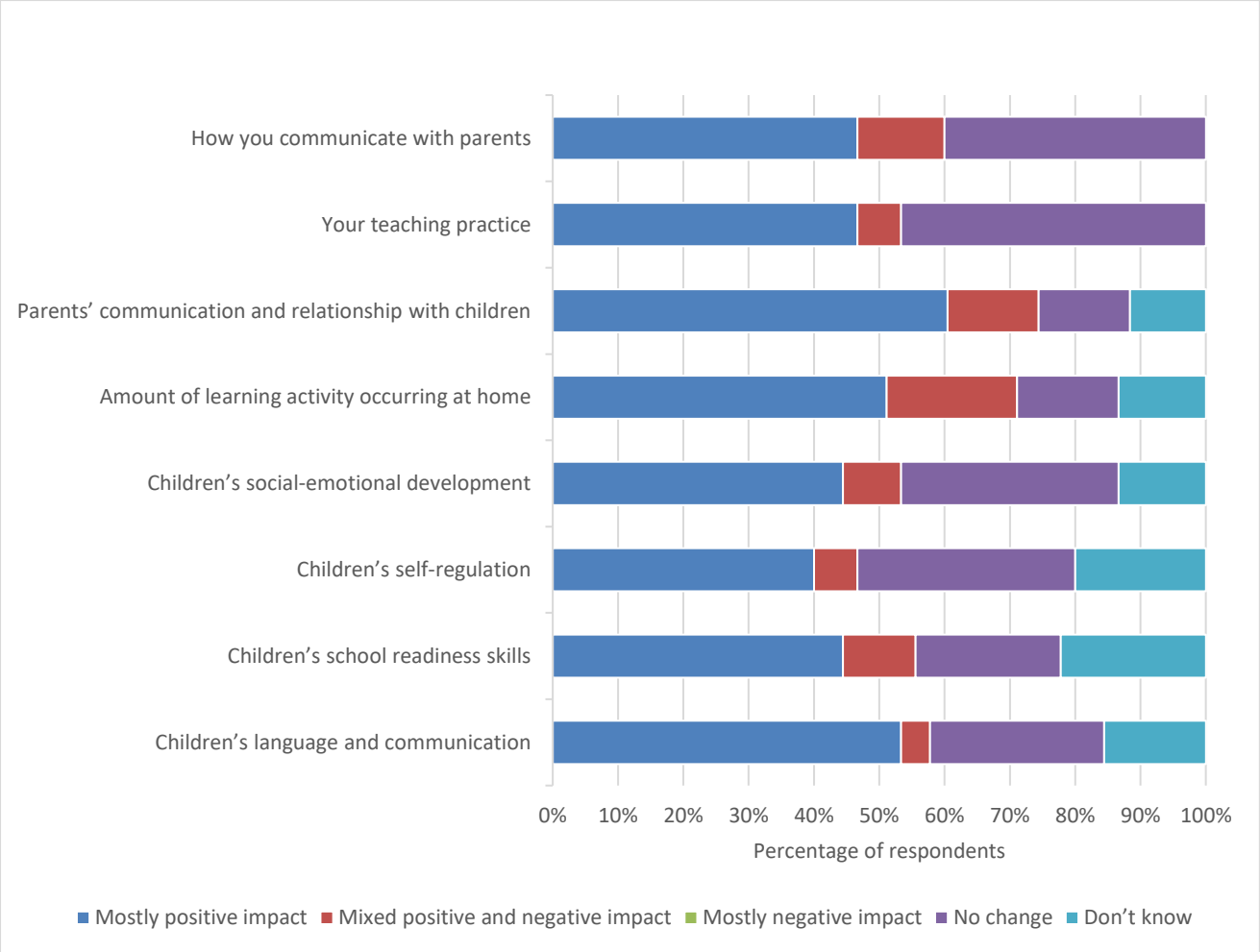
> *'Actually to sit and have a conversation and a game with them is more important than an app because it can be by chance sometimes on an iPad and it is that communication and that language that our children haven't got that they need. So, it is all of that and the parents modelling and showing them.'*
>
> *Pod Leader, longitudinal case-study interview, interview conducted at the beginning of the intervention period.*

Impact on nursery staff

The impact of EasyPeasy on teaching practice showed a mixed picture (Figure **5**). Here, 49% (22 out of 45 Pod Leaders responding to the Pod Leader practitioner survey) indicated that EasyPeasy had had a mostly positive impact on their teaching practice; 47% (n = 21/45) reported that there had been no impact on their teaching practice; and 40% (n = 18/44) reported that how they communicate with parents had not changed as a result of implementing EasyPeasy. This was supported by the findings of the longitudinal and best practice case studies and usual practice survey, where the schools considered that they already had good communication links with parents.

Figure 5: Pod Leaders' perceived impact on a range of outcomes from using EasyPeasy

Measured from responses to the Pod Leader practitioner survey administered to nurseries at the end of the intervention period (n = 44).



Pod Leaders were also asked to consider feedback they had received from other colleagues within their setting in relation to EasyPeasy. Half of all Pod leaders (n = 23/44, 52%) received mostly positive feedback regarding EasyPeasy, 20% (n = 9/44) received both positive and negative feedback, and 7% (n = 3/44) mostly negative comments. The remainder of Pod leaders did not receive feedback from other staff. Of the positive comments received by Pod Leaders from other colleagues in their nurseries, the most common feedback was that the games were simple and useful:

> *'[The games were] well pitched, clearly explained, really easy to engage with in the home and with some fantastic benefits to those who engage.'*
>
> *Pod Leader, Pod Leader practitioner survey open response, completed at the end of the intervention period.*

Of the Pod Leaders responding to the practitioner survey, 75% (n = 33/44) reported that they would probably or definitely recommend EasyPeasy to other nurseries. Reasons for recommending EasyPeasy included:

- the quality of activities, with activities being 'well pitched, clearly explained, really easy to engage with in the home';
- that EasyPeasy would be 'good for children's learning and development';
- that it was 'a really good idea for parents who need support with communicating with their child';
- that it was 'very clear and simple to use for both parents and practitioners';
- that it improved parental engagement with the nursery as it was 'easy to monitor which parents are engaging and easy to give game tips and extension ideas so they fit with the work happening in class';
- that 'it is a positive way for children and parents to positively engage with technology'; and
- a range of unspecific positive comments about engaging with the EasyPeasy programme.

