FLEXIBLE PHONICS, A TWO-ARMED CLUSTER RANDOMISED TRIAL

Evaluation Report

October 2023

Updated March 2024

Anneka Dawson, Clare Huxley, Helen Gray, Kate Alexander, Charlotte Edney, and Susie Bamford

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

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

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

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

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

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The evaluation team are largely based at the Institute of Employment Studies with support from Dr Susie Bamford, independent consultant with the cost evaluation. Dr Anneka Dawson and Dr Helen Gray were co-principal investigators and Dr Clare Huxley was the project manager. Kate Alexander worked across the implementation and process evaluation and Dr Charlotte Edney worked across the impact evaluation.

Contact details for correspondence
Anneka Dawson
Institute for Employment Studies,
City Gate,
185 Dyke Road,
Brighton,
BN3 1TL

Changes to the report

Update March 2024: This update includes pre-intervention scores on the primary outcome measure for pupils in schools randomised to the intervention and control groups for each of the subgroups considered in the analysis. It also includes pre- and post-intervention scores on the primary outcome measure for those who took part in testing at both time points for each subgroup.

Acknowledgements

The authors are indebted to the staff at all of the participating schools for their time taking part in the intervention and evaluation tests, interviews, and surveys. Schools have been through so much over the last two years and the staff's dedication to the children and continued enthusiasm and commitment to the project despite all the obstacles thrown at them by Covid-19 is admirable. It has been very much a team effort—with the delivery team—to navigate the challenges of the pandemic and a huge thank you to Professor Rob Savage and Amy Fox for collaborating so well at each step of the journey. Thanks also to Diotima Rapp, Celeste Cheung, Maria Pomoni, Andrea Mills, and Sarah Tillotson at the EEF for their guidance and continued support. This project has benefitted from the expertise of Dafni Papoutsaki for contributing to the original design and early impact work (including the randomisation); we are also grateful to Georgie Akehurst at IES for carrying out observations and interviews for the project, Jade Talbot at IES for arranging interviews, Seemanti Ghosh and Ben Brindle for additional data analysis, and Mandi Ramshaw and Rebecca Duffy at IES for formatting the report. Finally, thank you to the Education Endowment Foundation for funding this research and for supporting the changes to the planned evaluation and delivery throughout the pandemic.

This paper uses data from the Flexible Phonics evaluation dataset matched to the Department for Education (DfE) pupil-level spring census and the DfE school-level census for the 2020/2021 academic year. The analysis and interpretation of these datasets is the responsibility of the authors alone. The analysis was carried out in the Secure Research Service, part of the Office for National Statistics.

This output request has been granted publication-level clearance (as of 9 January 2024).

Executive summary

The project

The Flexible Phonics intervention aims to help reception class teachers and teaching assistants (TAs) deliver new strategies designed to optimise the teaching of reading to all reception pupils (aged four to five years). The intervention fits around existing phonics programmes, with strategies being incorporated approximately three to four times a week or even daily. Flexible Phonics approaches teach children to add another step after they have blended phonemes, to recognise whether they have successfully identified a word or if they need to use alternate strategies to do so. This 'set-for-variability' approach could enable children to read unfamiliar exception words independently (words that break phonic rules, such as 'the', 'two', or 'above').

The programme developers, Professor Savage and Amy Fox were based at University College London at the time of the trial. The project duration was from September 2020 to July 2021, with intervention delivery from January 2021 to July 2021. During delivery, between January and March 2021 schools were closed to most pupils due to the Covid-19 pandemic. The three half-days of initial training for teachers and TAs took place online due to restrictions. The three follow-up sessions offered to each school were also delivered remotely rather than using in-person school visits as intended. Schools received books to implement intervention strategies with pupils and they could access an online resource bank and ongoing virtual support—by email and over Zoom. The intervention delivery time was reduced from 20 weeks to 14 weeks due to the pandemic. When classes resumed, the pandemic continued to disrupt programme delivery.

Flexible Phonics was evaluated using a randomised controlled efficacy trial looking at the impact of the programme on children's word reading. A total of 123 schools were randomly allocated to the intervention or to continue their 'business as usual' provision. The initial pupil sample after the randomisation was 3,166. The process evaluation included training observations, case study interviews, online surveys, and interviews, which were remote rather than face to face due to the Covid-19 lockdowns.

Key conclusions	
1.	Pupils who participated in Flexible Phonics made the equivalent of one month less progress, on average, in early word recognition than pupils who did not receive the programme. This result has a moderate to high security rating.
2.	Pupils who participated in Flexible Phonics made the equivalent of zero months' progress, on average, in reading comprehension and correcting deliberately mispronounced words than children in other schools.
3.	Exploratory subgroup analyses found pupils who were eligible for free school meals who participated in Flexible Phonics made the equivalent of no months' additional progress in word recognition compared to similar children who did not receive the programme. There was marginal evidence that in Flexible Phonics schools that also received the Nuffield Early Language Intervention (NELI), pupils made more progress in word recognition than in Flexible Phonics schools that did not register for NELI.
4.	Teachers and TAs in Flexible Phonics schools reported that it was relatively straightforward to integrate the programme into existing phonics practice. However, a minority of educators were unclear about which elements of the programme were compulsory to deliver, so future delivery could seek to emphasise these aspects.
5.	Around 100 teachers and TAs surveyed in Flexible Phonics schools suggested that there was no change in their confidence or overall practice regarding phonics teaching, although confidence was already high at the start of delivery. They indicated that children engaged in activities well and approached reading with confidence and increased resilience.

EEF security rating

These findings have a moderate to high security rating. This was a well-designed two-arm cluster randomised controlled efficacy trial, which tested whether the intervention worked under developer-led conditions in 123 schools. The security or interpretation of the findings is impacted by the number of pupils who were not assessed at the end of the invention (20%) and the high insistence of take-up of a concurrent intervention (NELI) within the sample of schools.

Additional findings

Pupils in Flexible Phonics schools made, on average, one month less progress in the primary outcome of early word recognition than those in the control group equivalent. This is our best estimate of impact, which has a moderate to high security rating. As with any study, there is always some uncertainty around the result: the possible range of impacts

found for this programme include negative effects of three months less progress and positive effects of up to two months of additional progress.

Pupils who participated in Flexible Phonics made the equivalent of zero months' progress in the secondary outcomes of reading comprehension and correcting deliberately mispronounced words than children in other schools. The possible range of impacts found for these secondary outcomes include negative effects of 3 months less progress and positive effects of up to 2 months of additional progress. The results of this evaluation do not align with the proposed theory of change. The impact analysis shows that there was no clear positive effect of the intervention on the primary early word recognition and secondary outcomes of reading comprehension and mispronunciation correction. The reasons for the absence of any positive impact on both primary and secondary outcomes is unclear, but this result held regardless of whether prior attainment was considered. In the survey, there was little evidence of improvement related to the short-term outcomes focused on improving teacher and TA confidence in teaching phonics, however, confidence was already high at baseline. Positive qualitative data from some teachers and TAs indicated that children were more willing to attempt reading new words, try different approaches, and less concerned about getting the word 'wrong'.

Exploratory subgroup analysis found pupils in Flexible Phonics schools who were eligible for free school meals (FSM) made, on average, the equivalent of zero months' progress in early word recognition compared to FSM pupils in other schools. However, additional exploratory analysis showed that Flexible Phonics was more effective on children's word recognition in schools that also registered to receive the Nuffield Early Language Intervention (NELI). Previous research has found improved vocabulary and letter-sound knowledge among children receiving the NELI intervention, which would support both general reading and the use of Flexible Phonics strategies. However, it is also possible that schools receiving NELI may have had greater motivation and resources for language and literacy support or another unobserved differences.

The evaluation does not support existing evidence of a Flexible Phonics study in Canada (Savage et al., 2018), where there was evidence of improved reading skills and positive outcomes for spelling and reading in struggling readers aged five to seven years. It is possible that the age difference affected the level of impact in this evaluation, despite phonics teaching beginning earlier in England with pupils aged four to five who may have been at a similar developmental phase of reading to pupils in the earlier study. It is worth noting that 'set for variability' is an advanced strategy that is taught after children have learned the foundational elements of phonics. It is possible that this could have impacted the findings of this research project as the majority of pupils were identified by teachers and TAs as not meeting age-related expectations for language and communication development. This could have been due to disruption to education because of the Covid-19 pandemic. This disruption also shortened the programme delivery period and it is possible that teachers and TAs may have been focused on returning to normal practice, which may have further impacted findings.

Cost

In this evaluation there was no cost to schools: the programme was provided free of charge and online training meant that no travel or cover costs were incurred. Cost estimates for a more typical year suggest the programme would cost schools, on average, £19.58 per pupil, over three years.

Impact

Table 1: Summary of impact on primary outcome(s)

Outcome/ Group	Effect size (95% credibility interval)	Estimated months' progress	EEF security rating	No. of pupils	EEF cost rating
Early Word Recognition raw score	-0.05 (-0.2, 0.1)	-1		2,539	£ £ £ £
FSM subgroup, Early Word Recognition raw score	0.02 (-0.21, 0.25)	0	N/A	436	£ £ £ £

Introduction

Background

Studies show that systematic phonics is effective in supporting younger readers to master the basics of reading (Torgerson et al., 2018; Camilli et al., 2008; Galuschka et al., 2014). The Flexible Phonics approach is an optimisation of phonics by linking phonics to lexical and semantic information ('direct mapping') and strategies to allow independence in reading of the 'deep' (irregular) orthography of English, which admits many exceptions to phonic rules ('set for variability').[1] This approach aims to reinforce phonics learning through reading words in a meaningful context, such as reading specific recommended children's books as well as teaching children strategies to help with learning new, irregular words.

Evidence suggests that combining phonics teaching and book reading is more effective than teaching phonics alone (Hatcher et al., 1994, 2004, 2006). Further studies have found that explicitly linking phonics learning with a relevant reading task—an approach used in the Flexible Phonics intervention—was more effective than regular phonics teaching or a vocabulary learning task. Shapiro and Solity (2008) taught phonics to children aged six to seven and then explicitly linked this to reading selected children's books that contained a high density of grapheme-phoneme mappings that had been taught as part of the phonics. This approach improved reading outcomes over regular phonics teaching. Further research by Savage (2019), Savage et al., (2020), and Yeung and Savage (2020) has shown that, ideally, this linking to texts should be within the immediate future, such as in the same lesson, if possible and based on the simplicity principle that allows the greatest number of reading from the grapheme-phoneme correspondences (GPCs) they have just learnt. In the direct mapping condition of these studies, children articulated grapheme-phoneme mapping that they had recently learnt as part of the shared reading of children's books.

The teaching of phonic strategies has been shown to positively affect reading outcomes (Savage et al., 2007) and several studies have shown that focusing on variable vowel pronunciation positively impacts learning (Lovett et al., 2014; Savage and Stuart, 2001, 2006). Tunmer and Chapman (2012) demonstrated that phonics decoding skills can be broken down into component subskills. Schools in England are encouraged by the DfE to adopt a complete systematic synthetic phonics (SSP) programme (DfE, 2022a). In current best practice synthetic phonics programmes, children are first taught the speech sounds (phonemes) for letters (graphemes) and then they are taught to blend speech sounds to read full words, for example, 'c'-'a'-'t' to read 'cat' and 'c'-'a'-'tch' to read 'catch'. However, some models of word-reading propose an additional step after blending where learners compare the blended sounds with words known to them in their mental lexicon—their existing oral vocabulary. In particular, additional processing applied in cases where there is variation in the pronunciation of vowels (in irregular words) has been identified in models as either 'set for diversity' (Gibson, 1965) or 'set for variability' (Venezky, 1999).

Recent studies have found that phonics approaches that explicitly teach strategies for set for variability are more effective than standard phonics testing. An experimental study by Steacy et al. (2016) found that phonics teaching that included a focus on variable pronunciations of vowels—set for variability—yielded better reading outcomes (specifically on the pronunciation of words with variable vowels) compared with phonics teaching that did not incorporate this aspect. Several studies have found that teaching set for variability as a strategy for correcting irregular words that have been incorrectly pronounced with a regularised pronunciation improved children's ability to self-correct when they attempt to read new irregular words (Dyson at al., 2017; Zipke, 2016). Furthermore, several studies have proposed that this additional processing step can be applied to all words, including words with regular pronunciation (Elbro et al., 2012; Elbro and de Jong, 2017; Kearns at al., 2016), which suggests that teaching set for variability strategies may help with reading of all words and that there may be longitudinal effects seen on development (Steacy et al., 2019). Finally, a study using set for variability in remote teaching in Australia found that children who were struggling with reading and exposed to set for variability techniques as part of targeted lessons showed greater improvements in reading than when they were not receiving the intervention (Kohnen, Banales, and MacArthur, 2020).

[1] The terms 'direct mapping' and 'set for variability' are described more fully on page 10.

A randomised controlled trial (RCT) study in Canada found that an intervention combining these two strategies (direct mapping and set for variability) was more effective than best practice phonics teaching when taught to struggling readers aged five to seven, with additional positive outcomes for spelling and reading (Savage et al., 2018). The current study will use a RCT to test whether a similar intervention incorporating direct mapping and set for variability approaches would lead to improved reading outcomes for children of all abilities compared with current best practice phonics teaching in England. Further to this, as the intervention in Canada was taught by research assistants, the current study also investigated the feasibility of reception teachers and teaching assistants (TAs) delivering this intervention as part of everyday teaching. As phonics teaching in reception is mixed within schools (and can be led by teachers or TAs in different sized groups) a cluster RCT was the most appropriate design so that contamination across teachers or groups did not take place. The Flexible Phonics intervention aims to build on current best practice by training reception teachers and TAs to apply new approaches within phonics teaching (direct mapping and set for variability), which can help children with reading new irregular words.

As well as potential benefits to children's reading and to current phonics practice in the U.K., this study makes an especially valuable contribution to the evidence base that the EEF is developing. At the time of commissioning this evaluation, the EEF had funded ten phonics projects but none had focused specifically on reception class learners. The Flexible Phonics evaluation, therefore, fills a gap in the EEF's phonics portfolio. Further to this, the Flexible Phonics study contributes to a stated priority of the Early Years Professional Development round, which was to improve the training of reception teachers.

The IES evaluation of Flexible Phonics used a two-arm efficacy level randomised controlled trial at school level as it would not have been feasible for two different phonics programmes to be taught by the same teachers and TAs within the same school and because the intervention includes *all* teachers and TAs in the training. It would not have been possible to randomise at class level as some schools streamed for phonics across the reception year and taught groups of children from different classes with similar levels of phonics knowledge. Over half of the schools included in the final analysis (72) streamed phonics either within their class or across the year. Even where schools were teaching phonics to class groups there would still have been risk of contamination: staff who teach phonics may work with more than one class, staff may cover each other's classes if the need arises, and staff in control classes may still have noticed strategies that their colleagues were using with children in day to day interactions and activities. The control condition continued 'business as usual' phonics practice. The implementation and process evaluation included observations of the training, case study interviews, and pre- and post-intervention surveys with school staff, interviews with the delivery team, and theory of change (ToC) development, all of which provide a broad and detailed account of the intervention and how it has been received.

Intervention

The Flexible Phonics intervention helps reception teachers and TAs in the classroom delivery of new strategies designed to optimise the teaching of reading to all children. The work fits well around existing phonics programmes that can be delivered broadly as usual. A novel aspect of Flexible Phonics is that it teaches children additional strategies so they can be flexible in their approach to reading all words by trying out alternate approaches if they are unable to identify a word in their initial attempt. This could be particularly powerful in enabling children to independently read novel exception words (words that break phonic rules, such as 'the', 'two', 'between', 'above'). Children learn how to use phonics in close conjunction with authentic children's texts to become confident, motivated, readers. The theory of change models for the intervention at the start and the end of the project are shown in Appendix D and Figure 1, respectively.

The TIDieR framework for the intervention is as follows.

Name

Flexible Phonics.

Why

Systematic phonics now has a lot of evidence but there is still value in exploring whether it can be more effective in supporting children as they learn to read. Recent evidence suggests combining direct mapping and set for variability strategies can help to do this.

Who (recipients)

All pupils in reception year were targeted (aged four to five) but there may be added benefits for low achieving pupils. All reception class teachers and TAs are the direct recipients of the training and then deliver the intervention to their reception pupils in lesson time.

What (materials)

Those who were allocated to the intervention condition received three half-days of professional development training. Remote training using video-conferencing software, such as Zoom, was used in this version of the intervention during the pandemic. Intervention participants also received a copy of a teacher manual and access to the UCLeXtend online platform. The UCLeXTend platform included a discussion forum, videos of training activities, audio files for teaching activities, the training manual pdf, FAQs, training slides, Mentimeter feedback responses from the training sessions, and also teaching materials developed and shared by other schools in the trial. Following the training, there were three follow-up visits with research assistants (known as support partners). These were also delivered using video-conferencing software during the 2020/2021 academic year but would normally have been in-person, so that support partners could observe the classroom context and provide further feedback and guidance around delivering the intervention. Participating schools also received free children's books to the value of £400 per school, which could be used to implement the strategies. The delivery team selected books from existing commercially available children's literature that they felt were high quality, that is, age appropriate, well-written, with engaging stories and appealing illustrations. This included popular classics such as *The Gruffalo*. Ongoing telephone and email support was provided as needed on schools' request.

What (procedures)

The training introduced the two strategies for the teachers and TAs to implement in their teaching as follows.

1. The first strategy, direct mapping (DM), requires children to read texts that include several examples of the grapheme-phoneme correspondences (GPCs) that they have just learned. In the first instance, these will be carefully selected pre-existing decodable texts or specifically crafted controlled texts before real books are introduced slowly and strategically. While many models of phonics teaching link phonics and texts, DM aims to do so more thoroughly, consistently, and on the same day as children learn the specific GPCs, aiming to ensure that children understand phonics in context.

2. The second strategy, set for variability (SfV), explicitly teaches pupils to add in another step after they have blended phonemes to graphemes where pupils 'set for variability'. This is a metacognitive step where pupils recognise that they have not been able to successfully identify a word by blending phonemes and that they will need to use alternate strategies to identify the word. In SfV, pupils consider what the word may be by thinking about the distance between these blended sounds and known words, and potential spelling to sound inconsistencies. For example, when they sound out the phonemes 'c'-'a'-'t', the sounds they make bear little resemblance to the actual word 'cat'. SfV encourages pupils to take a moment to consider what the word may be from the words that they know. This enables children to better recognise all words but can also be especially useful when learning to recognise exception words (for example, 'wasp'). In comparison with other phonics programmes, SfV makes this metacognitive step following the blending of phonemes much more explicit and can enable children to be more flexible when selecting strategies to decode difficult words.