The remaining 25% of Pod Leaders (n = 11/44) reported that they probably would not recommend EasyPeasy to other nursery settings. Reasons for not making a recommendation were varied and included:

- that parents 'hadn't responded as [the nursery] had hoped despite being reminded' and that parents that did sign up 'didn't take part in the games each week';
- that 'the parents who did sign up were the ones who do not need the support';
- that the Pod Leader did not feel that the nursery, the parents, or children 'got much out of the whole experience';
- that 'although EasyPeasy may have had a positive impact on [their nursery's] children/families, it is difficult for [the Pod Leader] to show that in [their] nursery setting, and, therefore, [the Pod Leader] wouldn't recommend it as [they] can't prove it added value';
- one Pod Leader had received feedback from 'parents [who] explained that they would not recommend to other parents';
- that games were more suitable for younger children'; and
- that the administrative requirements of the evaluation were high: 'The project has been a lot of extra work, emails, filling-out paperwork etc. for possible "impact" to measure.'

Other factors affecting the perceived outcomes of the intervention

Here we discuss key themes emerging from the process evaluation that relate to the delivery and to the intervention itself. These are categorised as factors for successful delivery, barriers to successful delivery, and considerations for future improvement.

Research questions	Dimensions
What are the key success factors required for the EasyPeasy intervention to work well?	Quality
What are the barriers to successful delivery of the intervention— a) for parents? b) for setting practitioners?	Quality; monitoring of control/comparison
What areas of the programme could be further developed following completion of the project?	n/a

Factors for successful delivery

Pod Leaders suggested that two main factors affected successful parental engagement; these both involved forming closer links between the nursery and parents:

1) regular communication and reminders from the nursery for parents to use EasyPeasy; and
2) engagement by the nursery in the use of the EasyPeasy games as a link between home life and nursery practice.

As well as explaining the methods used to implement EasyPeasy within their setting, Pod Leaders were also asked to reflect on implementation.

> *'[At the start of the programme] you would invite all the parents in for a morning session and then we could all watch [the game] together on the big screen and then have the equipment out and the parents play with the children whilst you are there, and then perhaps you could have some display with the pictures and then maybe you could put up a chart that every week when they come in they could get a sticker that they could put on the chart. Maybe every half term, if there was a certificate we could print off to send home if they had done four out of the five weeks or … we just need to reward them, I think, and then it would be a bit more obvious who has and who hasn't [used EasyPeasy], and I think if you had a chart they might start talking to other parents a bit more.'*
>
> *Pod Leader, longitudinal case-study interview, School B, interviewed at the end of the intervention period.*

The developer commented in the post-intervention interview that a strength of the EasyPeasy programme is in the straightforward sign-up process—also to the training session offered to Pod Leaders when setting up the programme.

Barriers to successful delivery

Pod Leaders felt that the main barrier to delivery was a lack of time (Pod Leaders reported in the Pod Leader practitioner survey that it took approximately an hour a week to coordinate the use of EasyPeasy within their setting). The nurseries had many other activities or priorities during term time and home life was normally busy for parents.

Other barriers to successful delivery raised by Pod Leaders in the case-study interviews were:

- Language barriers. Pod Leaders commented that '[parents] do need a lot of support and the language barrier is huge'.
- A lack of information to be able to effectively discuss EasyPeasy with parents. Pod Leaders commented that more guidance to help explain EasyPeasy to parents would have been beneficial. 'We have difficulties sometimes with being part time and one in afternoons and another teacher doing mornings, with the handover and the information anyway … you kind of feel you are bombarded with the amount of information and then you are passing that information when maybe you haven't actually had the full background to what that is about anyway so then [parents] are asking you follow-up questions and you have no idea what you are talking about.'
- That some settings had not involved all the staff within a nursery in EasyPeasy and in other nurseries there was change in staff within the nursery and this was seen as less helpful for monitoring the programme at pick-up and/or drop-off times. 'There [are] lots of changes and changes of staff, so we would need to speak to the staff involved.'
- Change of Pod Leaders during the 20-week intervention with no handover for the details of the intervention. One Pod Leader commented that they felt they had 'been left on a bit of an island with it and it is just, "Can you deal with that?"'
- Lack of demographic data in weekly email (as noted above).

Pod Leaders felt that the best time for starting the intervention would be in the preparation period for nursery, that is, before the start of the academic year.

'When [EasyPeasy] went live it was immediately after a half term and so I was trying to input the parents data [into the dashboard] but because of how busy it was I couldn't input it until I had time to do it, which was over half term which meant I couldn't tell the parents exactly what was going on and so I sent them all a text and so they did get a text and they were told any further questions to come and speak with [me] and nobody bothered to come and ask me about it. Whereas if I had been able to give them a bit more of a heads up about it then maybe that would have helped a little bit more.'

Pod Leader, best practice case-study interview, interviewed at the end of the intervention period.

Some parents participating in case-study focus groups listed a further two barriers when running the programme at home:

- some games were not easy to understand; and
- there were a few technical issues, such as long login times or errors on web pages.

'Sometimes it doesn't download on my phone and it just says this page has an error and then I am closing it and having to open it…'

Parent, cross-sectional case-study parent focus group, interviewed at the mid-point intervention period.

In the post-intervention interviews with the developer, the developer expressed concerns about communication difficulties with schools while implementing the evaluation project, especially around integrating the intervention into the school's normal practice.

It worried that these difficulties would lead to negative impacts on:

- research results, for example, sample size was not sufficient to reflect the results, particularly as the initial 134 schools to sign up translated into 101 schools completing the trial;
- EasyPeasy's relationship with schools; and
- the reputation of EasyPeasy

The developer explained, for example, that Pod Leaders' confidence might have been damaged during the pre-testing period due to concerns having been expressed by Pod Leaders around the delivery of the tests. For example, some Pod Leaders were concerned that children did not understand the accents of the people who delivered the tests and were concerned about the way that children responded to the test questions. Such concerns are not unique to this trial and to mitigate against these concerns, pre-agreed assessment plans, detailed training of assessors, and monitoring assessor quality were all implemented to provide confidence in the assessment process. A sensitivity analysis has been conducted in the impact evaluation analysis (above) using quality scores assigned to test administrators as part of the quality assurance process, to minimise the impact of variation between administrators.

Considerations for future improvements to the intervention

This section considers whether there are any ways that the intervention could be improved. It includes suggestions from Pod Leaders, parents, and the developer.

Suggestions from Pod Leaders to improve the EasyPeasy intervention included:

- using the intervention with two-year-olds;
- differentiating the games more widely to provide parents with more examples of how to play the games with higher- and lower-ability children or older and younger children (as the current differentiation was commented to be too basic); and
- using EasyPeasy particularly for children who score low in language and struggle with confidence.

The developer suggested in its post-intervention interview that it supported the idea of 'stay and play' sessions in school and that future use of the intervention should see it carried out in an integrated way, as a 'whole-school community type' of intervention, rather than one targeting a specific group in the setting or as a subset of normal practice.

Parents indicated different preferences regarding improvements to the EasyPeasy games. Some reported that they would like to have more games featuring practical skills such as measuring or more challenging games such as numbers, adding, or subtracting. Others favoured games in which older siblings could participate, and the desire was also expressed for games helping to develop their child's imagination. The developer reported that the number of games in the EasyPeasy programme has been increasing. During the academic year in which the intervention took place, EasyPeasy increased the number of games available from 45 to 65.

The developer also felt that if the local leadership prioritised early years education, home learning, and parent engagement, there would be a strong influence on headteachers' decisions about integrating the EasyPeasy programme into the nursery's normal practice. As a result, leadership decisions affect the nature of a whole school's engagement with the EasyPeasy programme.

The developer remarked that policymakers would play a crucial role in the future direction of the programme. Without coordinated support for the programme at this level, headteachers and practitioners alone would have few resources (time and money) to invest in the intervention. EasyPeasy report that the project has now reached over 10,000 homes and nearly 300 early years settings.

Up to this point, it has been EasyPeasy's policy to work in partnership with schools with a high percentage of children eligible for the Pupil Premium, and it expects that this will continue in the future. The developer was not considering working directly with parents on a B to C (business to consumer) model in the near future.

The developer suggested that to achieve higher parental engagement, practitioners and teachers would need to be equipped with the 'contextual intelligence' to combine the programme with their local knowledge and their own expertise in child development.

Control group activity

This section considers the 'business as usual' practice that was ongoing in the control schools during the period of the intervention. Business as usual was assessed using a usual practice survey, completed by both control and intervention schools at the beginning and end of the trial.

Pod Leaders in 96 out of 102 (94%) nurseries (n = 49 intervention, n = 47 control) completed the usual practice survey at the beginning of the trial (pre-randomisation), while 71 out of 101 (70%) were received in the post-intervention period (n = 38 intervention, n = 33 control). This provided information about the activities of the control group compared to the intervention group during the intervention period.

As far as we are aware, there was no sense of resentment or demoralisation within the control group, and schools in this group continued to deliver standard provision without there being any drop-outs from the study as a result of their random allocation. However, one school in the control group was closed down after the pre-randomisation testing period, and before the intervention started. All the participating children at this school were sent to two intervention schools in the same area. Another school withdrew during the post-intervention period; the school did not feel it had time to conduct the post-intervention data collection as it already had another three projects running at the same time.

Usual practice surveys with nurseries in both the control and intervention group were conducted as part of the implementation and process evaluation. Findings suggest that there was little difference between the intervention and control schools in the contact methods or the frequency of contacts used in the schools. This situation did not change significantly at the post-test stage. The differences are very small and consistent between the control and intervention groups.

The most common methods used by nurseries to communicate with parents (in both the control and intervention schools) was face-to-face, telephone, and parents' evenings, with over 80% of nurseries reporting using these communication methods.

Before the intervention started, 66% of control nurseries and 71% of intervention nurseries reported sending activities home for parents to work on with their children daily or weekly. This showed no significant change over the course of

the trial, with 67% of control schools and 74% of intervention nurseries reporting sending activities home for parents to work on with their children daily or weekly at the end of the intervention period.

As EasyPeasy is an online product, it was interesting to investigate how many nurseries were already using an online product for communication with parents: 23% of control nurseries and 35% of intervention nurseries already used an online learning journal to communicate with parents daily or weekly. This increased to 33% of control group nurseries post-intervention and remained constant at 32% of intervention nurseries.

EasyPeasy sends a new activity weekly to parents, and many Pod Leaders in the intervention group chose to provide additional reminders and activities for parents to encourage engagement. Therefore, the frequency of communication between nurseries and parents, and the means of this communication, was also interesting to investigate. The most commonly used methods for communicating with parents on a daily or weekly basis was face-to-face (98% of control schools pre-test and 100% post-intervention; and 100% of intervention nurseries pre and post-intervention) and by letter (77% of control nurseries pre-test and 77% post-intervention; 69% of intervention nurseries pre-test and 74% post-intervention).