The three online support appointments for reception teachers and TAs ran in February and March, March and April, and April and June 2021 in this project and would be expected to run in a similar way in future. In this approach, the first appointment is offered to each class in a school (so four appointments for a four form entry school); staff are offered group appointments, if preferred (which some schools preferred in this project). The appointments are approximately 30-minute sessions held as twilight sessions from 3.30 to 6.00 pm. They enable staff to ask questions, get advice on best practice implementation, and for the Flexible Phonics support partners to deal with any misconceptions about the programme and provide further clarification. For example, some schools needed reassurance that Flexible Phonics was compatible with the government requirement to use a systematic synthetic phonics programme, and further clarification that it was meant to be used alongside their existing phonics programme and not intended as a replacement.

The online platform, UCLeXtend, was also available for the duration of the intervention, with resources including videos of the training sessions, short videos of key lessons, audio files for some of the teaching activities, training manual, FAQs, and slides, as well as any other training documents. The platform also included a discussion board for all trained teachers and TAs to join through which they could also ask for ad hoc additional support. Best practice and resources provided by partner schools were shared on schools' behalf by the Flexible Phonics team through this medium, or schools could upload and share resources directly themselves.

Teachers could also upload videos of their own practice for feedback through video calls with University College London (UCL) staff, if they chose, for specific further feedback. A monthly email bulletin also provided updates from UCLeXtend, including resources shared from other schools, the highlighting of any relevant articles on topics of concern for schools identified in the support appointments and training—such as working with children with English as an additional language (EAL) or children with special educational needs (SEND)—and sharing answers to frequently asked questions raised during the training or in online appointments more widely.

Proactive support for schools was provided by the Flexible Phonics support team by email between online support appointments. Schools could also contact their allocated Flexible Phonics support partner by phone or email, as needed.

Who (provider)

Professor Savage and his team at UCL Institute of Education (IOE)[2] delivered the training and follow-up sessions to the teachers and TAs who delivered the strategies within their normal phonics practice (both in whole-class and small-group delivery) after children had learned grapheme to phoneme correspondence. A phoneme is the smallest spoken unit of sound (for example, the word 'rain' has 3 phonemes; 'r'-'ai'-'n'). A grapheme is the written symbol that represents that sound, which can be a single letter or a sequence of letters.

How (format)

The strategies were delivered in normal phonics lessons.

Where (location)

The schools in this project were recruited from greater London.

When and how much (dosage)

The original intention was for the intervention to be delivered over five months from January 2021 to the end of May 2021. However, in this project, delivery was approximately three and a half months due to Covid-19 (see Changes section below). The expectation of the intervention is that all phonics lessons will incorporate the strategies after training until the end of the school year, which is normally three to four times a week depending on the school.

Adaptation

Teachers tailored and differentiated the content to suit children. There was freedom for teachers to adapt and modify as they go (although there was a defined core that they were required to follow).

[2] Professor Savage moved to York University in Canada in summer 2021 after delivery was completed.

Control condition

The control condition was business as usual phonics practice and schools allocated to the control condition received £1,000 at the end of the academic year when post-testing was complete.

Changes to the intervention due to Covid-19

Schools in England were partially closed from 5 January to 8 March 2021 because of the pandemic and only delivered in-person teaching to the children of key workers or vulnerable children during this time. In response to this, the intervention delivery time was adapted to run until mid-June. While some schools tried out some Flexible Phonics activities with children while teaching remotely in January and February, schools did not start teaching Flexible Phonics until 8 March when face to face teaching resumed with whole classes again. However, some schools used this time to plan their delivery of Flexible Phonics once schools reopened fully. The delivery team continued to offer support via support partners, the project manager, and Professor Rob Savage until mid-June. Overall, this reduced the delivery time by approximately one and a half months but the delivery team felt that alongside the remote teaching and planning undertaken in schools, this should still have been a sufficient time period for the intervention to elicit an effect based on previous studies of the intervention with children in other countries.[3] Schools continued to deliver Flexible Phonics until the end of the school year. As mentioned above, delivery of the training and follow-up visits were also moved to remote delivery. A summary of Covid-19 related changes to the project are presented in **Error! Reference source not found.**.

[3] The Savage et al. (2018) SSR study was run for ten hours contact time per child over 10 to 11 weeks in small groups with at-risk learners. Effect size on comparable isolated word reading outcome there was 0.41.

Table 2: Summary of Covid-19 changes to the project

Area	Planned activity	What happened	Delivery/evaluation
Training for teachers/TAs	Training in-person over 3 sessions	Virtual training over 3 sessions	Delivery
Delivery of the intervention in schools.	Approximately 5 months delivery.	Approximately 3.5 months delivery.	Delivery
Support partners observing practice in schools.	Support partners would visit schools 3 times.	Support partners had 3 visits virtually with schools and were unable to do observations of practice so had to rely on school reports of their delivery.	Delivery
Intervention Delivery and Evaluation Analysis (IDEA) workshops.	Two in-person IDEA workshops.	One meeting was in-person (in 2019) and then there were two follow-up IDEA meetings with delivery team. The second workshop included support partners too.	Delivery/evaluation
Observations and case studies.	In-person observations including 3 training sessions and 2 support partner visits. Case studies, which included observing teaching. All to be carried out in-person.	Virtual attendance at all the training was possible to carry out observations. No in-school observations were possible for support partner visits or case studies. Case study interviews were carried out virtually/ by telephone.	Evaluation
RQ4. 'Does the Flexible Phonics intervention provide value-added improvement to reception children's word reading ability compared to good phonics teaching alone in schools identified with good phonics practice?'	Research question was due to be answered by collecting data from schools on their historic phonics performance.	Research question was removed as data could not be collected from schools and public data was not appropriate.	Evaluation
Pre-testing.	All schools tested in-person.	17 schools had remote testing due to school restrictions.	Evaluation
Nuffield Early Language Intervention (NELI).	NELI would have been delivered to a small number of schools.	NELI was delivered to 42% of the schools as part of the government's Covid-19 support strategy. This was therefore explored in the analysis.	Delivery/evaluation

With respect to the evaluation of the project, the pandemic also led to several changes. Partial school closures and limitations on external school visitors meant that in-person observations and case studies for the evaluation were not

possible and all took place by phone or video conferencing instead. The second IDEA workshop[4] was moved to April 2021 (from February 2021) after schools had fully opened again and there was more certainty over the future of the project. This IDEA workshop took place by video conferencing and a follow-up ToC discussion also took place in June 2021 at the end of delivery. Other important changes to the evaluation include the removal of one of the research questions, RQ4, 'Does the Flexible Phonics intervention provide value-added improvement to reception children's word reading ability compared to good phonics teaching alone in schools identified with good phonics practice?', as historic phonics screening check information (which was originally going to be used to determine schools with good phonics practice) was not collected from schools because we did not want to place extra burden on them during the pandemic. An alternative option of using the publicly available data on progress between Key Stage 1 and Key Stage 2 to identify good phonics practice was also rejected due to the high proportion of schools showing 'average' progress.[5]

In addition, some of the pre-testing was completed remotely in 17 schools as there were still restrictions on external visitors coming to schools during the pandemic in June and July 2021. This was piloted with a couple of children from schools not involved in the project by Qa Research, an external assessment organisation that worked with the evaluation team to conduct the pre- and post-assessments with children for the trial. For the pilot exercise, Qa staff went through the test remotely with two children of the same age to confirm this was a viable means of administration. Test administrators used the same procedure for the tests, but used a video-conferencing programme, such as Zoom, to speak to the children and show them the pictures. A TA was present with the children at all times but sat behind the child and was given instructions by the assessor not to interfere with the testing process or prompt the children. Assessors were briefed on strategies for responding appropriately if a staff member was interfering in the test and the test administrators checked in with each assessor after testing to monitor for any interference from school staff. If there had been problems with this method, then the test administrators would have followed up with the school before any subsequent remote testing days. However, there were no issues with staff interference at pre-test.

Finally, after schools in the trial had been randomised into the intervention and control conditions, the evaluation team became aware that some schools participating in the Flexible Phonics trial had also signed up for a language and early literacy skills programme, the Nuffield Early Language Intervention (NELI), which was rolled out as part of the government's COVID-19 support strategy. NELI was designed to improve the spoken language skills of reception-age children such as those who had fallen behind during the pandemic. Priority was given to schools with a high proportion of disadvantaged pupils and pupils who needed additional support took part in small group and one to one sessions. The EEF has provided a list of all the Flexible Phonics schools that have signed up. Forty-two percent of schools in the trial had also taken up NELI (50 of 120 settings)—39% of control schools (24 of 61) and 44% of treatment schools (26 of 59). While the distribution of schools participating in the NELI is relatively balanced across the groups, the fact that around two-fifths of schools participated in NELI reduced the ability to estimate the impact of Flexible Phonics independently of NELI.

Theory of change

After the initial IDEA workshop in October 2019 with the delivery team (UCL) and the evaluators (IES), an initial model was developed to demonstrate the theory of change (ToC) underlying the Flexible Phonics intervention. The initial theory of change model shown in Appendix D described the rationale for the intervention, the overall ToC, inputs, activities, enabling factors, expected outputs, short term outcomes and mediators, and the expected long term outcomes and impacts.

A series of revised ToC models was created to incorporate changes made to delivery as a response to the ongoing Covid-19 pandemic as well as an increase in the sample size achieved at baseline data collection. A final model was

[4] Intervention Delivery and Evaluation Analysis (IDEA) workshops are a required element of EEF evaluations where the evaluation and delivery team meet to develop or update a Theory of Change logic model that sets out the elements of the programme and how they are expected to work, to identify which components are required for compliance to the intervention, and to identify relevant measures for the programme outcomes and impacts.

[5] Available at: https://www.compare-school-performance.service.gov.uk/. We rejected the option of oversampling schools that do not have good phonics teaching to enhance the prospects of being able to explore differential impacts because of the risks this posed to recruiting sufficient schools for the trial.

created after the ToC discussion in June 2021, which reflects the model as it was delivered for this project (Figure 1). The revisions to the model included:

- A note was added to state the number of participating schools recruited at the time of randomisation (123), which was higher than originally anticipated as recruitment was so successful.

- The description of the two follow-up sessions changed from in-person visits at schools to three online appointments.

- The three follow-up sessions no longer included observation of a phonics lesson (as they were no longer in-person). The sessions were an opportunity for schools to ask questions and receive advice on implementation and the delivery of the intervention.

- 'Ongoing support' had not been defined in the initial model 'Activities' section. This was updated to include proactive support for schools that was provided by the Flexible Phonics support team by email between February and June, where relevant resources and best practice were shared proactively with their schools in between the follow-up sessions. During this time, schools could contact their allocated support partner for any ad hoc support as required.

- Initially 'resources' had not been defined in the 'Activities' section as they had not been developed. The 'Support for schools' section now includes an online platform of resources (UCLeXtend), which was available for all trained staff and included a discussion forum as well as a monthly email bulletin sent to schools and described in the TIDieR framework section of the report.

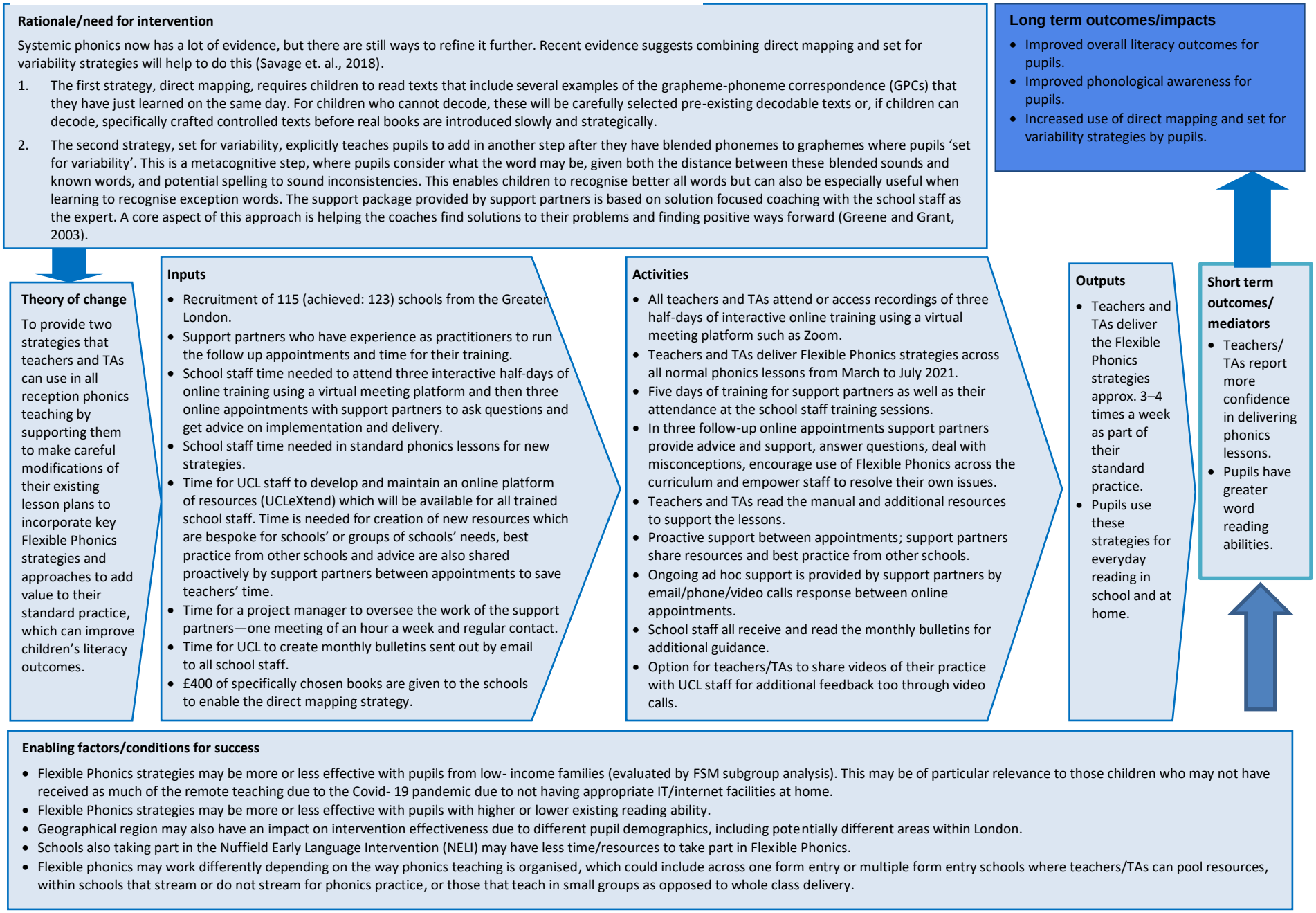
- The option for teachers and TAs to share videos of their practice with UCL staff for additional feedback through video calls was also added to the 'Activities' section.

- A note has been added to the 'Enabling factors/conditions for success' section regarding the possible implications of lack of digital access for low-income families, which was particularly important during school partial closures during this project, and, in addition, the possible impact of schools also taking part in the Nuffield Early Language Intervention (NELI).

In addition to the changes, a further two-part optional online workshop about Flexible Phonics ran in July 2021 for Year 1 teachers over two twilight sessions from 3.30 to 5.00 pm. This was added at a late stage (end of April 2021); it is not known if this will be used in any future roll-out of this intervention so, as agreed with the delivery team, this has not been included in the theory of change model at present.

The delivery team felt that in future the training and follow-up visits would most likely take place in person as originally planned (perhaps with some element of a hybrid model with some sessions in person and some online) but that catch-up sessions for those who missed a session could be delivered remotely (see Future Scale-up section in Implementation and Process Evaluation).

Figure 1: Theory of Change



Evaluation objectives

The primary research question is:

- RQ1. Does the Flexible Phonics intervention improve reception children's word reading ability (measured by the York Assessment for Reading Comprehension (YARC) Early Word Recognition subscale)?

The secondary research questions are:

- RQ2. Does the Flexible Phonics intervention improve reception children's literacy outcomes (measured by general literacy test)?

- RQ3. What is the differential impact of direct mapping and set for variability skills on children's word reading ability?

- RQ4. Does the Flexible Phonics intervention provide value-added improvement to reception children's word reading ability compared to good phonics teaching alone in schools identified with good phonics practice?

  Update: this question was removed. It was not possible to answer this research question as Year 1 Phonics Screening data was not collected from schools during recruitment as explained in the section above.

- RQ5. Does the Flexible Phonics intervention improve word reading ability differentially for children eligible for free school meals?

- RQ6. Does the Flexible Phonics intervention improve word reading ability differentially for children who had low scores at pre-test?

- RQ7. Does the Flexible Phonics intervention improve reception children's phonics skills one year later at the end of Year 1?

  Update: this question was removed. It was originally due to be considered in an addendum report in 2023, but a decision was taken not to proceed with this analysis, for reasons which are explained in the Conclusions chapter.

- RQ8. Does the impact of the Flexible Phonics intervention differ depending on whether the school was participating in the Nuffield Early Language Intervention (NELI)?

RQ8 was added to the list of formal research questions when it became apparent that a large proportion of schools were participating in NELI as part of the government's Covid-19 support strategy.

The original protocol from August 2020 is available here:
https://d2tic4wvo1iusb.cloudfront.net/documents/projects/EEF_trial_protocol_flexible_phonics_v2.pdf?v=1630925243

This was updated in June 2021 following the school partial closures in early 2021 caused by the pandemic and is available here:

The statistical analysis plan is available here:
https://d2tic4wvo1iusb.cloudfront.net/documents/pages/projects/20210609_EEF_SAP_Flexible__final.pdf?v=1630925245

Implementation and process evaluation research questions

The IPE assessed the eight key implementation dimensions set out below and identified moderating and contextual factors that influence impact and explain quantitative findings.

Fidelity

- IPE1. Are schools delivering the interventions and the trial as intended?[6]
- IPE2. Could the intervention be rolled out on a larger scale so that the intervention is delivered as intended?
- IPE3. What adaptations would be required to roll out the intervention on a larger scale and how might these affect the integrity of how the intervention is delivered?

Dosage

- IPE4. Do teachers and/or TAs teaching reception receive all intended training?
- IPE5. How often do participating teachers and TAs deliver Flexible Phonics strategies in phonics teaching?

Quality

- IPE6. How well is initial training and follow-up support received by teachers, TAs, and senior leadership at the school?
- IPE7. Is it necessary to conduct cascading training and has this been monitored or supported?
- IPE8. How effectively do teachers and TAs use Flexible Phonics strategies?
- IPE9. What facilitates or hinders effective implementation?
- IPE10. Would teachers and TAs find additional support helpful in maintaining quality—what and from whom?
- IPE11. Are there unintended or negative effects of the intervention?
- IPE12. What are TAs' and teachers' perceived benefits and outcomes of the intervention?