Pod Leaders were asked whether parents were part of a supportive social network: 34% of control schools said that, in general, most parents/guardians were part of a supportive social network at the beginning of the trial; this had increased to 52% at the end of the intervention period. At the start of the trial, 45% of intervention schools reported that, in general, most parents/guardians were part of a supportive social network. This remained consistent throughout the period of the trial (47% at the end of the intervention period).

Conclusion

Key Conclusions	
1.	Children in schools receiving EasyPeasy did not make any additional months' progress in language development compared to children in control schools, as measured by a composite, summary language score. This finding has a moderate to high security rating.
2.	There were small increases in 'word structure' and 'concepts and following directions' language subscales (equivalent to one month's additional progress) compared to children in control schools, but no additional months' progress in 'sentence structure' or 'expressive vocabulary'.
3.	Mixed results were found for children's social, emotional, and behavioural outcomes. Children who received EasyPeasy made small increases in sociability, cognitive self-regulation, and emotional self-regulation compared to the control group. However, effects on externalising, internalising, and prosocial behaviour and behavioural self-regulation favoured the control group.
4.	Parents receiving EasyPeasy reported improvements in the home learning environment. This included large increases in 'modelling', 'responsivity', and 'variety of activities and interactions'. These results are less secure than the main findings due to the small number of parents assessed, and should be treated as exploratory.
5.	Engagement from parents for the continued use of EasyPeasy was considered to be a challenge for nurseries. The most effective ways of encouraging parent participation included integration of the games into the classroom, introducing parents to the games at 'Stay and Play' sessions, and parents sharing comments.

Interpretation

The EasyPeasy intervention aimed to improve parental engagement with their children, and through this, to accelerate children's cognitive development, self-regulation, and language and communication skills. Previous smaller-scale trials (Jelley, Sylva and Karemaker, 2016; Jelley and Sylva, 2018) had found EasyPeasy to have a positive effect on parent-reported child cognitive self-regulation and on parents' self-efficacy regarding their sense of control (discipline and boundaries).

The results of this large-scale cluster randomised controlled efficacy trial found no statistically significant evidence that EasyPeasy is effective at improving language skills in children aged three to four. The observed effect size was positive but very small, and did not equate to any months' additional progress (effect size 0.04, 95% CI -0.10 to 0.18). The difference was not statistically significant. Very small, non-statistically-significant effects (effect sizes ≤0.06) in favour of the intervention group were observed in the individual CELF subscale scores (Sentence Structure, Word Structure, Expressive Vocabulary, Concepts, and Following Directions), each equating to a month or less of progress. Similarly, small effect sizes in favour of the intervention group were found for the CSBQ subscales of Sociability, Emotional self-regulation, and Cognitive self-regulation (effect sizes 0.04, 0.06, and 0.14, respectively) and no consistent patterns were seen. Conversely, small effect sizes in favour of the control group (without access to EasyPeasy) were found for the Externalising, Internalising, Prosocial behaviour, and Behavioural self-regulation subscales (effect sizes -0.08, -0.08, -0.02, -0.02, respectively). None of the differences in CSBQ outcomes between groups were statistically significant.

Participating in the intervention did make a significant difference to the HOME subscale measures of Modelling, and Variety of activities and parental interaction (effect sizes 0.66, and 0.72, respectively). The overall HOME score also favoured the intervention group, and the difference was statistically significant (effect size 0.76, 95% CI 0.11 to 1.42, $p = 0.02$). Reasonably large effect sizes were also observed for the remaining four HOME subscales (ranging from 0.42 to 0.59), but these were not statistically significant. The HOME analyses were only conducted on a small subset of participants, by design, and so were underpowered.

The impact of EasyPeasy on the post-test Core Language Standard Score was not seen to be substantially altered by having EYPP status (interaction effect $p = 0.58$), having EAL (interaction effect $p = 0.99$), or by gender (interaction effect $p = 0.12$). EAL parents reported that the EasyPeasy games were simple to understand, instructions were easy to follow, the oral English in the example video clips provided by EasyPeasy was easy to comprehend, that games could be played in their native language, and the games were useful outside of the English context.

All results were robust to sensitivity analyses, including repetition of the primary ITT analysis with an additional covariate for the grading of the assessor and including a random effect for assessor. Children in nurseries where the Pod Leader

attended the live webinar training tended to perform slightly better than in nurseries where the Pod Leader either didn't attend the training at all (control nurseries) or watched the recorded training, but the difference was not statistically significant. Just over half of Pod Leaders reported that they found the training very useful and felt very confident to delivery EasyPeasy. We acknowledge that defining compliance with the intervention as the Pod Leader attending the live webinar training, rather than watching a video recording, has some limitations. For instance, the video recording plus follow-up call and emails may have been comparable to attending the webinar. Other measures of compliance were considered, such as parental engagement in terms of how often they played the games with the child, but it was not possible to measure this at a pupil level.

It was not possible to gauge the impact of EasyPeasy by the level of parental engagement as the implementers of EasyPeasy were unable to provide this information to the evaluation team. However, the findings from the Pod Leader practitioner survey indicated that in general, at a Pod level, the numbers of parents accessing EasyPeasy reduced significantly during the intervention period. Although Pod Leaders and parents viewed the potential of EasyPeasy positively, acknowledging the quality of the games, Pod Leaders were not confident of the real depth of the programme's impact, as the level of active parental involvement in using EasyPeasy was lower than the expectations of the Pod Leaders. It may therefore be hypothesised that the impact of EasyPeasy may have been reduced through lack of engagement by parents with the intervention.

Pod Leaders perceived a range of purposes of EasyPeasy. Some Pod Leaders perceived EasyPeasy to be used by parents within the home, and others perceived it as a vehicle to help the school to engage parents and/or to link home and classroom activities. The perceived purposes of EasyPeasy ultimately impacted on how Pod Leaders viewed their role in association with the programme and how it was rolled out within the setting. Some Pod Leaders adopted a 'facilitator' role whereby they assisted parents in using EasyPeasy and answering their questions, whereas others adopted a 'leader' role proactively combining EasyPeasy games with classroom activities. The role that Pod Leaders adopted often determined their methods to promote and engage the use of EasyPeasy among parents. However, in all cases, Pod Leaders were in agreement that keeping parents engaged was the key to successful implementation of EasyPeasy. Pod Leaders acknowledged that they undertook additional approaches to keep parents engaged, or to re-engage those who were not active. These took the form of Stay and Play sessions, visits to homes, advice on how to differentiate the games, and the setting of games as homework. Pod Leaders also commented that engaging with the parents who would benefit the most from the intervention was often difficult and that the parents who did use EasyPeasy were not the ones who needed the additional support to engage with their children. Pod Leaders considered the main barrier to delivery to be a lack of time to engage with EasyPeasy, both by themselves and by parents, however, Pod Leaders also reported that when parents engaged with the EasyPeasy games, the parents valued the time they had spent with the child. The convenience of EasyPeasy being accessible on mobile phones and games using easily available materials was appreciated by parents.

EasyPeasy was generally perceived positively with around half of Pod Leaders receiving mostly positive feedback from parents regarding EasyPeasy. Two thirds of Pod Leaders felt that most games were appropriate for three- to four-year-olds while the remaining third felt that the games were too easy. Three quarters of intervention nurseries said that they would recommend EasyPeasy to colleagues in other nurseries. Several of the nurseries had adapted the intervention to use the EasyPeasy games within their own setting, in classroom based activities. Although this was not the prescribed method of using EasyPeasy at the start of the intervention, at the end of the intervention, the implementers of EasyPeasy considered that future use of the intervention should see it carried out in an integrated way, as a 'whole-school community type' of intervention, rather than one targeting a specific group in the setting or as a subsection of normal practice. It would be interesting to monitor whether increasing the time children spend interacting with EasyPeasy games and activities through integration of the programme into nursery practice increases the measurable difference on child self-regulation and language and communication compared to the light touch programme in the current trial.

Limitations

The timing of the trial proved challenging for recruitment of parents by nurseries and for the implementation of the intervention. Recruitment of nurseries commenced at the start of the academic year and due to a lack of uptake by parents within nurseries, was still ongoing at the point that pre-testing commenced. Pre-testing was therefore carried

out across a longer time-period than was originally planned. Nurseries were randomised in two batches, based on when they were pre-tested, to allow an earlier start date for the nurseries with earlier pre-testing dates. This also meant that post-testing would not be required up to the end of the academic year for all schools in the trial, only those in batch 2. The delay to the intervention starting for batch 2 schools meant that Pod Leaders were setting up EasyPeasy over the February half-term holiday and the first week of text messages from EasyPeasy were received the first week back after the holiday. Nurseries commented that this meant that they had less opportunity to promote EasyPeasy to parents and, therefore, less opportunity to encourage higher participation from the start. Nurseries commented that promoting EasyPeasy to parents before the start of the academic year and implementing it for all parents from the beginning of the academic year, may have led to higher engagement by parents.

The evaluation element of the trial created considerable additional administrative work for all Pod Leaders. Although participating schools (in both the treatment and control arms) received the use of EasyPeasy for free as an incentive to participate, Pod Leaders did not receive any additional incentive to take on the additional role as part in the project. During the study, all Pod Leaders were asked to collect parental participation agreement forms, to attend online training, to fill in the CSBQ questionnaire and the usual practice survey twice during the study, to provide child demographic information at the beginning of the intervention, and to provide data on children's school destinations after the intervention. In addition, Pod Leaders in the intervention schools also completed the practitioner survey at the end of the intervention. The Pod Leaders from the six case-study schools organised parent focus groups and were also interviewed themselves. This work was additional to their normal workload and extra support was not provided in nurseries. Concern was expressed by EasyPeasy that this may have led to nurseries seeing EasyPeasy as being more time consuming to administer than if they were using EasyPeasy under normal circumstances outside the trial. Pod Leaders from three schools commented in the Pod Leader practitioner survey that the additional evaluation component made them less likely to recommend EasyPeasy to colleagues in other nurseries.

As discussed in the interpretation section above, a limitation of the current evaluation was the lack of information available about the level of individual parental engagement with EasyPeasy. It is, therefore, not possible to identify what level of 'dose' of EasyPeasy may be required to be effective in improving children's cognitive development, self-regulation, and language and communication skills.

Future research and publications

Although no significant effect sizes were observed for the impact of EasyPeasy on children's cognitive development, self-regulation, and language and communication skills, the findings from the Pod Leader practitioner survey suggest that this may have been due to low parental engagement. In order to understand in detail whether the level of parental engagement with EasyPeasy has a significant impact on children's cognitive development, self-regulation, and language and communication skills, further research is required with detailed pupil-level information about the level of parental engagement.

The impact of EasyPeasy has, at present, only been measured immediately following the intervention. The longer-term impact of the programme on pupils' language and communication, and on self-regulation, will be measured using pupils' Early Years Foundation Scale Profile (EYFSP) data in November 2019. An addendum to the present report containing the analysis of the EYFSP data will be published in the spring of 2020.

References

Caldwell, B. and Bradley, R. (1984) *Home observation for measurement of the environment (HOME)* (revised edn), Little Rock, AR: University of Arkansas, Little Rock.

Dunn, G., Maracy M. and Tomenson B. (2005) 'Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods', *Stat Methods Med Res*, 14 (4), pp. 369–95.

Foster, M. A., Lambert R., Abbott-Shim, M., McCarty, F. and Franze, S. (2005) 'A Model of home learning environment and social risk factors in relation to children's emergent literacy and social outcomes', *Early Childhood Research Quarterly,* 20, pp. 13–36. DOI: 10.1016/j.ecresq.2005.01.006.