Reach

- IPE13. Do all intended pupils receive Flexible Phonics teaching?
- IPE14. Do some pupils receive more Flexible Phonics teaching than others?

Responsiveness

- IPE15. Do teachers and TAs engage well?
- IPE16. Is the intervention acceptable and practicable in schools' contexts?
- IPE17. Do senior managers perceive the intervention as worthwhile and cost-effective?

Programme differentiation

- IPE18. How does the intervention enhance, or differ from, existing phonics teaching?

[6] The delivery team rated compliance on a series of five key features of programme delivery, which were observed by support partners during the three follow-up support sessions. These key areas were agreed with the delivery team during the IDEA workshops and included direct mapping, vocabulary, oral flexibility, print-based flexibility in mispronunciation, correction of phoneme strings, and continuous phonation. Schools did not have to be compliant in all areas as some may not be relevant, depending on what children in a class need support with.

- IPE19. Before Flexible Phonics implementation, what was business as usual and how was this embedded in wider approaches to reading?

Control group monitoring

- IPE20. What phonics teaching and wider reading strategies are used in control schools?
- IPE21. Does the behaviour of control schools change during the trial?

Adaptation

- IPE22. Have schools adapted the intervention—how and why?

Ethics and trial registration

IES applied for ethics approval through the internal IES Ethics Committee once the recruitment materials were finalised in January 2020, which included a review by two senior staff within IES and the chair of the ethics board who is a member of Institute Management Team. Headteachers, senior leaders, or EYFS leads firstly received a flyer on the project and, if they were interested, received an information pack of details of the project (Appendix E) and then signed a Memorandum of Understanding (MOU) to agree to take part in the project, detailing the responsibilities of all parties for the trial and data sharing (Appendix F). A project lead was designated in the MOU; this person then received further communications on the project as it progressed. Schools were also invited to webinars in July 2020 with the UCL, IES, and Qa teams to hear more about the project and ask questions.

Parents received an information sheet detailing the trial and data sharing (Appendix G) and had the chance to withdraw their child's data from being shared with the evaluation team at the start of the trial by telling the school. An accompanying letter contained a link to a privacy notice that explained how the data was being used, stored, shared, and deleted (Appendix H). After randomisation, the parents could contact IES directly to have their child's data removed from the data stored by the evaluation team.

Teachers and TAs had the chance to opt out of taking part in the IPE when they received the invitation to complete the survey or take part in an interview which contained information on what the research would contain and data sharing.

The delivery team applied for ethics approval through UCL IOE for delivery of the trial, which was approved separately.

The study is registered with the ISRCTN registry with the study ID ISRCTN18428598 and can be accessed at: http://www.isrctn.com/ISRCTN18428598

Data protection

All participants interviewed verbally consented to indicate that they understood the research aims, agreed to the interview being recorded and transcribed, and were given assurance of anonymity. As mentioned, schools signed a MOU identifying the requirements of the project and how the data was used, shared, and stored (see Appendix F).

There was also a privacy notice for school staff (Appendix I) explaining how information collected was used and stored, and to communicate to participants their right to withdraw from data processing. This was available online, with the link provided in school and teacher letters, and included in email briefings to take part in the surveys and interviews. We also developed a data-sharing agreement between IES, UCL, and the EEF stating data to be shared by whom, how, and why to ensure full data security.

Our evaluation approach involved personal data collection including pupil name, date of birth, gender, and Unique Pupil Number (UPN). IES accessed and linked this pupil data to background and school data held on the National Pupil Database (NPD) at two points in time. The first NPD extract included whether or not the pupil is eligible for free school meals, to be used in this initial round of analysis for this report. A second request was planned for a longitudinal follow-up using the results of the Year 1 Phonics Screening Check for this cohort of pupils, but it was decided in collaboration with the EEF not to continue with this aspect of the research after reviewing findings from the main impact analysis. IES

has matched the above pupil data to data on pupil outcomes collected throughout the study. This includes data from questionnaires and assessments administered as part of the project including a standard assessment of literacy skills and a measure of mispronunciation correction as well as data on outcomes available through the NPD.

IES's legal basis for processing personal data is 'legitimate interests'. The evaluation of Flexible Phonics fulfils one of IES's core business purposes (undertaking research, evaluation, and information activities) and is, therefore, in its legitimate interest: that processing personal information is necessary for the conduct of the evaluation. For the purposes of conducting the evaluation to assess the impact of Flexible Phonics, IES and UCL IOE are data controllers of personal data of school staff and pupils obtained from schools and other sources such as the NPD. Personal data was shared with trusted processors, such as test administrators and transcribers as well as members of the delivery and evaluation teams, solely for the purposes of proper delivery, management, and evaluation of the project. At the end of the project, data will be submitted to the EEF's data archive. At this point, the EEF will become a data controller and the archive manager will be the data processor.

IES will securely delete all personal data within six months of the project finishing, that is, once the final draft of the evaluation report has been submitted and the trial data has been submitted to the EEF archive. UCL IOE will keep the data for five years.

Project team

Delivery team

Professor Rob Savage, York University, Canada (formerly at UCL IOE): developer of the Flexible Phonics intervention, delivered training and catch-up training sessions for intervention school staff teaching phonics to reception class.

Amy Fox, National Literacy Trust (formerly at UCL IOE): project manager. Amy was responsible for managing all delivery activity, contributed to programme design, including training and manual development and creating research tools, liaising with IES and Qa, managing the support partners, and quality assuring support given to teachers.

Flexible Phonics support partners (research assistants) from UCL IOE delivered the follow-up and support sessions with individual intervention schools: Alice Robinson, Clare Whalley, Denise Amankwah, Greta Boldrini (UCLeXtend lead, January to July), Ameena Khan Sullivan, Sophia Gowers (UCLeXtend lead), and Sam Dexter (January to April).

Evaluation team

Dr Anneka Dawson, co- principal investigator of the evaluation. Anneka led on the implementation and process evaluation, overseeing the assessments, and quality assuring materials.

Dr Helen Gray, Learning and Work Institute (formerly at IES), co-principal investigator of the evaluation. Helen led on the impact assessment.

Dr Clare Huxley, IES, project manager. Clare was responsible for managing all research activity, liaising with UCL, drafting research tools, and coordinating members of the research team.

Dr Susie Bamford, IES. Susie supported the cost evaluation analysis.

Dr Charlotte Edney, Nuffield Family Justice Observatory (formerly at IES). Charlotte supported the impact evaluation analysis from November 2021.

Kate Alexander, IES. Kate supported the IPE.

Dr Dafni Papoutsaki, IES. Dafni originally supported the impact evaluation analysis until June 2021.

Georgie Akehurst, IES. Georgie supported the IPE until June 2021.

Methods

Trial design

Table 1: Trial design

Trial design, including number of arms		Two-arm, cluster randomised controlled efficacy trial with pupil-level outcomes
Unit of randomisation		School
Stratification variable(s) (if applicable)		None
Primary outcome	Variable	Early Word Recognition
	Measure (instrument, scale, source)	Early Word Recognition subscale raw score (0–30) from the York Assessment for Reading Comprehension (YARC)
Secondary outcome(s)	Variable(s)	Early Word Reading composite measure Mispronunciation Correction Literacy over the longer-term
	Measure(s) (instrument, scale, source)	For literacy: The sum of standardised scores derived from each of the four YARC subscales, i.e., early word recognition, letter sound knowledge, sound deletion, and sound isolation. For Mispronunciation Correction: An adapted version of Tunmer and Chapman's Mispronunciation Correction Test (2012) as used in Dyson et al. (2017) using the words most commonly used in English children's books.
Baseline for primary outcome	Variable	Early Word Recognition
	Measure (instrument, scale, source)	Early Word Recognition subscale raw score from YARC
Baseline for secondary outcome(s)	Variable	Early Word Recognition and Letter Sound Knowledge composite measure
	Measure (instrument, scale, source)	Constructed from the standardised scores for the Early Word Recognition and Letter Sound Knowledge subscales from YARC

The evaluation was an efficacy level randomised controlled trial (RCT). As the intervention involved training all reception teachers and TAs, randomisation was at the school level to two groups. Half of the participating schools were randomly allocated to the treatment group, the other half to the control group. The control schools were asked to continue their usual approaches to phonics teaching.

The study measured the impact of Flexible Phonics for pupils in reception class in the school year 2020/2021. Children participated in a pre-test of reading ability—the York Assessment for Reading Comprehension (YARC), using the early word recognition and letter sound knowledge subscales—prior to randomisation to verify that treatment and control groups are well-matched. Information on the balance between the two groups pre-intervention enabled assessment of the likely robustness of findings and accounts for some of the variance in the post-test, meaning sample size is optimised.

To reduce the costs of testing and minimise the time and administrative burden on schools, the pre-test and post-test were administered to one class per school. Where there was more than one reception class per school, the class was

selected at random by IES from a list of teachers provided by the school. While only one class per school took part in the pre- and post-tests, all teachers or TAs were invited to take part in the training and training materials could be shared with colleagues who did not attend the training but cascading was not an intended part of the programme design. Therefore, the transfer of teachers between classes should not have impacted on whether pupils received the intervention. Also, there was less movement of teaching staff between classrooms this academic year due to Covid-19.

Participant selection

The intervention was targeted at children in reception classes who were expected to turn five in the 2020/2021 academic year. All children in reception classes in the schools recruited to the trial were eligible to participate. All teachers and TAs of reception-aged children in the schools assigned to the treatment group were invited to attend training and participate in other activities to equip them to teach Flexible Phonics.

Schools participating in the English Hubs programme were not eligible to participate in the trial. All other schools with reception-aged children in Greater London were eligible as long as they were not participating in another EEF reception year trial other than the Nuffield Early Language Intervention Effectiveness trial (which was so widely spread it was impossible to avoid). The EEF selected the geographical location of the trial (Greater London) to avoid, where possible, overlap with other EEF trials evaluating interventions targeting language and communication in reception year. Greater London was also convenient for the delivery team, which was based at University College London at that time.

The delivery team (from University College London, Institute of Education) was responsible for recruitment from October 2019 to July 2020 and was expected to recruit approximately 115 to 125 schools so that approximately 100 would go forward to participate in the trial, allowing for attrition between signing up and being randomised. The delivery team was quite successful at building interest in the trial and was able to maintain a waitlist so it could replace schools that dropped out before randomisation. Twenty-six schools dropped out between recruitment and randomisation and were replaced with waitlist schools until 123 were randomised in December 2020 (therefore, 149 schools were recruited in total). Three of the schools assigned to the intervention group withdrew from the trial after randomisation and then one school each from the intervention and control groups were withdrawn as it had not been possible to assess their pupils within the post-test period. Therefore, the final sample for analysis was 118 schools. The team used tweets on Twitter, announcements in sector newsletters, and an online information session to attract schools and get initial sign-ups.

Outcome measures

Baseline measures

The York Assessment of Reading for Comprehension (YARC) test is suitable for four- to seven-year-olds and covers four dimensions: sound isolation, sound deletion, letter sound knowledge, and early word recognition. Two subscales (early word recognition and letter sound knowledge) from the YARC Early Reading test[7] were used to assess baseline performance. An overview of the four subscales is given below, including internal reliability scores.

- The **early word recognition** test measures reading attainment in young readers. Children are asked to read 30 single words that are graded in difficulty. Half of the words have regular correspondence between the graphemes and phonemes, that is, letter to sound mapping, and half are irregular. The test's internal reliability using Cronbach's alpha is 0.98. This test is a measure of overall literacy outcomes, a longer-term impact on which is expected in the ToC.

- The **letter sound knowledge** test measures alphabetic knowledge. Children are shown lower case letters and digraphs, one at a time, and are asked to say what sound the letters make. A digraph is two letters that combine to make one sound, such as 'sh'. The core test comprises 11 singleton letters and six digraphs. The extended test comprises 26 singleton letters and six digraphs. The core test's internal reliability using Cronbach's alpha is 0.95. This test is a measure of phonological awareness, a longer-term impact on which is identified in the ToC.

[7] Available from GL at https://www.gl-assessment.co.uk/products/york-assessment-of-reading-for-comprehension-yarc/

- The **sound isolation** test measures phoneme isolation skills, which are a component of phonemic awareness. Children hear a series of 12 nonsense words and are asked to identify either the first or the final sound in the word. The test's internal reliability using Cronbach's alpha is 0.88. This test is a measure of phonological awareness, a longer-term impact on which is identified in the ToC.

- The **sound deletion** test measures phoneme deletion skills, which are a component of phonemic awareness. Children hear a series of 12 words accompanied by a picture of what they represent, and they are asked to repeat the word but 'take away' a sound from the word. The test's internal reliability using Cronbach's alpha is 0.93. If the Sound Isolation and Sound Deletion scores are combined, this combined score has an internal reliability of 0.95 using Cronbach's alpha. This test is a measure of phonological awareness, a longer-term impact on which is identified in the ToC.

To reduce testing time and burden on the school, only the early word recognition (measured on a scale of 0–30) and extended letter sound knowledge (0–32) subscales were used as a pre-test. The delivery and evaluation teams agreed that these two subscales were the most relevant measures for expected outcomes of the intervention, as the early word recognition subscale reflects children's ability to read, while letter sound knowledge captures children's initial phonics knowledge of mapping letters to sounds (graphemes to phonemes) and is more appropriate than the other subscales given that the pre-test is conducted at the start of reception year when the early years curriculum first introduces phonics teaching (DfE, 2021).

These two tests were used in combination as the baseline for the secondary outcomes, while the early word recognition subscale was used as the baseline for the primary outcome. The baseline tests were administered in November to December 2020.

Primary outcome

The primary outcome was the early word recognition subscale score as the intervention theory of change model, developed in collaboration with the delivery team, identified that word reading is where the programme developer envisaged most of the impact would be seen. Qa Research carried out the pre- and post-tests as an independent test administrator, and administrators were blind to whether the school had been assigned to the treatment or control group. After data entry, the test administrators ran a series of logic checks on key information, for example, date of birth, checked field completion, and checked for anomalous values. The evaluation team also checked that the raw score on the primary outcome measure matched the number of correct responses on individual items to verify that the raw score had been calculated correctly.

It was necessary to carry out some of the testing remotely due to the Covid-19 pandemic in a total of 17 schools, as discussed in the Introduction. For remote testing, the assessment was conducted via video-conferencing with the independent assessor in one location speaking to and sharing materials on-screen with a child at their school who was accompanied by a TA or teacher. TAs and teachers were instructed not to help the child with the test questions and tasks. More detail regarding remote testing can be found in the Introduction in the section Changes to the Intervention Due to Covid-19'.

Secondary outcomes

All four subscales from the YARC (described above) were used as a post-test, with a composite measure of early word reading constructed by standardising the raw scores on each of the subscales and then combining them into a single metric.

To capture the differential impact of the direct mapping and set for variability strategies, we used an adapted version of the Mispronunciation Correction Test (MCT) developed by Tunmer and Chapman (2012) to assess the impact of set for variability. During the MCT, children are asked to play a game with a puppet that they are told sometimes says words wrong. The children are then presented with examples of irregular words that are incorrectly pronounced with a regularised pronunciation. The children are asked to 'correct' the puppet, which requires them to consider other possible pronunciations of the word they have just heard, that is, implement a set for variability strategy. The number of words that a child successfully 'corrects' is then used as a measure of their ability to use set for variability strategies. In their 2012 paper, Tunmer and Chapman report delivering the test as two sessions conducted two weeks apart (one presenting

the word in isolation and one where the words are embedded in sentences) but for this study, children were tested once with the puppet pronouncing the words embedded in sentences following the methodology used by Dyson et al. (2017).

The MCT was administered at post-test only due to the limited time to carry out testing in the period prior to the delivery of the intervention and only for a randomly-selected subsample of 15 pupils per school, or all pupils if there were only 15 in the class. This reduced the burden on schools and pupils and kept costs to a minimum, but the reduction in sample size means that it is necessary to interpret findings carefully as it is less likely that the estimate of impact will be statistically significant. We explored the means by which set for variability affects the primary outcome measure in order to isolate the impact of direct mapping and compare the differential impact of each strategy. Again, all the post-tests were administered by Qa Research and testing took place in June and July 2021.

In addition to the post-test outcome measures, it was planned that data from the Year 1 Phonics Screening check recorded on the NPD would be used to explore the impact of the intervention on longer term outcomes and the findings published as an addendum to this report, once the required data was available. However, it was decided in collaboration with the EEF not to continue with this aspect of the research after reviewing findings from the main impact analysis. The planned analysis is discussed in further detail in the section on future research in the Conclusions chapter.

Sample size

More schools participated in the trial than was expected at the time that the protocol was drafted (123 were randomised, rather than 100) but the number of pupils per class who took part in the pre-test for the primary outcome measure was slightly lower than expected (22 rather than the 23 predicted at the time of randomisation). A total of 3,166 pupils took part in the pre-test, with 1,567 of these in schools randomised to the intervention group and 1,599 in the control group. Using similar assumptions to those made when the protocol was drafted and with an ICC of 0.15, the impact of increasing the scale of the trial was to reduce the MDES from 0.23 to 0.21 standard deviations for the main sample and from 0.37 to 0.33 for FSM pupils.

As some pupils did not take part in post-intervention testing, the final sample for analysis was reduced compared with the numbers randomised to the intervention and control groups. It was also possible to observe some of the parameters which had to be assumed in the earlier calculations of the MDES. The correlation between pre- and post- test scores on the primary outcome measure was higher than expected (0.54 compared with 0.40 expected when the trial was designed). Schools also accounted for less of the variation in test scores than expected and the ICC was 0.13 rather than 0.15. The 58 schools in the intervention arm where some pupils took part in both the pre- and post-tests contributed a total of 1,256 pupils to the final analysis sample, compared with 1,283 pupils at the 60 schools in the control group. The MDES for the primary outcome was 0.17 standard deviations—below the MDES at the protocol or randomisation stages. As the MDES for the analysis sample was below that anticipated at the design stage, the achieved sample size appeared sufficient to detect the expected impact from Flexible Phonics.

The NPD measure of whether a pupil has been eligible for free school meals at any point in the past six years (EVERFSM_6_P), observed in the 2020/2021 academic year, was used to determine the number of FSM pupils in the final analysis sample. The class average was higher than expected, at four pupils per class rather than three. The correlation between pre- and post-test scores on the primary outcome measure was 0.63, which exceeded the correlation predicted earlier in the trial. The ICC was slightly lower at 0.12 compared with 0.15 assumed at the time of randomisation. With 211 FSM pupils in the intervention group and 225 in the control group, the MDES for the final analysis sample was 0.29 standard deviations. While this was lower than the MDES expected for this subgroup when the trial was being designed (0.37 standard deviations), and even below the MDES anticipated at the time of randomisation, it was still considered unlikely to be possible to detect any impact from Flexible Phonics on FSM- eligible pupils.