Hart, B. and Risley, R. T. (1995) *Meaningful differences in the everyday experience of young American children*, Baltimore, MD: Brookes.

Howard, S. J. and Melhuish, E. (2016) 'An Early years toolbox for assessing early years executive function language, self-regulation, and social development: Validity, reliability, and preliminary norms', *Journal of Psychoeducational Assessment*, 35 (3), pp. 255–275.

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': https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Guidance_Final.pdf

Jelley, F., Sylva, K. and Karemaker, A. (2016) 'EasyPeasy parenting app: Findings from an efficacy trial on parent engagement and school readiness skills': http://www.suttontrust.com/wp-content/uploads/2016/11/EasyPeasyEvaluation_V5_FINAL.pdf

Jelley, F. and Sylva, K. (2018) 'EasyPeasy Evaluation in Newham: findings from the Sutton Trust Parental Engagement Fund (PEF) Project': https://www.suttontrust.com/wp-content/uploads/2018/05/EasyPeasyNewham-FINAL.pdf

Lugo-Gill, J. and Tamis-LeMonda, C. (2008) 'Family resources and parenting quality: Links to children's cognitive development across the first 3 years', *Child Development*, 79, pp. 1065–1085.

Melhuish, E. C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B. and Phan, M. (2008) 'Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school', *Journal of Social Issues*, 64, pp. 157–188.

Melhuish, E. (2010) *Impact of the home learning environment on child cognitive development: Secondary analysis of data from 'Growing Up in Scotland'*, Edinburgh: The Scottish Government.

Merrell, C., Little, J. and Coe, R. (2014) 'Is the Attainment gap among primary aged children decreasing?', in Wood, C. and Scott, R. (eds), *Harnessing what works in eliminating educational disadvantage: A tale of two classrooms*, London: Demos.

Saghaei, M. and Saghaei, S. (2011) 'Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials', *Journal of Biomedical Science and Engineering*, 4, pp. 734–739. DOI: 10.4236/jbise.2011.411090: http://www.scirp.org/journal/PaperInformation.aspx?PaperID=8518

Senn, S. (1994) 'Testing for baseline balance in clinical trials', *Stat Med,* 13 (17), pp. 1715–26.

Snow, C. E., Burns, S. and Griffin, P. (1998) *Preventing reading difficulties in young children*, Washington, DC: National Academy Press.

Tamis-LeMonday, C. S. and Rodriguez, E. T. (2009) '"Parents" role in fostering young children's learning and language development', *Encyclopaedia on Early Childhood Development* (3rd edn), December 2014.

Totsika, V. and Sylva, K. (2004) 'The home observation for measurement of the environment revisited', *Child and Adolescent Mental Health*, 9 (1), pp. 25–35.

Tymms, P., Merrell, C., Hawker, D. and Nicholson, F. (2014) 'Performance indicators in primary schools: A Comparison of performance on entry to school and the progress made in the first year in England and four other jurisdictions: Research report', London: Department for Education: https://www.gov.uk/government/publications/performance-indicators-in-primary-schools

Wiig, E. H., Second, W. A. and Semel, E. (2006) 'CELF Preschool 2 UK Examiner's Manual (Second Edition)', London: Pearson Assessment, p. 130. ISBN: 978 0 749126 21 6.

Appendix A: EEF cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. More information about the EEF's approach to cost evaluation can be found **here**. Cost ratings are awarded as follows:

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

Appendix B: Security classification of trial findings

Security ratings provided by independent peer reviewers

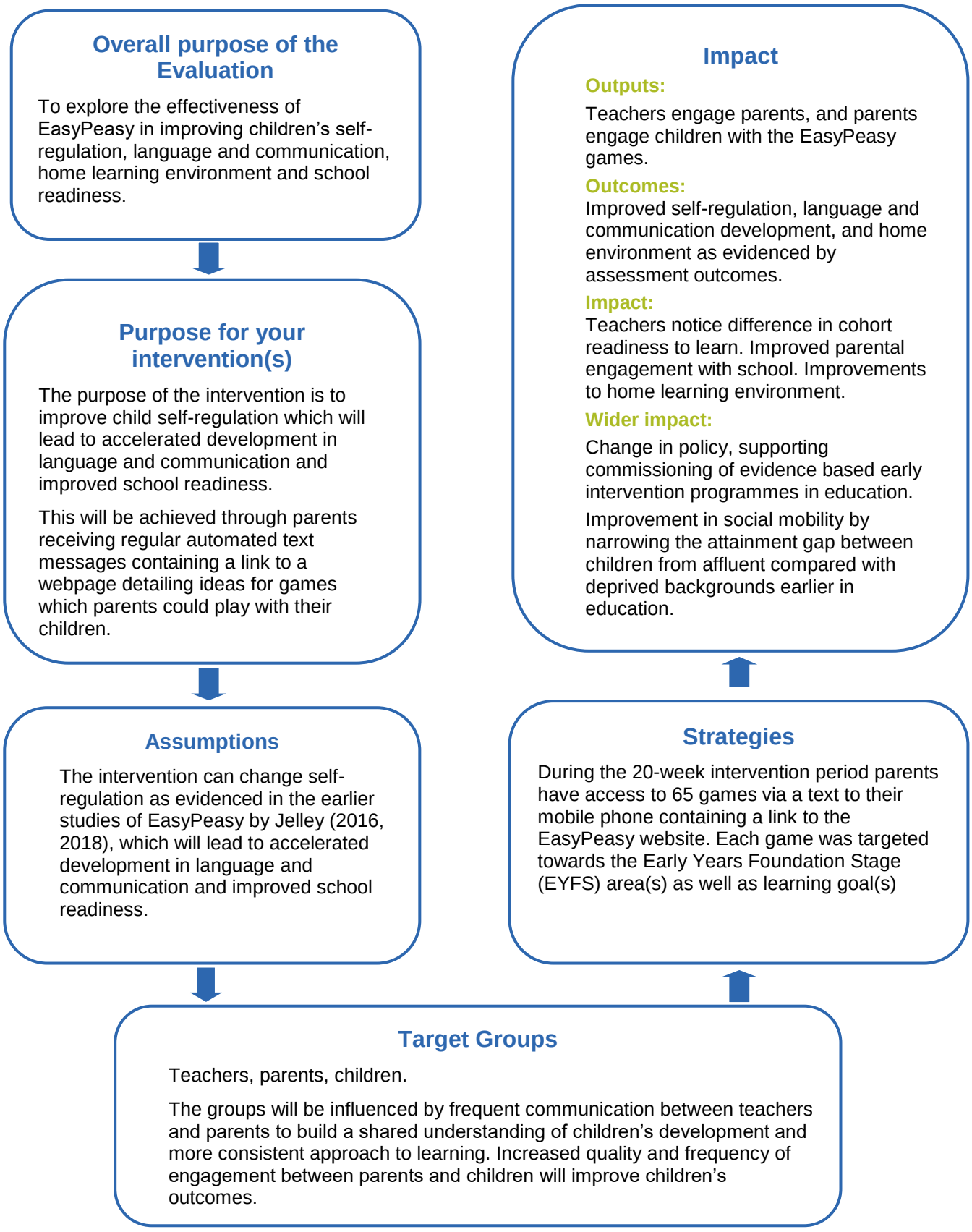
Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	Power	Attrition[12]					
5 🔒	Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%					
4 🔒	Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity	MDES < 0.3	11-20%	4		Adjustment for Balance **[-1]**		
3 🔒	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%		→		→	3
2 🔒	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			Adjustment for threats to internal validity **[0]**		
1 🔒	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%					
0 🔒	No comparator	MDES > 0.6	over 50%					

- **Initial padlock score:** lowest of the three ratings for design, power and attrition = The design is a randomised controlled trial which has been powered to 0.226 and there was pupil level attrition of 6.4%. Therefore, the initial score for the trial should be 4 padlocks.
- **Reason for adjustment for balance** (if made): There was imbalance of pre-test data of more than ES 0.05 and there was also imbalance in Early Years Pupil Premium scores. Therefore 1 padlock reduction is recommended.
- **Reason for adjustment for threats to validity** (if made): No threats to validity are present so no adjustment is needed.

Final padlock score: initial score adjusted for balance and internal validity = 3 padlocks.

[12] Attrition should be measured at the pupil level (even for clustered trials) and from the point of randomisation to the point of analysis.

Appendix C: Theory of Change



Appendix D: Model fit summary

Figure 6 and Figure 7 below are a histogram and qq plot of the standardised residuals from the primary analysis model. The symmetry of the distribution displayed in the histogram, and the closeness of the points to the line in the qq plot indicate that the assumption of the normality of the residuals is met.

Figure 6: Histogram of the standardised residuals from the primary analysis model.

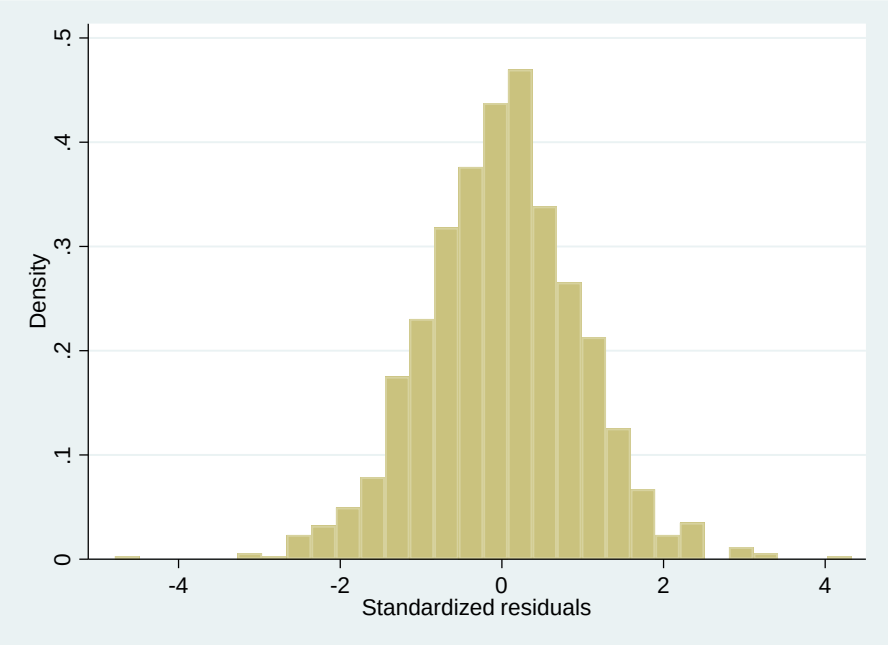


Figure 7: QQ plot of the standardised residuals from the primary analysis model.