Randomisation

All schools recruited to the trial were asked to supply a list of reception teachers. Where the school had more than one class of reception-aged children, the evaluation team selected one teacher per school at random and children in the class taught by this teacher were selected to participate in the pre- and post-tests. The random selection of classes in multi-form entry schools took place on 9 October 2020, prior to the administration of the pre-test. It was carried out in

Stata and the seed for the random number generator recorded to ensure that the allocation could be replicated. A random number from a uniform distribution was assigned to each teacher and they were then sorted in ascending order by school. The teacher with the lowest value on the random number within each school was selected to be the focus of class testing

Following the pre-test, schools were randomly assigned to the intervention and control groups. This was done in two batches, as testing was completed. This was necessary as the Covid-19 pandemic made it more difficult to arrange testing and in some cases it was necessary to reschedule appointments. Randomising schools in batches made it easier for the delivery team to arrange training sessions for schools allocated to the intervention group before the Christmas break. The randomisation dates were six days apart: 83 on 3 December 2020 and 40 on 9 December 2020. Again, schools were allocated a number at random from a uniform distribution and then sorted in ascending order. A sort order variable was derived based on the sequence of random numbers and schools with an odd number on the sort order were assigned to the intervention group, while those with an even number were allocated to the control group. The Stata syntax used to assign schools to either trial arm is included in Appendix J.

The delivery team was supplied with a list of schools assigned to the intervention group so that teachers could be invited to training. Test administrators were blind to whether schools had been allocated to the intervention or control groups at both the pre- and post-test stages. Before the post-test was conducted it was necessary to select 15 pupils at random (in classes with more than 15 pupils) to take the MCT. A random number from a uniform distribution was assigned to all pupils included in the data extract supplied by schools at the start of the trial. Within each school, pupils were sorted in ascending order of this random number and the first 15 in each class were selected to take part in the MCT. The evaluation team were not blind to whether schools were part of the intervention or control groups, but the test administrators were.

Statistical analysis

Primary analysis

The primary outcome used to measure the impact of the intervention was the Early Word Recognition subscale raw score (0–30) from the York Assessment for Reading Comprehension. The trial was designed as a cluster randomised controlled trial, meaning that schools were assigned to either the treatment or control group. All pupils within a specific class in a certain school were either given the intervention or not given the intervention. Given this nested structure, pupil outcomes are likely to be correlated within schools. This can result in standard errors in conventional OLS regression models being underestimated and produce a significant treatment effect deeming the intervention a success where in fact it is due to the estimation method.

This evaluation, therefore, uses a multilevel modelling methodology with outcomes measured at the pupil level. Commonly in educational research, determining whether an intervention is effective is based on an assessment of the probability of rejecting the null hypothesis. This analysis takes an alternative approach by using Bayesian multilevel methodology. Bayesian multilevel modelling with weakly informative priors is likely to improve the precision of the impact estimates compared with other techniques (Xiao, Kasim and Higgins, 2016). Given this finding, the analysis for the current study was carried out using an analytical package (called EEFanalytics.ado) developed by the EEF to run within the statistical analysis software package, Stata 16. Specifically, the analysis used the `crtbayes` command in EEFanalytics.ado. The prior distribution is based on an earlier quasi-experimental study of direct mapping and set for variability in Canada (Savage et al., 2018). Bayesian credibility intervals are reported.[8]

The main analysis of the primary outcome measure used the raw pre-test score on the Early Word Recognition subscale as the measure of prior attainment.

[8] The analysis was based on a Markov chain Monte Carlo sample size of 100,000 with a burn-in period of 2,500, repeated for ten chains.

Secondary analysis

Unlike the main analysis of the primary outcome, the analysis of the secondary outcome measures was not informed by any prior expectations about the likely size of the effects. The only other difference between the analysis of primary and secondary outcome measures was in the measure of prior attainment, as explained in the section on the baseline outcome measures.

Analysis in the presence of non-compliance

Schools were regarded as compliant with the Flexible Phonics programme, if both of the following were true:

1. The school complied with the training attendance measure. This was the case where the teacher of the class that was selected for testing attended all three training sessions or watched videos of the training sessions and attended a catch-up tutorial. All schools assigned to the intervention group met this requirement and so no schools were non-compliant on the attendance measure.
2. The school delivered the intervention to the required standard. In this case, the delivery team drew on facilitated discussions with school staff to assess whether the Flexible Phonics programme was being implemented as intended.

The assessment of compliance in the delivery of Flexible Phonics was made by scoring phonics practice over three follow-up sessions with schools. Compliance was assessed on a four-item scale ranging from zero, which indicated that there was no implementation, to three, meaning that there was expert, adaptive delivery. Only those aspects of teaching practice that were relevant at each point in time were assessed during a given observation. The highest score recorded at any of the sessions for each of the areas of Flexible Phonics practice was used to determine whether the school was compliant in delivering Flexible Phonics as intended. Schools had to attain a score of two or three in the assessment of direct mapping practice. They also needed to achieve a score of two or three or 'not applicable' on assessments of either Oral Flexibility or Print-Based Flexibility.

A Complier Average Causal Effect (CACE) analysis was used to estimate the impact of Flexible Phonics on pupils in schools where Flexible Phonics was assessed as being delivered as intended. This used a two-stage least squares regression, with assignment to the treatment or control group used as an instrumental variable.[9] In the first stage, the probability of compliance was estimated using the binary compliance measure as the dependent variable. The second stage used the predictions from the first stage to estimate the impact of compliance on the primary outcome measure of Early Word Recognition, controlling for the pre-test score on Early Word Recognition (EWR). The analysis takes into account the fact that pupils are clustered together within schools when calculating standard errors.

Missing data analysis

The attrition rate of pupils in both trial arms was compared to assess the potential impact of missing data on the findings of the primary analysis. The number of missing observations were reported at school and pupil level across the pre- and post-test measures for the primary outcome. A binary variable was derived to indicate whether the post-test score on the primary outcome measure was missing and a probit regression was used to establish which baseline pupil and school characteristics predicted the likelihood of the post-test EWR raw score being missing.

As a sensitivity analysis, missing values were imputed jointly over clusters using the multivariate normal model for the primary and secondary outcomes. The Stata mi suite of commands were used for running the imputations and the model was set to generate ten imputations. Controls were included for gender, age in months, low reading score indicator, treatment indicator, pre-test score, and free school meal eligibility. However, it was not possible to use the mi commands with the Bayesian approach because they produce p-values that are not compatible with Bayesian statistics. This limitation was not identified at the time that the SAP was drafted and so it was necessary to deviate from the SAP. Instead of following the intended approach, the main analysis was estimated using an Ordinary Least Squares (OLS) regression. Missing values were then imputed as described above, and OLS re-estimated with the inclusion of these

[9] The analysis was carried out using the `ivregress` command in the Stata statistical analysis software package.

imputed values. The coefficients across the two models were then compared to assess the degree to which missing values may be driving any results found in the main analysis.

Subgroup analyses

Subgroup analysis was conducted to examine whether the effect of the intervention differed among three different groups of pupils: FSM pupils, low-ability pupils, and pupils at schools that were *not* participating in the Nuffield Early Language Intervention (NELI). Pupils with low scores were defined as those who scored less than the median on the combined pre-test standardised Letter Sound Knowledge and Early Word Recognition scales. The estimation strategy remained as described in the primary outcome analysis, but with the analysis restricted to each of the three subgroups. In addition to this, an interaction term between the treatment identifying variable and the subgroup of interest was included to identify the differential impact of the intervention on the subgroup of interest. In the analysis for the FSM subgroup, the number of chains was reduced from ten to five. This was due to having to run the analysis in the Secure Research environment, which imposed a time constraint on how long the code could be run for. The reduced number of chains made this a less time-consuming process and is highly unlikely to make a difference to results.

Additional analyses and robustness checks

A number of additional robustness checks were undertaken. A descriptive analysis investigated whether randomisation was effective in ensuring that the intervention and control groups were balanced by comparing the characteristics of the groups prior to the intervention being delivered. This was done at both the school and pupil level and the absolute standardised difference across a range of characteristics was calculated. At the time of drafting the Statistical Analysis Plan, the intention was to produce a version of the primary analysis that included controls for any characteristics that were not balanced between the trial arms at baseline. 'Imbalance' was defined as having an Absolute Standardised Difference in excess of ten. In practice, the only characteristic where the level of imbalance met this threshold was participation in NELI. As this was the subject of subgroup analysis anyway, the additional analyses to address this imbalance at baseline is covered in the section on the NELI subgroup analysis.

Additional analyses were conducted to test the sensitivity of the impact estimates to the exclusion of controls for prior attainment. Also, the pre-test had to be conducted remotely in 17 schools due to the pandemic. As a result, it was decided to estimate the impact of Flexible Phonics in the subset of schools where all testing took place face to face. This was to assess whether the impact of Flexible Phonics was more apparent if the method of testing was in line with the original design of the study.

Estimation of effect sizes

Estimated impacts were calculated in accordance with the EEF analysis guide to aid comparability with other trials. This involved estimating Hedges' g based on total variance, rather than within-cluster variance. This also gave a more conservative estimate of impact compared with using within-cluster variance. The effect size equation is as follows:

$$ES = \bar{Y}t - \bar{Y}c \,\sqrt{(\sigma s\ 2 + \sigma error\ 2\ )}$$

where:

ES is the estimated effect size;

$\bar{Y}t - \bar{Y}c$ is the adjusted difference in mean outcomes between the treatment and control group; and

$\sqrt{\sigma s\ 2 + \sigma error\ 2}$ is the pooled unconditional variance of the treatment and control groups, taking into account both school level variance and pupil level variance.

Bayesian 95% credibility intervals are also reported and the parameters used in the calculations of the effect size are reported in Appendix C.

Estimation of ICC

The ICC for the post-test and pre-test is reported at school level. In schools with more than one reception year class per year group, one class was randomly selected to take part in testing. This means the ICC could also be interpreted as being at the class level.

Longitudinal analysis

In addition to these secondary outcomes, it was planned that the Year 1 statutory phonics screening test would be analysed to assess whether the intervention had had a long term impact on literacy. However, it was decided in collaboration with the EEF not to continue with this aspect of the research after reviewing findings from the main impact analysis.

Implementation and process evaluation

Research methods

Our implementation and process evaluation drew on the EEF guidance (Humphrey et al., 2016; EEF, 2019a) and used a multiphase, mixed methods design involving:

- two IDEA workshops and reviewing programme materials;
- observations of one pilot training day, and review of UCL pilot reports;
- observation of three online training half-days for the main trial and five online follow- up support sessions;
- online surveys (baseline and post-treatment/endline) of reception teachers and other staff to gather evidence about business as usual and changes to practice;
- school case studies;
- interviews with the delivery team; and
- analysis of data collected by the delivery team, for example, attendance and cost data.

The IDEA workshops

IES and UCL's team explored the intervention as part of an initial session shortly after set-up (October 2019) and then a second session took place in April 2021 after schools had reopened to all children in March 2021 and the team had an understanding of how schools were operating and delivering phonics teaching. There was a final short one and a half-hour session in June 2021 where the support partners had an opportunity to provide input into the model as well. Building on the set-up meetings, the team co-developed the TIDieR framework and theory of change, identified enabling factors and conditions for success, examined training and delivery materials, and revisited evidence about the interventions. The delivery team were then able to expand their initial logic model to include these additional aspects and reclassify the intervention activities and outcomes through a framework of inputs, activities, outputs, short-term outcomes and mediators, and long-term outputs and impacts. This laid a solid foundation for the evaluation and enabled the team to tackle key questions, such as an appropriate compliance measure, and to continue developing the ToC in light of the changes due to Covid-19 and the adaptations to online delivery.

Observations

As part of the pilot stage, the evaluation team was only able to observe one training day; no further attendance was possible due to the Covid-19 pandemic.[10] UCL also provided the team with two pilot reports to learn more about the intervention and how it developed through the pilot, which further informed the theory of change as well as the team's observation, interview, and survey materials for the main trial.

[10] We were due to also attend a follow-up session and two school support sessions.

The evaluation team observed three half-day training online sessions on the virtual meeting platform for the main trial (instead of the original two in-person training sessions planned) and five follow-up sessions (which were also virtual instead of in-person) to understand expectations for delivery and whether things had changed from the pilot and how effective the changes had been, to underpin the IPE. The observations also helped the team to develop well-tailored research instruments (case study topic guides, post-intervention survey of teachers and TAs). By observing all three online training half-days, the team were able to observe all of the training material being delivered and identify any possible differences between training cohort groups (of which there were six in total).

Teacher and TA surveys

The surveys of teachers and TAs were developed by IES using online survey software, SNAP, which allowed completion on computers and mobile devices. The baseline survey captured usual practice prior to randomisation and information on broader approaches to teaching reading, phonics, and spelling. The evaluation team used resources such as the Ofsted report on features of a 'good' and 'outstanding' reception curriculum, which includes characteristics of strong phonics teaching (Ofsted, 2017) and EEF guidance on improving literacy at Key Stage 1 (EEF, 2017) to formulate questions. The original research plan was to match individuals' responses at baseline and endline in order to support the identification of value-added impact, for example, in combination with data on past reading attainment, it would be possible to explore links between usual practice and past performance, differences in practice between control and intervention schools, as well as whether any overall differences in teaching practice at the outset have a bearing on the effectiveness of flexible phonics. Unfortunately, due to a technical error in the implementation of the baseline survey, it was no longer possible to identify individual responses in the baseline survey and, therefore, to measure changes between baseline and endline for individual respondents or to identify which respondents belonged to the intervention or control groups. However, it was still possible to explore business as usual and approaches to teaching reading in the endline survey, although it required respondents to report previous practice, so responses relied on recall.

The survey helped to identify the extent to which results may be explained by control schools improving phonics teaching (due to compensatory rivalry or other drivers) or the displacement of other literacy activities in treatment schools. The endline survey repeated questions about phonics teaching and, for treatment schools, covered experiences of taking part, staff time, and resources required (to inform the cost-per-pupil estimate). It included questions on adaptations made to the programme (beyond expected differentiation to meet the needs of individual pupils) and about participating in the NELI trial, where relevant, to understand how this may have also affected how staff support children's language development. The surveys were sent to approximately 613 teaching staff, including teachers and TAs, at least at both timepoints. The surveys were sent to all teachers and TAs of participating classes (not just the classes tested as part of the impact evaluation). Due to an error with the implementation of the baseline survey it was not possible to identify individual's responses to the baseline survey and match these to their responses in the endline survey. This meant it was not possible to identify and review any possible differences between non-responders and those who have completed the survey at endline, such as role, experience, type of school, confidence in phonics teaching at baseline, and so forth.

Case studies

Eight case studies were chosen to allow detailed qualitative exploration of delivery. Case study visits were planned to observe teaching and included interviews with the reception teachers or TAs who were involved in teaching phonics to the class participating in the study, literacy or early years leads, and a senior leader. However, due to Covid-19, case studies were conducted online and the observation component was not possible so the interviews were extended instead to include questions exploring practice and how they have integrated the Flexible Phonics approach with their usual approach. The sample was selected to include schools of different sizes and types and some schools that were participating in the NELI trial as well as looking at different types of pre-intervention phonics teaching. In each case study, we originally planned to observe teachers and TAs using Flexible Phonics strategies and use this data to inform assessment of fidelity and shape questions for follow-up interviews. However, as this was not possible due to pandemic restrictions, the evaluation team undertook longer interviews with teachers, TAs, a literacy lead or an early years lead, and a senior leader to explore how they had integrated the strategies into their phonics teaching and to elicit examples. The team conducted individual interviews with senior leaders and teachers and TAs separately to ensure open and honest discussions. Interviews explored:

- the training received;

- the materials;
- the workload and time requirements of Flexible Phonics, the costs incurred;
- facilitators and barriers to implementation;
- adaptations and the reasons for them—including views on how children with SEND or EAL respond;
- how Flexible Phonics compares to usual practice;
- pupils' outcomes;
- suggested improvements; and, where relevant,
- any changes to how they supported children's language and communication as a result of participating in the NELI trial.

Interviews lasted around 30 to 45 minutes. The team had planned to interview around three to four participants per case study so up to a total of 32 participants and achieved a total of 28 interviews. Findings from the case study interviews are reported as a mix of individual staff members' views and school-level practice and experience as relevant.

Delivery team interviews

The evaluation team also conducted six telephone interviews with UCL towards the end of intervention delivery, including Professor Savage, the project director, the project manager, and four of the seven support partners. As schools in England were closed from 5 January to 8 March 2021 and only delivering in-person teaching to the children of key workers or vulnerable children, contracts for some of the support partners were extended to offer support to schools for a longer period, once all children were back to in-person delivery. For this reason, support partners were interviewed at three timepoints to capture their experiences before their contract ended: late March, late May, and late June. These interviews explored delivering training, school engagement and participation, and enablers and barriers to successful implementation. This provided a rich picture of how training and support was delivered to schools across geographical areas, school types, and pre-existing phonics programmes.

Delivery team monitoring information

Finally, the evaluation team analysed data collected by UCL on compliance and contact via their support methods (telephone and email). This covered teacher and TA attendance at training (which will be used for the compliance analysis) and a rating of their fidelity to the intervention rated by the support partners during the follow-up visits. Contact by telephone or email was overall quite limited, so there was no data to assess on this. The team also collected cost data from UCL to calculate the cost per pupil, including fees charged and length of training visits.

Interviews with schools that withdrew

If a school withdrew from programme delivery or collection of child assessment data, UCL would inform the evaluation team so that they could be invited to participate in either a short telephone interview or answer questions on an email form in order to explore any reasons for their withdrawal. At the point when qualitative fieldwork was taking place, there were only three schools that had dropped out since baseline testing; the team conducted interviews with two of these—one staff member at one school and two at the other.

Analysis

Interviews were digitally recorded with the agreement of participants and transcribed verbatim. We analysed IPE data using Framework, drawing themes and messages from an analysis of interview transcripts, observations of training, and other materials collected by evaluation and project teams as a pragmatic, cost-effective approach for this amount of qualitative data. Data was collected using the methods described in Table 2 and analysed according to the research questions listed.

Framework is an Excel-based qualitative analysis tool that ensures that the analytical process and interpretations from it are grounded in the data and tailored to the research questions. Relevant information from interview transcripts or other sources is extracted and summarised against key themes with key quotes noted. Thematic analysis was used to identify key issues, views, and experiences regarding programme implementation and delivery. Framework was designed to ensure a systematic and consistent treatment of all units of data (for example, transcripts of interviews). It also allows for the analytical framework to be refined and modified in the early stages of its use.

The context of the information is retained and the page of the transcript from which it comes is noted so that it is possible to return to a transcript to explore a point in more detail or to extract text for verbatim quotations. Framework allows for full within-case analysis (looking in detail at each individual case) and between-case analysis (comparing individual cases and groups of cases), and it is the ability to interrogate data at both these levels that adds real richness and depth to the analysis and interpretation. Organising the data in this way allows us to compare the full range of experiences and accounts and patterns across different groups of people.

Observations and themes identified in the qualitative data through the fieldwork were then compared with quantitative data gathered, such as survey findings and training attendance, to test whether perceptions are reflected across the settings overall.

Table 2: IPE methods overview

Research methods	Data collection methods	Participants/data sources	Data analysis methods	Research questions addressed	Implementation/ logic model relevance
Theory of change development	Two IDEA workshops	Evaluation team and delivery team	Theory of change	IPE 5, 15.	Theory of change, Inputs, Activities
Material review	Reviewing intervention materials	Training materials, support materials	Literature review, thematic analysis	IPE 2, 3, 16, 18.	Inputs, Activities
Observations of pilot	Observations of pilot training day	Delivery team (Prof. Savage) and reception teachers/TAs from pilot schools	Observation framework, Theory of change	IPE 2, 7, 10, 18.	Input, Activities, Outputs
Observations of trial	Observation of three online training half days for the main trial and three follow-up training sessions	Delivery team (Prof. Savage and RAs) and reception teachers/TAs from half of schools receiving the intervention (~25–35 schools)	Observation framework	IPE 2, 4, 7, 10, 15, 18.	Input, Activities, Outputs
School staff survey	Online surveys (baseline and post-treatment) of reception teachers/other staff	reception teachers/TAs from all schools participating in the study (≤120)	Descriptive: frequencies, cross-tabs, t-tests, ANOVA, regression	IPE 5, 7, 9, 10, 12, 13, 15, 17, 19–21.	Activities, Outputs, Outcomes, Enabling factors
School case studies	Case study including observations and interviews	8 intervention schools, 8 or more reception teachers/TAs, up to 8 literacy or Early Years Leads, up to 8 senior leaders;	Teaching observation framework, Extraction framework	IPE 1–3, 4–5, 6–12, 13–14, 15–17, 18–19, 22.	Inputs, Activities, Outputs, Outcomes, Enabling factors
Delivery team interviews	Interviews with UCL	Intervention designer/trainer: Prof. Savage, Amy Fox project	Extraction framework	IPE 1–4, 7, 15, 18.	Inputs, Activities, Enabling factors

		manager, RAs undertaking support visits			
Monitoring Information	Analysis of intervention data collected by UCL	Training attendance, summary of compliance, satisfaction surveys and cost data	Thematic analyses, Descriptives: frequencies, cross-tabs, average cost per child.	IPE 2–4, 7, 13, 15.	Inputs, Activities, Outputs

Costs

Details of costs were obtained from the endline quantitative survey and in-depth qualitative interviews with school staff and the delivery team. These findings have informed our costs analysis and are also reported as part of the IPE section.

Endline survey

Teachers and TAs in both the intervention and control groups filled out the endline survey at the end of the trial. In the control group, they were asked what costs were involved with delivering their usual phonics programme—to briefly describe the type of expense and give an idea of the approximate cost. Teachers and TAs in the intervention group were asked whether there had been any extra costs involved with delivering the Flexible Phonics programme and, if so, to briefly describe the type of expense and give an idea of the approximate cost. The intervention group also answered questions about the average number of hours needed each week to prepare and deliver flexible phonics as well as whether any cover was required for training or follow-up sessions. Finally, the intervention group were asked to report whether they had required support from senior or specialist staff at their school to deliver Flexible Phonics and, if so, to estimate the total number of hours support they received.

Qualitative interviews

In-depth qualitative interviews were carried out with teachers and TAs—and senior leadership team members or subject leads in selected case-study schools—to facilitate a more detailed understanding of any costs or resources needed to participate in the programme. Thirty-five staff were interviewed across eight schools. Teachers and TAs were asked how much time they needed to plan and deliver the programme, how much time they had needed for training, and if there were any extra costs with delivering it at their school (for example, cover for training, extra staff, resources, books, or printing and photocopying). They were also asked whether they had needed any support from the senior leadership team. SLT members and specialist leads were asked whether staff had needed support to deliver the programme, how much time they had needed, and whether there had been any extra costs associated with delivery.

In-depth qualitative interviews were also completed with members of the delivery team. This included the project director, the project manager, and six of the eight support partners. The project director and project manager were asked what their costs were in delivering the programme and how much time they spent on the project. This included time for delivering the training, time for training and supporting the support assistants, time for delivering the programme more generally, any unanticipated costs, any impacts of the Covid-19 pandemic on time or costs, and any anticipated changes to costs or time if the programme were delivered at scale. The project manager also completed a costs pro-forma with detailed information about costs and time needed by the delivery team during the training and delivery of Flexible Phonics. The support assistants were asked to estimate the average hours needed per week to support schools with delivery.

Limitations

The number of respondents from the survey answering the costs questions about the type of expense and the approximate cost was small (intervention group, n = 7; control group, n = 31). It is unclear why there were fewer responses from those in the intervention group as there was the same number of respondents from the intervention and control groups (120 and 120 respectively) and the question wording was similar for both. However, respondents in the intervention group would have answered a larger number of questions before being presented with the costs question compared to respondents in the control group, as they were asked to answer questions about their experience of the

intervention before being asked about the costs, so may have been more survey fatigued at that point. Also, respondents from the intervention group may have found it more complicated to work out costs as they were not delivering their usual phonics approach. Therefore, we have treated this data as qualitative data and not included it in the cost estimate models. The number of respondents answering the questions about their time was larger and so can be treated as more secure (n = 72 for the prep time question and n = 69 for the delivery time question).

Price per pupil per year calculation

The main model reported in the Costs section shows little to no costs were incurred for the schools in this trial and therefore no further calculation was needed for the per pupil per year cost as it still sits at zero. In Appendix P we provide an alternative cost model that provides estimated costs assuming face to face training and follow-up sessions as originally intended, and also that schools would need to pay a programme fee to cover costs for printing training manuals, direct mapping books, and other such costs that were provided for free by the delivery team's institution, UCL. This alternative model calculates a price per pupil based on a three-year time horizon, following the EEF cost evaluation guidance (EEF, 2019b) and is based on the average number of pupils per class in this trial. Costs are calculated for a two-form school with an average of 27 pupils per class.

Timeline

Table 3: Flexible Phonics project timeline

Dates	Activity	Staff responsible / leading
Jun–Oct 2019	Set-up meetings and first IDEA workshop.	Delivery team and evaluation team
Jan–Feb 2020	Recruitment of pilot schools.	Delivery team
Dec 2019–Jul 2020	Recruitment of trial schools.	Delivery team with support from evaluation team
Oct 2019–Jun 2020	Pre-trial development of programme; observation of pilot training session.	Delivery team Evaluation team
Jul 2020	Webinars—school information session for participating schools.	Delivery Team with support from Evaluation and Assessment Teams
Sep–Dec 2020	Collection of pupil data; collection of pre-test data; business as usual survey of teachers/TAs; randomisation.	Delivery team and evaluation team (overseeing Qa Research test administrator)
Jan–Jun 2021	School training days by end Feb; school follow-up sessions completed by mid-June; second IDEA workshop in April 2021; observation of school training sessions; schools deliver Flexible Phonics.	Delivery team and evaluation team
Apr–Jun 2021	Collection of data from delivery team; school case studies (observation, senior leader, teacher/TA and literacy lead/early years lead/SENCO interviews).	Evaluation team
Jun–Jul 2021	Administration of post–test assessments.	Evaluation team (overseeing Qa Research test administrator)
Jun–Jul 2021	Post-intervention survey of teachers/TAs.	Evaluation team
Autumn 2021–spring 2022	Analysis of project and evaluation data.	Evaluation team

Dates	Activity	Staff responsible / leading
Dec 2021–Aug 2022[11]	Evaluation report writing.	Evaluation team
30 Aug 2022	First draft of evaluation report.	Evaluation team
Sep 2022	Obtain NPD data for Year 1 phonics.	Evaluation team
Autumn 2022	Analysis of Year 1 phonics and evaluation data for addendum report.	Evaluation team
Jan–Feb 2023	Addendum report writing.	Evaluation team
Mar 2023	First draft of addendum report.	Evaluation team

[11] Report writing was delayed due to access issues with the Secure Research Service.

Impact evaluation results

Participants

The intervention was targeted at children in reception classes at primary schools. Schools were recruited into the trial and could withdraw at various stages. The delivery team maintained a waitlist so that they could replace schools that dropped out. Figure 2 shows the numbers of schools and pupils in the trial at each of these different stages in the project. Of the 149 schools invited to take part in the trial, eight failed to submit teacher and class names before class randomisation. However, 18 of the schools that initially expressed a willingness to take part withdrew prior to randomisation to the treatment and control groups. In eight cases the withdrawal occurred before a single class was randomly selected for testing, while in three cases the class selection was made before the school withdrew. A further seven schools supplied pupil data before withdrawing.

In total, 123 schools were randomised to either the intervention or control group. Three assigned to the intervention group withdrew from the trial after randomisation. Reasons for withdrawing included staff shortages caused by Covid-19, the reception teacher in a one-form entry school being off on long-term sick leave with Covid-19, and a school closure planned for the following year, so parents were moving their children to other schools. This left a total of 120 schools in the randomisation sample, with 59 schools and 1,567 pupils in the intervention group and 61 schools and 1,599 pupils in the control group. After randomisation the initial sample was 3,166 pupils.

One school in each of the trial arms withdrew before the post-test. Some pupils were withdrawn from the project during the trial or were absent from school when either the pre-test or post-test was conducted. The final sample used for the analysis of the primary outcome consisted of a total of 118 schools, with 58 schools and 1,256 pupils in the intervention group and 60 schools and 1,283 pupils in the control group.

From the total initial sample, 627 pupils dropped out and final analysis included 2,539 pupils. The pupil attrition rate is 20%.

Figure 2: Participant flow diagram (two arms)

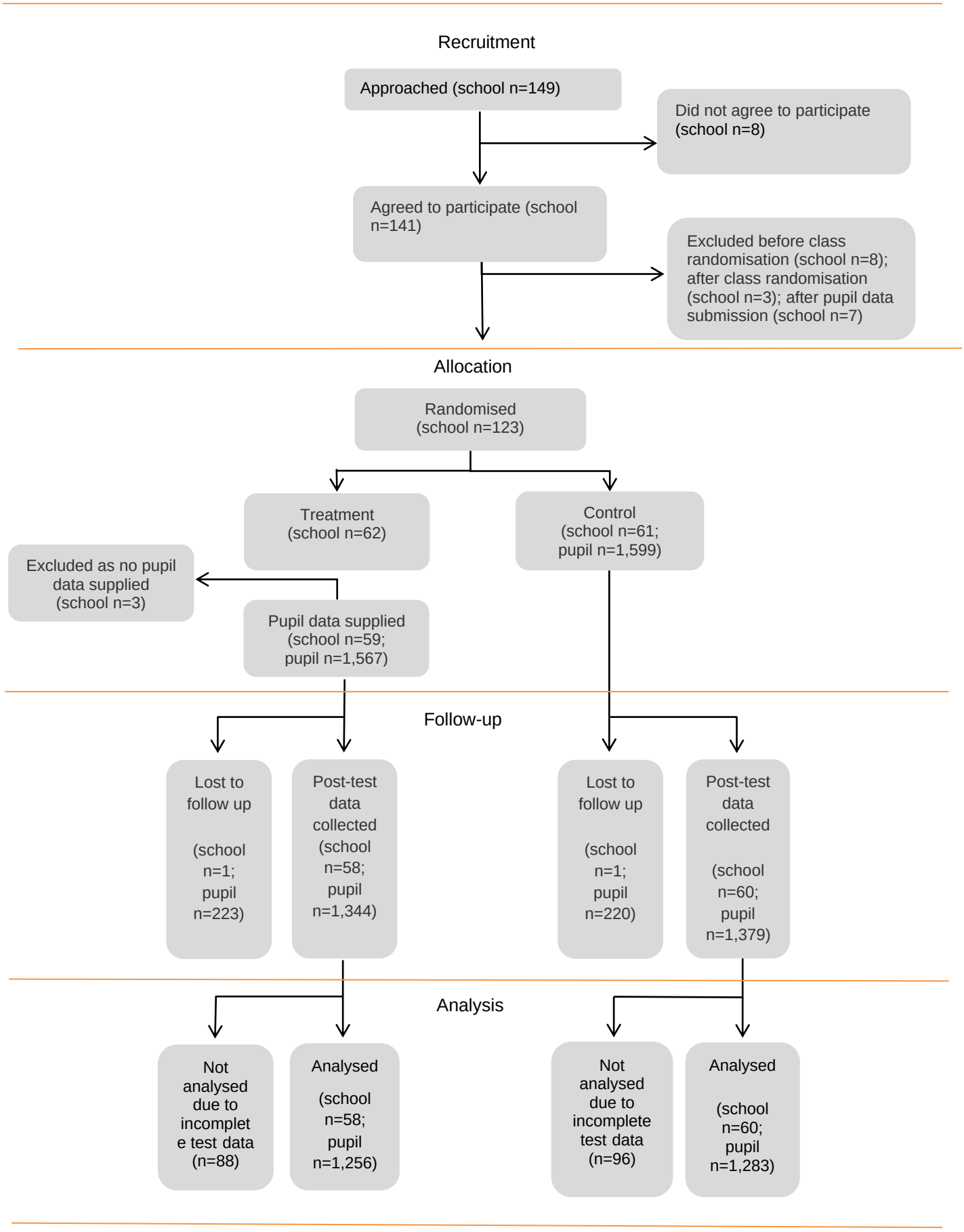


Table 5 shows the minimum detectable effect size (MDES) at different stages of the trial. The MDES for the analysis (0.17 standard deviations) is smaller than in both protocol and randomisation stages (0.23 and 0.21) for two reasons. Firstly, a larger sample size than expected was achieved. This came about through the over-recruitment of schools. Secondly, it was assumed pre-test outcomes would explain 40% of the variation in post-test outcomes but the observed correlation was considerably larger at 54%. The ICC indicates that 13% of the variation in the post-test score comes from variation across schools. Although this is slightly smaller than the estimated 15% at protocol and randomisation, it is still indicative that a multi-level model should be used to account for clustering.

The table also shows the MDES for the subset of pupils who were eligible for free school meals (FSMs). Again, the MDES of 0.29 standard deviations for the analysis sample was smaller than that anticipated at the initial design phase, or at the time of randomisation. This was because the number of pupils eligible for FSMs in the schools which participated in testing was higher than expected (an average of 4 pupils per school, rather than 3). Also, the correlation between pre- and post-test scores for the FSM subsample was higher than expected, at 0.63, compared with a correlation coefficient of 0.40 assumed at the time of randomisation. The correlation between pre- and post-test scores was also higher for pupils eligible for FSM than for the wider group of pupils who took part in the trial. The ICC was similar to that seen for the full analysis sample, at 0.12.

The calculations of MDES for the FSM subgroup at the design and randomisation stages did not allow for the fact that some schools would not have any pupils eligible for FSMs. Only 101 of the 108 schools in the analysis sample had any pupils eligible for FSMs, but the number of schools with pupils eligible for FSMs in each trial arm was fairly balanced (49 in the intervention group and 52 from the control group, as shown in Table 5). Overall, the number of pupils eligible for FSMs in the analysis sample slightly exceeded expectations, with similar numbers in the intervention and control groups (211 and 225 pupils respectively).

Table 4: Minimum detectable effect size at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
MDES (standard deviations)		0.23	0.37	0.21	0.33	0.17	0.29
Pre-test/post-test correlations	Level 1 (pupil)						
	Level 2 (class)	0.4	0.4	0.4	0.4	0.54	0.63
	Level 3 (school)						
Intracluster correlations (ICCs)	Level 2 (class)						
	Level 3 (school)	0.15	0.15	0.15	0.15	0.13	0.12
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided	Two-sided	Two-sided	Two-sided	Two-sided
Average cluster size		23	3	22	3	22	4

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
Number of schools	Intervention	50	50	59[12]	59	58	49
	Control	50	50	61	61	60	52
	Total:	100	100	120	120	118	101
Number of pupils	Intervention	1,150	150	1,567	186	1,256	211
	Control	1,150	150	1,599	183	1,283	225
	Total:	2,300	300	3,166	369	2,539	436

Attrition

Table 5 provides an overview of attrition rates for both the intervention and the control group. If pupils randomised to either of the two trial arms were missing pre- or post-test scores on the primary outcome measure of Early Word Recognition, this was regarded as attrition. Attrition could occur for one of several reasons: pupil absence on the day of the test, a pupil being withdrawn from testing by the teacher, or the pupil moving to another school during the trial. A pupil may be withdrawn by a teacher on the day of testing (instead of a parent withdrawing their child in advance of testing) because the teacher feels that the child would not be suitable for testing or where assessment may not be appropriate or cause distress because of a learning or language difficulty, behavioural difficulties, or other issues. During the pre-test, four children were withdrawn by a teacher because they had complex SEN needs and were not on the reception curriculum so this would have disrupted their routines and two children had not started yet at the time of assessment. The attrition rate was very similar for both arms, with around 20% randomised to either trial arm not forming part of the analysis sample. This was higher than the assumption of 15% attrition made when producing the power calculations. However, given the additional uncertainties created by the Covid-19 pandemic and the need to carry out the pre-test remotely in 17 schools, it is unsurprising that the actual rate of attrition differed from initial expectations. Potentially, a higher than expected rate of attrition might reduce the likelihood of detecting any impact from the intervention and mean that the impact estimates were not representative of the impact of Flexible Phonics across all participants.

Table 5: Pupil-level attrition from the trial (primary outcome)

		Intervention	Control	Total
Number of pupils	Randomised	1,567	1,599	3,166
	Analysed	1,256	1,283	2,539
Pupil attrition (from randomisation to analysis)	Number	311	316	627
	Percentage	20%	20%	20%

Pupil and school characteristics

Table 6 reports the baseline characteristics of schools and pupils randomised to the intervention and control groups. It also shows how the characteristics of schools participating in the trial compared to state funded primary schools across

[12] The figures for the randomised sample exclude three schools allocated to the intervention that were randomised in error. As these schools did not supply pupil data, they should not have been randomised and so were excluded from the evaluation dataset.

England as a whole. The section on imbalance at baseline assesses whether the observed differences between the intervention and control groups were likely to bias the impact estimates.

As schools participating in the trial were based in Greater London, they were all located in urban areas. Nationwide, just over two-thirds (68%) of state-funded primary schools are in major urban conurbations or cities and towns. Across England as a whole, state-funded primary schools are most likely to be academies or academy converters or free schools, whereas community schools were most common among schools participating in the trial. Roughly half of the schools in the intervention and control groups were community schools, compared with just over one-third of primary schools across England. The percentage of voluntary aided, voluntary controlled, and foundation schools in the intervention and control groups was similar to the percentage across England as a whole, but academy schools, free schools and those converting to academy status were under-represented in the intervention and control groups.

State-funded primary schools in England had an average of one in five (21%) pupils eligible for free school meals. This percentage was similar among schools randomised to the intervention group but was closer to one in four (24%) among schools in the control group. However, schools in either trial arm were very similar to primary schools nationally in terms of the percentage of pupils who were female (49%).

Across England as a whole, 17% of primary schools achieved an overall rating of ‘outstanding’. Nearly double this percentage (36%) of intervention schools were judged to be ‘outstanding’ compared with 25% of schools assigned to the control group. As a very small number of trial schools were actually rated as ‘inadequate’, it was necessary to group the ‘good’ and ‘inadequate’ categories together. The percentage of schools in the control group rated as ‘good’ or ‘inadequate’ was similar to the national average (75% and 74%, respectively), whereas only 64% of schools in the intervention group were assessed as ‘good’ or ‘inadequate’. None of the schools randomised were considered to require improvement or were without an Ofsted rating, compared with around 10% of all state-funded primaries.

Forty-four percent of schools assigned to the intervention group had some pupils participating in the Nuffield Early Language Intervention compared with 40% of schools in the control group. The evaluation team only became aware after randomisation that some Flexible Phonics schools had also signed up for NELI, which was rolled out as part of the government’s Covid-19 support strategy during the same school year. Remote testing of pupils at the pre-test phase was less common in schools that were part of the intervention group compared with schools in the control group (12% and 16%, respectively). Schools participating in the trial had a larger mean average number of pupils on the roll than the national average for state-funded primary schools. Whereas nationally schools had an average of 281 pupils, the average number of pupils at schools in the intervention group was 432. Schools assigned to the control group were slightly smaller, on average, with 412 pupils.

Turning to pupil-level characteristics: at the time of randomisation, the percentage of FSM pupils in the reception classes selected to take part in the trial was similar in both arms, at 19%; the percentage of female pupils was also similar, at 48%. Pupils who scored below the median in the EWR pre-test were defined as being of below average ability in terms of literacy. A slightly higher percentage of those in the intervention group had a low score on this measure compared with the control group (52% and 48%, respectively). Pupils randomised to the intervention group were more likely to receive support from NELI than those in the control group (12% and 8%, respectively), which could potentially bias the estimated impact of Flexible Phonics.

The intervention and control groups were very similar in terms of age: pupils in both groups had an average of 56 months (four years and eight months) at the time of the pre-test and 64 months (five years and four months) at the time of post-intervention testing. The reported effect size for the differences in pre-test scores between the intervention and control groups demonstrate that the intervention group performed worse than the control group on both the EWR measure and the composite measure of the standardised EWR and LSK subscales. This imbalance in test scores prior to the intervention might potentially bias the impact estimates. The likelihood of this is considered further in the section on Imbalance at Baseline.

The final rows of the table report EWR pre-intervention raw test scores for each of the subgroups considered.[13] Pre-test scores for pupils randomised to the intervention and control groups were similar for those who scored below the median in the EWR pre-test and those who attained a score at the median or above. Within schools where at least some pupils were participating in NELI, pupils at schools assigned to the control group for the Flexible Phonics trial had a higher pre-test score than pupils at schools that were assigned to the Flexible Phonics intervention group. The pattern of pupils in schools assigned to the Flexible Phonics control group achieving higher pre-test EWR scores than those assigned to the intervention group was also evident in schools where no pupils were participating in NELI.

On average the mean EWR pre-test score was higher for pupils eligible for FSMs who were assigned to the intervention group than those assigned to the control group. However, across both the intervention and control groups, pupils who were not eligible for FSMs had a higher EWR pre-test score than pupils who were eligible for FSMs. Within the subset of pupils who were not eligible for FSMs, those assigned to the intervention group had a lower EWR pre-test score than those assigned to the control group.

Table 6: Baseline characteristics of groups as randomised

School-level (categorical)	National-level mean	Intervention group		Control group	
		n/N (missing)[14]	Count (%)	n/N (missing)	Count (%)
Rural hamlet and isolated dwellings	4.40%				
Rural hamlet and isolated dwellings in a sparse setting	0.49%				
Rural town and fringe	10.51%				
Rural town and fringe in a sparse setting	0.50%				
Rural village	12.20%				
Rural village in a sparse setting	0.89%				
Urban city and town in a sparse setting	0.14%				
Urban major conurbation and urban city and town[15]	67.57%	59/59 (0)	59 (100%)	*/61 (<3)	* (100.00%)
Urban minor conurbation	3.31%				
School type					

[13] This text and the rows described in the table have been added since the previous version of the report.
[14] The exact number of missing values is only reported if these meet the threshold for statistical disclosure control, namely, three or more for school-level variables and ten or more for pupil-level variables. Where the number of missing values is below these thresholds the number of observations in a category (n and Count) are suppressed—indicated by an *—and the total number of observations (N) is adjusted to include the missing cases. Percentages are calculated excluding the missing cases from the denominator.
[15] These categories are combined due to low cell sizes for the randomised sample.

Academy converter; academy sponsor led and free schools[16]	36.78%	14/59 (0)	14 (23.73%)	*/61 (<3)	* (20.34%)	
Community school	34.70%	29/59 (0)	29 (49.15%)	*/61 (<3)	* (52.54%)	

School-level (categorical)	National-level mean	Intervention group		Control group		
		n/N (missing)[17]	Count (%)	n/N (missing)	Count (%)	
Voluntary aided schools; voluntary controlled schools; and foundation schools[18]	28.52%	16/59 (0)	16 (27.12%)	*/61 (<3)	* (27.12%)	
Percentage of pupils at school eligible for FSMs	20.68%	59/59 (0)	59 21.41%	*/61 (<3)	24.19%	
Percentage of pupils at school who are female	49.11%	59/59 (0)	59 49.24%	*/61 (<3)	49.32%	
Ofsted rating						
Outstanding	16.50%	21/59 (0)	21 (35.59%)	15/61 (0)	15 (24.59%)	
Good/inadequate[19]	73.63%	38/59 (0)	38 (64.41%)	46/61 (0)	46 (75.41%)	
Requires improvement	9.31%	0/59 (0)	0 (0.00%)	0/61 (0)	0 (0.00%)	
Null	0.55%	0/59 (0)	0 (0.00%)	0/61 (0)	0 (0.00%)	
Some pupils participating in NELI	-	26/58 (1)	26 (44.83%)	24/60 (1)	24 (40.00%)	
Remote testing at pre-test	-	7/59 (0)	7(11.86%)	10/61 (0)	10 (16.39%)	
School-level (continuous)	Mean (SD)	n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	
No. of pupils on role	281.05 (162.90)	59/59 (0)	432.20 (202.35)	61/61 (0)	412.43 (216.05)	
Pupil-level (categorical)[20]		n/N (missing)	Count (%)	n/N (missing)	Count (%)	

[16] Categories are combined due to low cell sizes for the randomised sample.
[17] The exact number of missing values is only reported if these meet the threshold for statistical disclosure control, namely, three or more for school-level variables and ten or more for pupil-level variables. Where the number of missing values is below these thresholds the number of observations in a category (n and Count) are suppressed—indicated by an *—and the total number of observations (N) is adjusted to include the missing cases. Percentages are calculated excluding the missing cases from the denominator.
[18] Categories are combined due to low cell sizes for the randomised sample.
[19] The 'good' and 'inadequate' categories are combined throughout due to the very low number of trial schools rated as 'inadequate'.
[20] As we do not have access to national pupil-level data, these cells are left blank.

Pupil eligible for FSM according to the pupil-level census		289/1,550 (17)	289 (18.65%)	*/1,599 (<10)	* (19.36%)	
Female		755/1,567 (0)	755 (48.18%)	773/1,599 (0)	773 (48.34%)	
Low-ability pupils		738/1,413 (154)	738 (52.23%)	689/1,445 (154)	689 (47.68%)	

		Intervention group		Control group		
Pupil-level (categorical)[21]		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Participating in NELI		180/1,534 (33)	180 (11.73%)	129/1,568 (31)	129 (8.23%)	
Pupil-level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Effect size
Age in months at time of pre-test		1,414 (153)	56.20 (3.61)	1,456 (143)	56.29 (3.58)	-
Age in months at time of post-test		1,344 (223)	63.56 (3.64)	1,379 (220)	63.58 (3.62)	-
EWR rawscore at time of pre-test		1,413 (154)	3.65 (6.24)	1,445 (154)	3.98 (6.54)	-0.33
Standardised pre-test for secondary outcomes		1,413 (154)	-0.08 (1.78)	1,445 (154)	0.09 (1.77)	-0.17
EWR rawscore at time of pre-test for low ability pupils		731 (176)	0.29 (0.84)	665 (181)	0.25 (0.57)	0.04
EWR rawscore at time of pre-test for pupils of higher ability		660 (176)	7.39 (7.44)	753 (181)	7.39 (7.58)	0.00
EWR rawscore at time of pre-test for pupils at schools participating in NELI		601 (176)	3.43 (6.32)	537 (181)	3.96 (6.70)	-0.53
EWR rawscore at time of pre-test for pupils at schools not participating in NELI		790 (176)	3.84 (6.22)	881 (181)	4.09 (6.51)	-0.25
EWR rawscore at time of pre-test for pupils eligible for FSMs		242 (31)	2.33 (5.43)	258 (47)	2.24 (5.35)	0.09
EWR rawscore at time of pre-test for pupils not eligible for FSMs		1149 (101)	3.94 (6.39)	1160 (102)	4.44 (6.76)	-0.50

[21] As we do not have access to national pupil-level data, these cells are left blank.

Table 7 reports the baseline characteristics of schools and pupils in the intervention and control groups from the sample included in the analysis. As the vast majority of schools randomised also took part in testing at the post-intervention stage, the baseline characteristics of the analysis sample were very similar to those of the randomisation sample.

Once again, as schools participating in the trial were based in Greater London, they were all located in urban areas. Community schools were over-represented in the final sample of intervention and control schools for analysis compared with primary schools in England as a whole, while academy, academy converters, and free schools were under-represented. The percentage of voluntary aided, voluntary controlled, and foundation schools was similar in both trial arms (28%) to the national average (29%).

The percentage of FSM pupils across schools participating in the intervention was similar to the national average (21%) but schools in the control group had a higher percentage at 24%. Schools participating in the trial had a similar percentage of female pupils to schools nationally, at 49%. As noted, as the trial was targeted at schools with an Ofsted rating of 'outstanding' or 'good', the analysis sample was unrepresentative of state-funded primary schools nationally in this regard. Over one-third (36%) of schools in the intervention group were considered 'outstanding' compared with 23% in the control group—and only 17% of schools nationally. No schools in the analysis sample were judged to require improvement or were without an Ofsted rating, compared to around 10% of state-funded primary schools across England as a whole.

Forty-five percent of schools in the intervention group analysis sample had some pupils who were receiving support from NELI compared with 40% of schools in the control group. Schools in the intervention group were less likely to have opted for remote testing at the pre-test stage than schools in the control group (12% and 17%, respectively).

As with the randomisation sample, the average number of pupils at schools participating in the trial was much larger than the national average of 281. As noted, this is unsurprising given that the trial schools were located in Greater London where multi-form entry is common. The average number of pupils in intervention and control schools in the analysis sample was very similar to that seen for the randomisation sample. Intervention schools in the analysis had an average of 429 pupils compared with 414 at control schools.

Turning to the characteristics of pupils within the reception classes selected to take part in the trial, a similar percentage of pupils in the intervention and control groups were eligible for FSM (17% and 18%, respectively). A slightly higher percentage of pupils were female in the intervention group than in the control group (50% compared with 48%). Pupils in the intervention group were slightly more likely to be of below-average ability based on their pre-test performance in the EWR subscale than those in the control group (51% and 47%, respectively). Those in the intervention group were also more likely to receive support from NELI than those in the control group.

Again, pupils in the intervention and control groups were a similar age at the time of the pre-test and post-test, at an average of 56 months (four years, eight months) at the time of the pre-test and 64 months (five years, four months) at the time of the post-test. As with the randomisation sample, the effect sizes indicate that pupils in the intervention group tended to perform worse than the control group on both the EWR pre-test and the composite measure constructed from the standardised EWR and LSK subscales. This highlights the need to control for pre-test scores in the analysis of the primary and secondary outcomes to adjust for differences in literacy between the two groups at baseline. Post-test scores for the EWR raw score, the full YARC and the MCT were all slightly higher for the control group than for the intervention group.

For the subgroups considered in the analysis, the patterns in the pre-test EWR data for the analysis sample were similar to those seen for the randomised sample.[22] Pupils who were below the median EWR raw score at the time of the pre-test who were at schools assigned to the control group performed better in the post-test than those in the intervention group. However, pupils with higher levels of performance in the pre-test at schools allocated to the intervention group achieved slightly higher post-test scores than those in the control group. Pupils at schools participating in NELI achieved

[22] The text in the remainder of this section and the rows described in the table have been added since the previous version of the report.

higher post-test EWR scores on average if they were part of the Flexible Phonics intervention group compared with those in the control group. However, for those who were at schools where no pupils took part in NELI, EWR post-test scores were lower for pupils in the Flexible Phonics intervention group compared with pupils in the control group.

Among pupils who were eligible for FSMs, those assigned to the intervention group had a higher average EWR pre-test score than those assigned to the control group. As with the randomised sample, this pattern was reversed for the subset of pupils who were not eligible for FSMs, with those assigned to the control group scoring more highly in the EWR pre-test than those assigned to the intervention group. Across both the intervention and control groups, pupils who were not eligible for FSMs on average had higher mean EWR pre-test scores than pupils who were eligible for FSMs. Similar patterns were evident in the EWR post-test, with those eligible for FSMs achieving lower post-test scores than those not eligible for FSMs, irrespective of whether they were assigned to the intervention or control group. Among those eligible for FSMs, the mean post-test score was higher for the intervention group than for the control group, whilst for pupils not eligible for FSMs, those assigned to the intervention group had a lower post-test score than those assigned to the control group.

Table 7: Baseline characteristics of groups as analysed

School-level (categorical)	National-level mean	Intervention group		Control group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
Rural hamlet and isolated dwellings	4.40%				
Rural hamlet and isolated dwellings in a sparse setting	0.49%				
Rural town and fringe	10.51%				
Rural town and fringe in a sparse setting	0.50%				
Rural village	12.20%				
Rural village in a sparse setting	0.89%				
Urban city and town in a sparse setting	0.14%				
Urban major conurbation and urban city and town[23]	67.57%	58/58 (0)	58 (100.00%)	*/60 (<3)	* (100.00%)
Urban minor conurbation	3.31%				
School type					
Academy converter; academy sponsor led and free schools[24]	36.78%	13/58 (0)	13 (22.41%)	*/60 (<3)	* (20.69%)
Community school	34.70%	29/58 (0)	29 (50.00%)	*/60 (<3)	* (51.72%)
Voluntary aided schools; voluntary controlled schools; and foundation schools.[25]	28.52%	16/58 (0)	16 (27.59%)	*/60 (<3)	* (27.59%)
Percentage of pupils at school eligible for FSMs	20.68%	58/58 (0)	58/20.98%	*/60 (<3)	*/24.41%
Percentage of pupils at school who are female	49.11%	58/58 (0)	58/49.22%	*/60 (<3)	*/49.39%
Ofsted rating					
Outstanding	16.50%	21/58 (0)	21 (36.21%)	14/60 (0)	14 (23.33%)
Good/inadequate[26]	73.63%	37/58 (0)	37 (63.79%)	46/60 (0)	46 (76.67%)
Requires improvement	9.31%	0/58 (0)	0 (0.00%)	0/60 (0)	0 (0.00%)

[23] These categories are combined due to low cell sizes for the analysed sample.
[24] Categories are combined due to low cell sizes for the analysed sample.
[25] Categories are combined due to low cell sizes for the analysed sample.
[26] The ‘good’ and ‘inadequate’ categories are combined throughout due to the very low number of trial schools rated as ‘inadequate’.

School-level (categorical)	National-level mean	Intervention group		Control group		
		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Null	0.55%	0/58 (0)	0 (0.00%)	0/60 (0)	0 (0.00%)	
Some pupils participating in NELI		26/58 (0)	26 (44.83%)	24/60 (0)	24 (40.00%)	
Remote testing at pre-test		7/58 (0)	7 (12.07%)	10/60 (0)	10 (16.67%)	
School-level (continuous)	Mean (SD)	n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	
No. of pupils on role	281.05 (162.90)	58/58 (0)	428.84 (202.45)	60/60 (0)	413.78 (217.61)	
Pupil-level (categorical)		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Pupil eligible for FSM according to the pupil-level census		211/1,256 (0)	211 (16.80%)	225/1,283 (0)	225 (17.54%)	
Female		626/1,256 (0)	626 (49.84%)	620/1,283 (0)	620 (48.32%)	
Low-ability pupil		641/1,256 (0)	641 (51.04%)	604/1,283 (0)	604 (47.08%)	
Participating in NELI		161/1,256 (0)	161 (12.82%)	114/1,283 (0)	114 (8.89%)	
Pupil-level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Effect size
Age in months at time of pre-test		1,256/1,256 (0)	56.22 (3.61)	1,283/1,283 (0)	56.30 (3.58)	-
Age in months at time of post-test		1,256/1,256 (0)	63.55 (3.64)	1,283/1,283 (0)	63.53 (3.58)	-
EWR rawscore at time of pre-test		1,256/1,256 (0)	3.75 (6.38)	1,283/1,283 (0)	4.08 (6.65)	-0.33
Standardised pre-test for secondary outcomes		1,256/1,256 (0)	-0.03 (1.80)	1,283/1,283 (0)	0.12 (1.79)	-0.15
EWR rawscore at time of post-test		1256/1256 (0)	13.44 (8.99)	1283/1283 (0)	13.96 (8.83)	-0.52
YARC post-test		1,256/1,256 (0)	-0.06 (3.39)	1,283/1,283 (0)	0.15 (3.14)	-0.22
MCT post-test		684/684 (0)	9.71 (9.08)	718/718 (0)	10.33 (9.11)	-0.62

		Intervention group		Control group		
Pupil-level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Effect size
EWR rawscore at time of pre-test for low ability pupils		641/641 (0)	0.26 (0.65)	604/604 (0)	0.26 (0.59)	0.00
EWR rawscore at time of pre-test for pupils of higher ability		615/615 (0)	7.39 (7.53)	679/679 (0)	7.47 (7.66)	-0.08
EWR rawscore at time of post-test for low ability pupils		641/641 (0)	8.09 (6.83)	604/604 (0)	8.72 (6.95)	-0.63
EWR rawscore at time of post-test for pupils of higher ability		615/615 (0)	19.02 (7.46)	679/679 (0)	18.63 (7.62)	0.39
EWR rawscore at time of pre-test for pupils at schools participating in NELI		544/ (0)	3.62 (6.49)	486/ (0)	3.90 (6.70)	-0.28
EWR rawscore at time of pre-test for pupils at schools not participating in NELI		712/ (0)	3.86 (6.29)	797/ (0)	4.18 (6.62)	-0.32
EWR rawscore at time of post-test for pupils at schools participating in NELI		544/ (0)	13.81 (9.24)	486/ (0)	13.01 (9.06)	0.80
EWR rawscore at time of post-test for pupils at schools not participating in NELI		712/ (0)	13.16 (8.79)	797/ (0)	14.55 (8.63)	-1.39
EWR rawscore at time of pre-test for pupils eligible for FSMs		211 (0)	2.56 (5.70)	225 (0)	2.34 (5.58)	0.22
EWR rawscore at time of pre-test for pupils not eligible for FSMs		1045 (0)	3.99 (6.48)	1058 (0)	4.45 (6.80)	-0.46
EWR rawscore at time of post-test for pupils eligible for FSMs		211 (0)	10.63 (9.24)	225 (0)	10.31 (8.64)	0.32
EWR rawscore at time of post-test for pupils not eligible for FSMs		1045 (0)	14.01 (8.84)	1058 (0)	14.74 (8.67)	-0.73

Imbalance at baseline

The absolute standardised difference between the baseline characteristics of the intervention and control groups was calculated to determine the likelihood that any of the observed differences between the two groups might bias the estimated impact of Flexible Phonics. An absolute standardised difference of ten or more indicates imbalance at baseline (Austin, 2009). As the differences were very similar for the randomisation and analysis samples, the results for the randomisation sample are reported in Appendix K, Appendix table. Table 8 assesses the scale of any differences between the key observed characteristics of pupils in the intervention and control groups in the final analysis sample.

The only marked difference between the intervention and control groups was in the number of pupils participating in NELI. As pupils in the intervention group were more likely to receive support from NELI as well as Flexible Phonics, there is a risk that any differences in outcomes between the intervention and control groups are due to NELI rather than Flexible Phonics. This possibility is explored in the subgroup analysis.

Table 8: Absolute standardised differences in baseline characteristics at pupil level—analysis sample

Pupil-level categorical variables:	Intervention group mean	Control group mean	Absolute standardised difference
Female	49.84%	48.32%	3.03
Low-ability pupils	51.04%	47.08%	7.92
Participating in NELI	12.82%	8.89%	12.67
Eligible for Free School Meals	16.8%	17.5%	1.96
Pupil-level continuous variables:			
Age in months at time of pre-test	56.22	56.30	2.19
Age in months at time of post-test	63.55	63.53	0.45
EWR rawscore at time of pre-test	3.75	4.08	4.95
Standardised pre-test for secondary outcomes	-0.03	0.12	8.82

** Statistically significant at the 5% level; *** statistically significant at the 1% level.

Outcomes and analysis

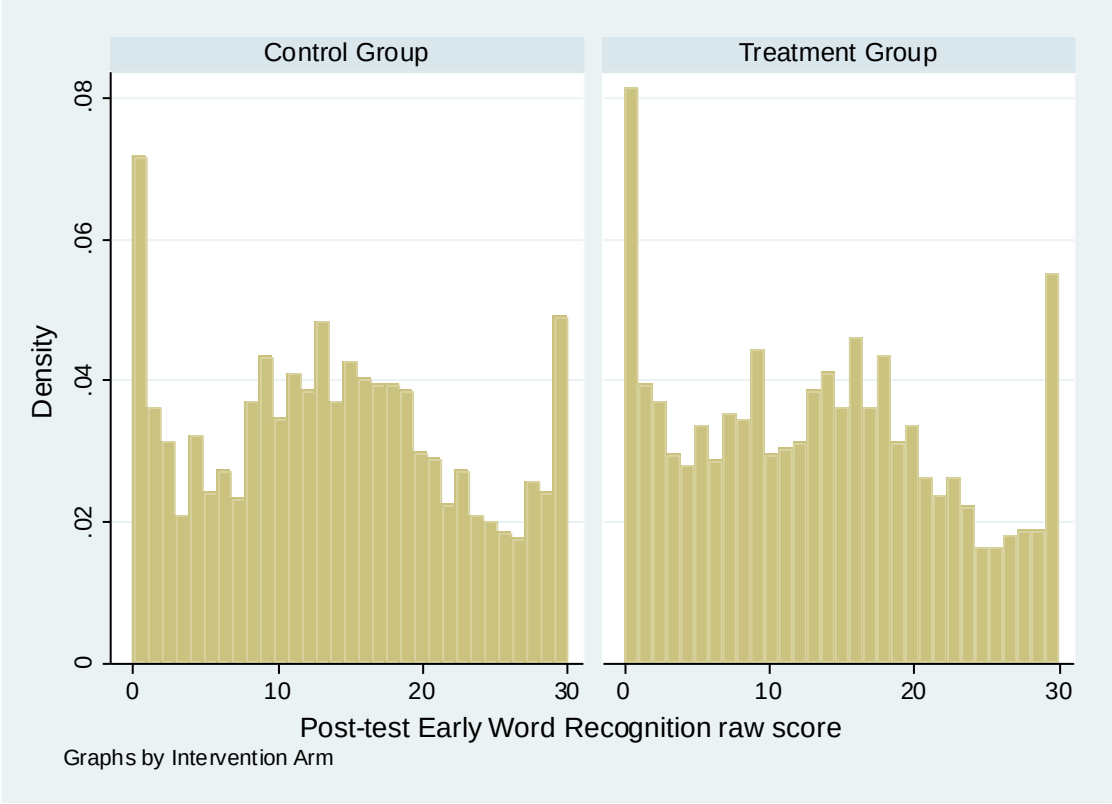
The results of the primary analysis are reported in Table 9. This shows that pupils in the intervention group had a slightly lower score on the primary outcome of Early Word Recognition (EWR) than the control group, before adjusting for the pre-test score. Figure 3 shows that the distribution of post-test scores was similar for the intervention (treatment) and control groups but even at the time of the post-test the most common score for either group was zero, with the maximum score of 30 being the second most common score.

When taking into account pre-test EWR scores, the intervention had an effect size of -0.05 standard deviations (SDs), equivalent to one month less progress. However, the Bayesian credibility intervals ranged from -0.2 to 0.1. This means that there is a 95% likelihood that the true impact of Flexible Phonics on the EWR score is between -0.2 and 0.1 SDs. It is therefore highly uncertain whether the impact of the intervention is positive or negative. The effect size was calculated using Hedges' g and the underlying parameters are reported in Appendix C Appendix table 2.

Table 9: Primary analysis

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)[27]
Early Word Recognition raw score	1,256 (0)	13.44 (-4.19, 31.07)	1,283 (0)	13.96 (-3.33, 31.26)	2,539 (1,256; 1,283)	-0.05 (-0.2, 0.1)

Figure 3: Histogram of post-test Early Word Recognition raw score, by trial arm



Secondary analysis

The results of the secondary analysis using the full YARC Early Word Reading standardised score are reported in the first row of Table 10. On average, pupils in the intervention group had a lower score on the YARC post-test than those in the control group before taking account of pre-test scores on the EWR and letter-sound knowledge subscales. Even after adjusting for the pre-test scores, the estimated effect of Flexible Phonics (calculated using Hedges' g) was to reduce the YARC score by 0.02 SDs. This was equivalent to zero months of additional progress, although there was a 95% likelihood that the true impact of the intervention was between -0.20 and 0.16 SDs. Again, this means that it is uncertain whether Flexible Phonics had a positive or negative impact on the YARC measure of Early Word Reading.

[27] Variance is presented in Appendix Table 2: Effect Size Estimation in Appendix C.

The set for variability strategy was expected to have an impact on performance in the Mispronunciation Correction Test (MCT). On this measure, the average post-test score for the intervention group was again lower than that of the control group. Even after adjusting for pre-test scores on the EWR and letter-sound knowledge measures, the estimated impact of the intervention was to reduce the MCT score by 0.04 SDs, equivalent to zero months of additional progress. However, once again there was a 95% chance that the true impact lay within the range of -0.24 to 0.16 SDs. As it was unclear whether the set for variability strategy reduced or improved performance in the MCT after controlling for pre-test performance, the impact of Flexible Phonics was ambiguous across all the primary and secondary outcome measures.

The distribution of post-test YARC and MCT scores for the intervention and control groups are shown in Appendix L and Appendix C. Appendix Table 2 reports the parameters underlying the calculation of Hedges' g.

Table 10: Secondary analysis

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)
YARC	1,256 (0)	-0.06 (-6.67, 6.54)	1,283 (0)	0.15 (-5.96, 6.27)	2,539 (1,256; 1,283)	-0.02 (-0.20, 0.16)
MCT	684 (0)	9.71 (-8.09, 27.52)	718 (0)	10.33 (-7.52, 28.18)	1,402 (684; 718)	-0.04 (-0.24, 0.16)

Analysis in the presence of non-compliance

As noted in the Methods section, all schools in the intervention group met the requirement of attending the Flexible Phonics training and, overall, two-thirds of schools (67%) were assessed as compliant on the binary measure of compliance with the Flexible Phonics programme, which took into account whether the intervention appeared to be delivered as intended as well as whether the school participated in Flexible Phonics training. This meant that 39 of the 58 schools in the intervention group analysis sample were judged to be compliant.

The Complier Average Causal Effect (CACE) analysis focuses on the primary outcome measure of EWR. It found that even when the programme was implemented as intended, it did not have a discernible impact on the primary outcome measure (see Appendix M for details). As the compliance measure was based on an assessment that teachers were delivering Flexible Phonics correctly, the CACE analysis provides evidence that the impact of Flexible Phonics on EWR is ambiguous even when the techniques are implemented as intended as outlined in the compliance measure agreed with the delivery team.

Missing data analysis

Table 5 in the subsection on Attrition summarised the pupil-level attrition for the primary outcome between treatment and control arms. Table 11 provides more detailed information on the availability of the pre- and post-test data for the primary outcome by treatment arm, while Table 12 shows missingness by intervention arm at the school level. As noted earlier, the level of attrition was slightly higher than expected when designing the trial (20% rather than 15%), but this is to be expected given the disruption caused by the pandemic, which could not have been foreseen when the protocol was drafted. Although 120 schools were randomised, two did not provide post-test data and thus were excluded from the analysis.

Table 11: Distribution of missingness, by trial arm, pupil level

	Intervention group	Control group	All
Total randomisation sample	1,567	1,599	3,166
EWR pre- and post- test score available	1,256	1,283	2,539
Missing either EWR pre or post test score	311	316	627
EWR pre-test score available	1,414	1,456	2,870
EWR pre-test score missing	153	143	296
EWR post-test score available	1,344	1,379	2,723
EWR post-test score missing	223	220	443

Table 12: Distribution of missingness, by trial arm, school level

	Intervention group	Control group	All
Total randomisation sample	59	61	120
EWR post-test score available	58	60	118
Missing either EWR pre- or post-test score	1	1	2

This section examines whether the missing data can be predicted based on other observed characteristics and, in particular, whether assignment to the intervention group is predictive of a missing post-test score. In line with the analysis imputing missing data, this section adopts a frequentist, rather than Bayesian, approach. Table 13 presents the results of probit regressions in which the dependent variable is a binary indicator of whether the post-test score on the primary outcome is missing. This indicator is regressed on pupil age, gender, free school meal eligibility, and intervention group. There was no statistically significant association between being in the intervention group and having a missing post-test score.

However, Table 13 shows that being FSM-eligible increases the likelihood of having a missing post-test score. Columns 1–3 present the results using the randomisation sample, which includes schools that did not take part in the post-test but were randomised to the intervention and control group. Columns 4–6 present the results for the sample of schools used in the analysis. The results are broadly similar.

Table 13: Probit regression predicting the likelihood that the primary outcome measure (EWR raw score) is missing

	Randomisation sample			Analysis sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-test Early Word Recognition raw score	-0.013* (0.006)	-0.007 (0.006)	-0.005 (0.006)	-0.009 (0.005)	-0.003 (0.006)	-0.002 (0.006)
Female		0.015 (0.057)	0.002 (0.057)		0.014 (0.064)	-0.001 (0.065)
Eligible for FSM		0.268* (0.125)	0.250** (0.089)		0.199* (0.084)	0.198* (0.090)
Treated	-0.009 (0.139)	-0.033 (0.142)	-0.040 (0.146)	0.007 (0.088)	-0.017 (0.088)	0.011 (0.086)
Age in months (at pre-test)		-0.004 (0.009)	-0.003 (0.009)		-0.003 (0.010)	-0.003 (0.010)
Pupil with a low score		0.091 (0.122)	0.095 (0.109)		0.094 (0.089)	0.089 (0.090)
Percentage of FSM pupils at school			0.003 (0.007)			0.002 (0.004)
School size (headcount)			-0.000 (0.000)			0.000 (0.000)
School type controls			Yes			Yes
Ofsted rating controls			Yes			Yes
Constant term	Yes	Yes	Yes	Yes	Yes	Yes
N	2,858	2,851	2,798	2,809	2,802	2,749

* p < 0.05; ** p < 0.01; *** p < 0.001.
Standard errors in parentheses. Errors have been clustered at the school level.

The Stata `mi` suite of commands was used to impute missing values across all variables that contained missing observations. This was done using a multivariate normal distribution. The `mi` commands are not currently supported within a Bayesian framework and so an Ordinary Least Squares (OLS) regression was estimated for both the primary outcome of EWR and the secondary outcome of the full YARC. As noted earlier, this was a deviation from the SAP. Only the pre-test score relevant to either the primary or the secondary outcome was included as a control variable, in line with the approach taken in the main analysis.[28] This OLS regression was undertaken purely for comparison purposes between the OLS main analysis sample and the analysis using the imputed sample.

The results presented in Table 14 show that, for the primary outcome, the analysis sample coefficient for the intervention group was very similar to the coefficient when using the sample with imputed missing values. In both cases, the intervention did not have a statistically significant impact on the primary outcome measure, in line with the findings of

[28] The analysis imputing missings was not carried out for the MCT as the test was only administered to a smaller subset of pupils than either the EWR subscale or the full YARC.

the Bayesian analysis. For the full YARC, the coefficient for the intervention group when imputing missing values was seven times as large as that seen in the main analysis sample. This large difference suggests that the subset of pupils with missing data performed very differently on the full YARC than those in the analysis sample. However, the coefficient was not statistically significant in either case. Again, this was consistent with the findings from the Bayesian analysis.

Table 14: Comparison of OLS final analysis sample regression with OLS regression using dataset of imputed values

	Primary outcome - EWR		Secondary outcome – Full YARC	
	OLS analysis sample (1)	OLS imputed missings (2)	OLS analysis sample (1)	OLS imputed missings (2)
Intervention arm	-0.247	-0.222	-0.040	-0.283
	(0.528)	(0.469)	(0.231)	(0.201)
Pre-test Early Word Recognition raw score	0.853***	0.864***		
	(0.025)	(0.025)		
Baseline for secondary outcomes			1.116***	1.134***
			(0.042)	(0.041)
Constant	10.486***	10.326***	0.014	-0.019
	(0.382)	(0.364)	(0.149)	(0.133)
N	2539	3166	2539	3166

* p < 0.05; ** p < 0.01; *** p < 0.001.
Standard errors are clustered at the school level. Standard errors in brackets.

Subgroup analyses

This section explores the impact of Flexible Phonics on three groups of pupils:

1. those with a score below the median on the combined pre-test standardised EWR and LSK subscales, referred to as low-ability pupils;
2. pupils at schools which did not participate in the Nuffield Early Language Intervention (NELI) alongside Flexible Phonics; and
3. pupils eligible for free school meals; this uses a measure from the National Pupil Database, which records whether pupils have been FSM-eligible at any point in the past six years.

RQ5 considered whether Flexible Phonics improved word reading ability differentially for FSM children while RQ6 sought to explore whether the impact of Flexible Phonics differed depending on the prior word reading ability of pupils. RQ8 was added to the list of formal research questions when it became apparent that a large proportion of schools participating in the trial were taking part in NELI under the government's Covid-19 support strategy. Table 15 shows the results of the analysis, which estimated the impact of Flexible Phonics on EWR for pupils who achieved a score below the median on the combined EWR and LSK pre-test. The effect size calculation (Hedges' g) indicated that Flexible Phonics reduced the primary outcome measure by 0.08 SDs for pupils with a low pre-test score. However, there was a 95% likelihood that the true impact was between -0.28 and 0.11. This means that it did not have a clear positive or negative impact for this subgroup.

Table 15: Primary outcome, pupils assessed as having a low score on the EWR and LSK pre-test

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)
Early Word Recognition raw score	641 (0)	8.09 (-5.29, 21.47)	604 (0)	8.72 (-4.89, 22.33)	1,245 (641; 604)	-0.08 (-0.28, 0.11)

Further analysis was carried out to assess whether there was a clear difference in the impact of Flexible Phonics between low-ability pupils and those who achieved a score that was median or above on the pre-test. This is reported in Table 16. Unsurprisingly, pupils in the control group who achieved a score below the median on the pre-test performed worse on the primary outcome measure than those who achieved a higher score on the pre-test. This is indicated by the fact that both the lower and upper 95% credibility figures for the low scoring pupil variable are below zero. The confidence intervals around the interaction term show that Flexible Phonics did not appear to differ in effectiveness depending on whether pupils were rated as having a low score or a higher score, based on their performance in the pre-test. The Flexible Phonics intervention does not appear to improve word reading ability differentially for children who had a lower score at pre-test.

Table 16: Regression results, interacting treatment and low-ability pupil status for the primary outcome

	EWR			
	Coefficient	Standard deviation	Lower bound of 95% CI	Upper bound of 95% CI
Treatment	0.241	0.578	-0.896	1.376
Pupil with a low score	-6.034	0.395	-6.809	-5.261
Interaction between treatment and low-ability pupil	-0.750	0.512	-1.753	0.252
N	2,539			

Each column shows selected coefficients from a regression of the outcome on treatment arm, low-ability pupil, treatment*low-ability pupil, and the EWR pre-test score. The analysis also accounts for the clustering of pupils within schools.

A similar number of intervention and control group schools had some pupils participating in NELI (26 and 24 schools, respectively). However, there were differences between the arms in the numbers of pupils who received support from NELI. Table 17 shows that among the subset of pupils at schools where no pupils participated in NELI, participation in Flexible Phonics was associated with a reduction in the primary outcome measure of 0.18 SDs. However, there was a 95% chance that the true impact of Flexible Phonics lay between -0.39 and 0.03 for this subgroup.

Table 17: Primary outcome, pupils at schools not participating in NELI

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)
Early Word Recognition raw score	712	13.16 (-4.07, 30.40)	797	14.55 (-2.37, 31.46)	1,509 (712, 797)	-0.18 (-0.39, .03)

To assess whether the impact of Flexible Phonics differed depending on whether the school had any pupils participating in NELI, further analysis was carried out using an interaction between whether any pupils at the school took part in NELI and whether the school was part of the intervention group. This is reported in Table 18. The 95% credibility intervals spanned zero and so it is uncertain whether Flexible Phonics was more or less effective in schools where some pupils participated in NELI. While the 95% credibility intervals reported in Table 17 also spanned zero, the analysis provided marginal evidence that Flexible Phonics was more effective in schools which participated in NELI as the lower bound was very close to zero. This suggests that perhaps other schools in the intervention group would have benefited from the additional catch-up support offered by NELI. Had this been available, it is possible that the Flexible Phonics programme would have been more effective.

Table 18: Regression results, interacting treatment and NELI status for the primary outcome

	EWR			
	Coefficient	Standard deviation	Lower bound of 95% CI	Upper bound of 95% CI
Treatment	-1.205	0.701	-2.581	0.171
School participating in NELI	-1.347	0.765	-2.850	0.151
Interaction between treatment and NELI participation	2.079	1.083	-0.048	4.205
N	2,539			

Each column shows selected coefficients from a regression of the outcome on treatment arm, whether the school had any pupils participating in NELI, treatment*participation in NELI, and the EWR pre-test score. The analysis also accounts for the clustering of pupils within schools.

Table 19 reports the impact of the intervention on the primary outcome measure of the Early Word Recognition raw score for the subset of FSM pupils. For these pupils, Flexible Phonics raised performance in the EWR test by 0.02 SDs but, again, the credibility intervals spanned zero meaning that there was a 95% chance that the true impact of Flexible Phonics was between -0.21 and 0.25 SDs.

Table 19: FSM subgroup analysis

Outcome	Unadjusted means					Effect size
	Intervention group		Control group			
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges' g (95% CI)
EWR: FSM subgroup	211 (0)	10.63 (-7.48, 28.74)	225 (0)	10.31 (-6.62, 27.24)	436 (211; 225)	0.02 (-0.21, 0.25)

The final table, Table 20, shows the coefficient and standard error for a version of the FSM subgroup analysis that included an interaction term to capture the additional impact of Flexible Phonics on EWR for those eligible for FSM compared to those ineligible and assigned to the control group. The fact that both the lower and upper bounds on the 95% credibility intervals for the term which captured pupils eligible for FSM but who were assigned to the control group were both negative reflects the fact that FSM pupils generally performed worse on the EWR test than ineligible pupils. However, as the 95% credibility intervals for the interaction term (Treatment x Eligible for FSM) span zero, the intervention was no more effective for FSM pupils than for other pupils.

Table 20: FSM subgroup analysis with interaction term

	Early Word Recognition raw score		95% credibility intervals	
	Coefficient	Standard deviation	Lower bound	Upper bound
Treatment	-0.445	0.536	-1.496	0.608
Eligible for FSM	-2.312	0.501	-3.291	-1.331
Interaction between treatment and eligible for FSM	0.457	0.721	-0.954	1.869
N	2,539			

Additional analyses and robustness checks

The analyses of the primary and secondary outcome measures were repeated excluding pre-test scores and are reported in Table 21. In all cases, the 95% credibility intervals spanned zero and the estimated effect sizes were very small, indicating that it was unclear whether Flexible Phonics had a positive or negative effect on each of the outcome measures. Irrespective of whether pupil performance prior to the intervention was taken into account, Flexible Phonics did not have a clear impact on any of the outcome measures it was expected to affect.

Table 21: Primary and secondary analysis excluding pre-test scores

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)
Early Word Recognition raw score	1,256	13.44 (-4.19, 31.07)	1,283	13.96 (-3.33, 31.26)	2,539 (1,256; 1,283)	-0.04 (-0.16, 0.08)
YARC	1,256	-0.06 (-6.67, 6.54)	1,283	0.15 (-5.96, 6.27)	2,539 (1,256; 1,283)	-0.02 (-0.16, 0.12)
MCT	684	9.71 (-8.09, 27.52)	718	10.33 (-7.52, 28.18)	1402 (684; 718)	-0.04 (-0.22, 0.14)

For the secondary outcomes (YARC and MCT tests), an additional analysis was conducted using the pre-test early word recognition raw score as the measure of prior attainment, as opposed to the composite baseline measure used in the main analysis. Table 22 shows the results of this exercise, which are in line with the main analysis. The estimated effects were small with a Hedges g of -0.07 in the case of the YARC and -0.06 for the MCT. The 95% credibility intervals spanned zero, indicating that it is unclear whether Flexible Phonics had a positive or negative effect on the secondary outcome measures.

Table 22: Secondary outcome analysis using the pre-test EWR score as control

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)
YARC	1,256	-0.06 (-6.67, 6.54)	1,283	0.15 (-5.96, 6.27)	2,539 (1,256; 1,283)	-0.07 (-0.23, 0.10)
MCT	684	9.71 (-8.09, 27.52)	718	10.33 (-7.52, 28.18)	1,402 (684; 718)	-0.06 (-0.25, 0.14)

The analysis of the impact of Flexible Phonics on the primary outcome was repeated excluding pupils at schools where the pre-test was conducted remotely to explore whether the apparent impact of Flexible Phonics varied with the method of testing. As shown in Table 7, a smaller percentage of schools in the intervention group than the control group took part in remote testing at the time of the pre-test (12% and 17%, respectively) and so it is possible that differences in the testing approach might have a bearing on the effectiveness of Flexible Phonics. This possibility is explored in Table 23. The effect size (Hedges' g) suggested that Flexible Phonics reduced performance on the primary outcome measure by a greater extent (0.12 SDs) for the subset of pupils who took part in face to face testing than for the wider analysis sample, although, once again, the 95% credibility intervals spanned zero. There was a 95% likelihood that the true impact of Flexible Phonics on pupils who took part in face to face testing was between -0.28 and 0.05 SDs.

One potential explanation for the stronger negative impact related to face to face testing is that remote testing in the pre-test phase might be a proxy for other actions to seek to reduce the risk of Covid-19 interrupting schooling. If schools that took part in face to face testing had higher absenteeism rates than schools that requested remote testing, this might explain these findings.

Table 23: Primary analysis for pupils who took part in face to face testing

	Unadjusted means				Effect size	
	Intervention group		Control group			
Outcome	n (missing)	Mean (95% CI)	N (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)
Early Word Recognition raw score	1,102	13.14 (-4.57, 30.85)	1,077	14.34 (-3.08, 31.75)	(1,102; 1,077)	-0.12 (-0.28, 0.05)

The original intention was to use a path analysis to assess the contribution of the direct mapping and set for variability strategies to the overall impact of Flexible Phonics on the primary outcome measure. However, as there was no causal pathway between the Flexible Phonics strategies and either the primary or secondary outcome measures, the assumptions underlying path analysis could not be met.

Implementation and process evaluation results

This chapter presents findings from the implementation and process evaluation strand of the research. Evidence from qualitative case studies with schools in the intervention group, interviews with the delivery team, surveys with teachers and TAs, and observations of training and follow-up sessions are used to explore compliance, fidelity, roll-out, business as usual approaches, and perceived outcomes.

Compliance

This section explores teacher and TA attendance and engagement with the Flexible Phonics training (compliance) as well as their views on the training content and format. It also explores engagement with, and views on, the three follow-up sessions that Flexible Phonics support partners conducted with schools in order to record which parts of the Flexible Phonics programme were being delivered, as well as offering support with understanding the programme and ideas for delivery.

This section focuses on the following research questions from the implementation and process evaluation aspect of the study.

- IPE4: Do teachers or TAs teaching reception receive all intended training?
- IPE6: How well is initial training and follow-up support received by teachers, TAs, and senior leadership at the school?
- IPE7: Is it necessary to conduct cascading training and has this been monitored and supported?
- IPE15: Do teachers and TAs engage well?

We will discuss the training first and then the follow-up sessions and support resources.

Training

Training attendance

In the endline survey, intervention group staff were asked whether they received the intended Flexible Phonics training. Most staff surveyed attended all three training sessions (87%); just 4% reported that they received no training, 8% had attended some but not all training sessions, and 1% attended some training sessions and received a separate catch-up session (total N = 121). Training attendance data provided by the delivery team showed that 304 individual staff (98%) either attended training in person or were sent catch-up videos to watch and seven staff (2%) did not attend any training or receive a catch-up video (total N = 311). Sixty-four percent of those who attended or watched a video of the training (195 individuals) were teachers or had similar roles such as assistant heads, phonics or EYFS leads, and reception teachers, and 36% (109) were TAs or similar roles, such as high-level teaching assistants, learning support partners, SEN support partners, and EY practitioners (total N = 304). Qualitative interviews with teaching and senior management team (SMT) staff found that, in most cases, staff did not attend the training because of time constraints, including one school which could not release any of its staff from teaching to attend the live training. Most staff who did not attend the training sessions reported watching recordings of all three sessions to catch-up, and staff members from one school also attended a follow-up session with Professor Rob Savage. Attendance information provided by the delivery team confirmed that staff at three schools had been unable to attend the online training and were offered a follow-up session with Professor Savage but only one school took up this offer. Qualitative interviews revealed that at one case study school, two TAs were not given time to catch up on the training so they mainly learned how to deliver Flexible Phonics through watching trained staff delivering sessions.

Engagement with the training

Views on the initial training were explored through qualitative interviews with teaching staff and SMT members at eight case study schools. Overall, staff received the training very positively. Many interviewed felt that the training was high quality, describing it as engaging, interesting and informative. One SMT member described feedback from their reception staff:

> *'Staff were really complimentary about the training and said it was really good to have … The training was engaging and well-presented and people were really passionate about phonics, which they really enjoyed' (SMT, School 85).*

Staff particularly appreciated the opportunities to contribute to discussions and engage in peer learning during the training sessions, which were provided both through break-out groups and whole group discussions. Many interviewees highlighted this as one of the key strengths of the training. During observations of training, evaluation team researchers observed attendees participating in break-out discussions, asking questions when given opportunities, and contributing to interactive activities. One reception teacher commented:

> *'Also, when we had that time, like five or less minutes to have a chat with some people, we could also share some ideas with no experience, but to have other people's view, that was really nice, I really liked that' (teaching assistant, School 118).*

Other staff highlighted the benefits of having a chance to reflect on what they had learned and, in later training sessions, how delivery was progressing. The online training included break-out rooms with staff from different schools to discuss their thoughts on the programme, their ideas for how they might implement it, and, if relevant, their experience of trying programme activities or strategies. However, one reception teacher reported that some people did not engage in break-out groups, which limited their usefulness. This was also observed by researchers who attended the training sessions. While for the most part there was a good level of discussion in break-out rooms, some training groups were more engaged while others were quieter, and in at least one case no attendees contributed and all had their cameras switched off. In interviews with the delivery team, the project director noted that some training cohorts were more active than others but felt that there was learning across all the cohorts. The project director also highlighted the role of using software (Mentimeter), which allowed participants to ask questions or make comments anonymously without feeling self-conscious, for helping get a sense of a group's needs or interests. However, the project manager expressed concern that TAs were sometimes less vocal in the breakout groups and possibly felt less empowered to contribute.

Understanding Flexible Phonics theory and delivery

Most staff felt that they had a good understanding of the programme from the initial training sessions, follow-up support, and the manual and felt confident in delivering the programme. In a training session observed by researchers, attendees were asked to rank their confidence in delivering the programme on a scale of one to five, with one being 'not at all confident' and five 'extremely confident'. During the second training session, the group responded with an average of 3, rising to 3.3 by the final session, indicating a modest rise in the level of confidence among attendees. Another cohort exhibited an increase from 2 in session two, up to 3.3 in session three, demonstrating the progress made. The training took place during the partial closure of schools so most schools were unable to try out delivering elements of the programme with their class between training as the delivery team originally intended so this may have reduced overall confidence levels. Findings from the endline survey confirmed that most teachers and TAs felt they had a good understanding of the programme: the majority agreed that they had a good understanding of the programme (93%, Total N = 116) including aspects such as direct mapping (85%, total N = 115), set for variability (88%, total N = 115), and strategies to use with children struggling with general phonics (87%, total N = 116).

A few teachers and TAs interviewed appreciated that the training covered the theory behind Flexible Phonics. These respondents said that the training updated their phonics knowledge and was a helpful refresher on phonics theory. Similarly, the majority of teachers and TAs who responded to the endline survey agreed that they had an understanding of how Flexible Phonics supported language processing during reading (89%, total N = 115). One staff member interviewed from a school that later withdrew reported that the training had given them useful ideas for teaching phonics. However, a small number of respondents responded negatively to the theoretical elements of the training. One phonics co-ordinator said that they found the technical terms hard to follow at first but after reviewing the handbook they were able to follow the rest of the training more easily.

Most teachers and TAs who answered the endline survey agreed that, after the training, they had felt ready to start planning ways to incorporate Flexible Phonics strategies into their phonics teaching (86%, total N = 115) and ready to start teaching the Flexible Phonics strategies (80%, total N = 116). In case study interviews, one teaching assistant reported that the examples and roleplays demonstrating Flexible Phonics strategies were particularly helpful in planning and delivering the programme and many staff who were interviewed highlighted the Flexible Phonics manual as being key to having a good understanding of the programme. One phonics co-ordinator reported that the manual helped them

to understand Flexible Phonics terminology, which in turn helped them to understand the training sessions. A reception teacher at another school said that the manual helped them to feel confident delivering Flexible Phonics as it provided a clear structure for delivery, which was consolidated by the training sessions and follow-up support. However, one delivery team support partner thought that a lot of the teachers and TAs they worked with had not read the manual. Overall, SMT members also reported that their staff were confident and ready to deliver Flexible Phonics after the training, without needing additional support from them. One TA reported that 'from the first session I felt prepared, and I could go in and do a lesson' (TA, School 110). One SMT interviewee particularly appreciated that staff could contact UCL with any questions following the training.

However, a few members of staff reported having a weaker understanding of Flexible Phonics. However, two of these interviewees were TAs who were not given time to watch the training sessions. A teacher at one school was unclear on whether the vocabulary element of the programme was a core, compulsory element of the programme or not, and a TA at another school felt confused about the mispronunciation-correction strategy. One phonics co-ordinator reported being a bit nervous about delivery at first as there were lots of elements of the programme to understand, but the coordinator felt that once they started delivery it went well.

Suggested improvements

Just one interviewee provided solely negative feedback on the training. They felt that the training was too long and lacked real life examples of delivering Flexible Phonics strategies:

> *'I didn't think it was the best training I'd ever been on. I thought it was very, very lengthy … the practical examples weren't great. I thought there could have been more practical examples and some actual examples in a school setting' (EYFS lead, School 13).*

Practical demonstrations were not possible due to the training taking place during lockdown and roleplay examples were filmed with adults for the same reason. The delivery team hoped that these could be included in future delivery.

A few other interviewees felt that the training was somewhat long and that there was too much content to take in, especially those who could not attend live and watched recordings. However, one teacher thought that the length of the training was necessary to provide a good understanding of the Flexible Phonics programme. Another reception teacher suggested delivering the training in smaller chunks to make it more accessible, and two interviewees suggested including less theory. One teaching assistant felt that the session on planning was not relevant to their role, so they would have preferred not to attend it.

Several interviewees also said that they would have liked more practical examples of Flexible Phonics strategies and examples including children, which was not possible as schools were partially closed during the lockdown. A few interviewees said they would have preferred the training to be delivered in person as they felt that this would have improved opportunities for peer learning. One teaching assistant reported that the break-out rooms did not always work well as the groups were small and sometimes did not include teaching staff, and some attendees did not contribute to group discussions. One reception teacher felt it would have been better to deliver the training when schools were open so that attendees could have delivered Flexible Phonics between sessions as intended:

> *'It was during lockdown so there wasn't actually a chance to have a go with the children being in class and then discuss and share that with other schools that are taking part in the project' (reception teacher, School 14R).*

However, a few other interviewees felt that the remote delivery had worked well as staff members who did not participate in the training live could watch the recordings, although teachers at two schools commented that it was quite a lot of video to watch. From researcher observations of the training, attendees and hosts appeared comfortable with Zoom, and technical issues were minor (for example, problems with feedback or people being removed from break-out rooms prematurely) and were addressed quickly. Although the original plan was for the training to be delivered face to face, the delivery team were considering using a mix of face to face and online training for future delivery at scale as they felt this had worked well.

Cascading training

There was little evidence of schools cascading training as most teaching staff either attended the training or watched the videos. In case study interviews, at one school where TAs were unable to attend either the training or a catch-up session, there was some informal cascading of training through the form of TAs observing Flexible Phonics sessions delivered by other staff at the school, however, this was not monitored or supported by the Flexible Phonics team. In another school, a reception teacher also shared in-depth lesson plans based on the training and manuals to support delivery.

However, there was evidence that some schools were considering cascading learning for teachers and TAs in other years, once the programme has finished. The project manager reported that 29 schools (a little under half of the control group) had requested training in Flexible Phonics for their Year 1 teachers and TAs as they felt it would be helpful to continue supporting this approach, and some highlighted that this cohort of children had been impacted twice by partial school closures during the Covid-19 pandemic. The project director felt it would be helpful to link up teaching in the two years, so they were planning to offer training for Year 1 teachers in the summer.[29] The project manager explained that the training would be broadly similar to the original training but the vocabulary taught would be updated to be appropriate to the Year 1 level. One support partner reported that several schools she was working with had told her that they had been thinking of doing their own training for teachers in other year groups, including up to Years 4 or 5, as they felt that it would be beneficial for all staff to have a background understanding of the approach.

Follow-up support sessions

The Flexible Phonics training also included three follow-up support sessions delivered by support partners from the UCL Flexible Phonics team. The sessions were scheduled so that the first support session occurred towards the end of the training sessions (late January or early February), the second session occurred around Easter, and the third session took place in the summer term. The proposed format for the follow-up sessions were that, for the first session, there would be individual support sessions for the teacher(s) and TA(s) for each reception class and then the second and the third sessions would be plenary sessions with all teachers and TAs. Another support partner noted that while earlier sessions might be attended by a range of teachers and TAs from reception classes, in some cases, this reduced to just the teacher(s) and TA(s) attending from the class that was participating in the trial.

Follow-up sessions attendance

Compared to the initial training sessions, a smaller majority of staff surveyed (68%) reported attending the follow-up support sessions. In case study interviews, the small number of staff who did not attend the follow-up sessions were TAs who could not attend due to not being