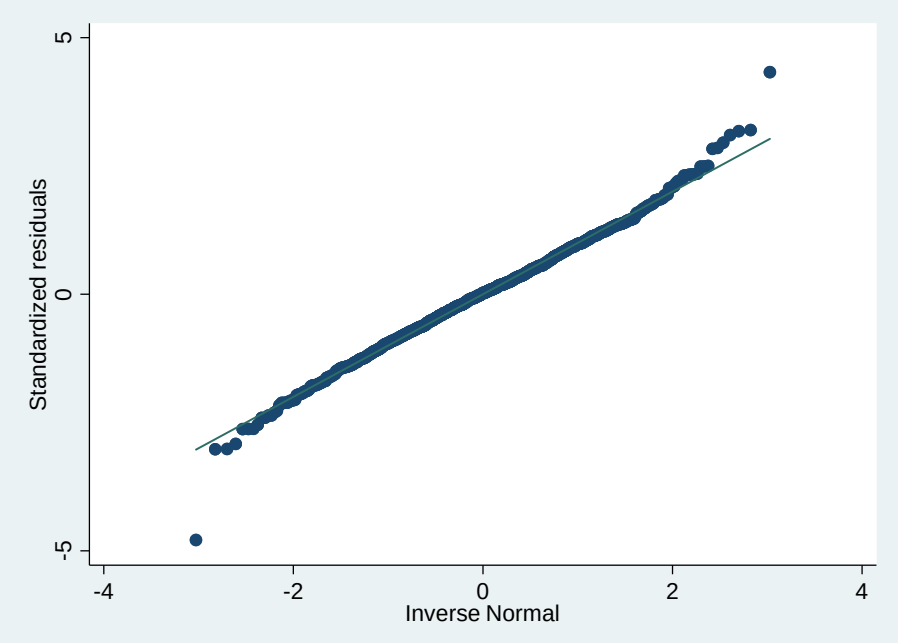
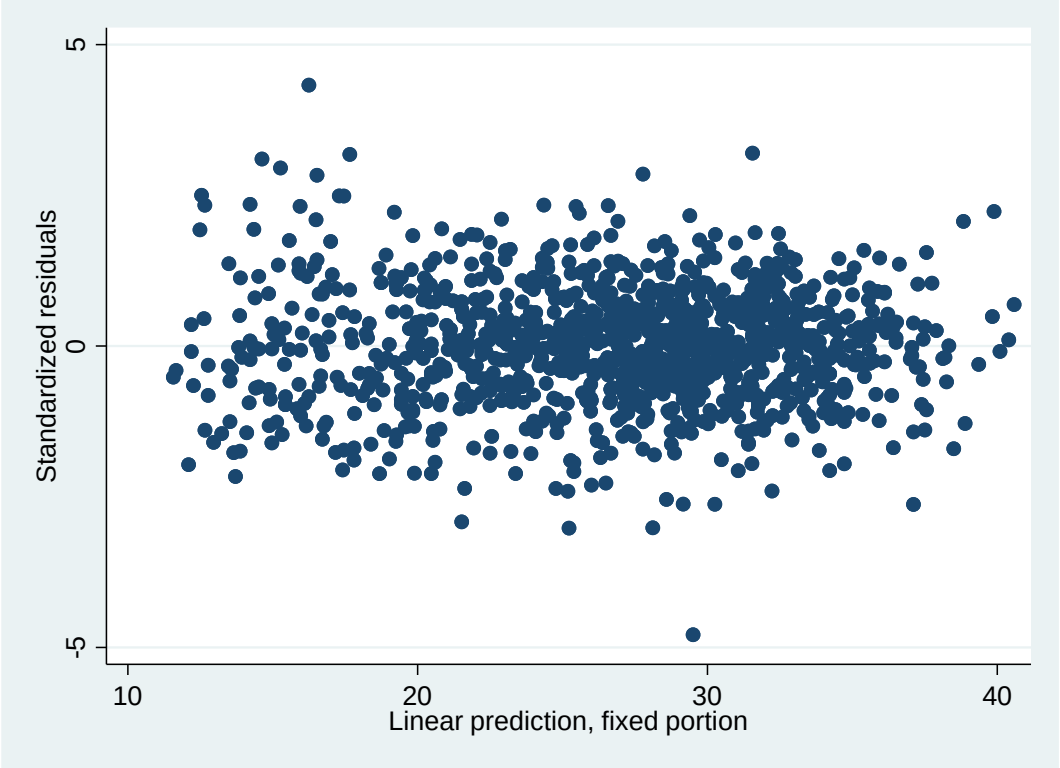


Figure 8 below is the scatter plot of fitted values against the residuals, from the primary analysis model. The random displacement of points in a rectangular shape with no clustering or systematic pattern demonstrates that the assumption of homoscedasticity is met.

Figure 8: Scatter plot of fitted values against the residuals from the primary analysis model.



Appendix E: CELF Preschool 2 UK raw score summary

Table 25: Summary of post-test CELF- Preschool 2 UK outcome measure raw scores.

	Raw means			
	Intervention group		**Control group**	
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**
Sentence Structure (raw)	562 (33)	13.2 (12.8, 13.5)	567 (43)	12.9 (12.5, 13.2)
Word Structure (raw)	562 (33)	12.8 (12.4, 13.2)	566 (44)	12.0 (11.6, 12.4)
Expressive Vocabulary (raw)	561 (34)	17.8 (17.2, 18.4)	563 (47)	16.8 (16.2, 17.4)
Concepts and Following Directions (raw)	553 (42)	10.1 (9.8, 10.5)	558 (52)	9.3 (8.9, 9.6)

Additional Appendices

Additional appendices are available in a separate document, which includes:

1. Memorandum of Understanding for schools (excluding Doncaster LA)
2. Memorandum of Understanding for schools (Doncaster LA)
3. Parent participation agreement
4. Parent home visit participation agreement
5. Pod leader case study interview participation agreement
6. Parent case study focus group participation agreement
7. Pod leader practitioner survey participation agreement
8. Developer interview participation agreement
9. General Pod Leader privacy notice
10. General parent privacy notice
11. Parent home visit privacy notice
12. Pod Leader case study interview privacy notice
13. Parent case study focus group privacy notice
14. Developer interview privacy notice
15. Usual practice survey (pre and post)
16. Pod Leader practitioner survey
17. Longitudinal case study Pod Leader interview schedule
18. Cross-sectional case study Pod Leader interview schedule
19. Best practice case study Pod Leader interview schedule
20. Longitudinal case study parent focus group interview schedule
21. Cross-sectional case study parent focus group interview schedule
22. Best practice case study parent focus group interview schedule

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

Education Endowment Foundation

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

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn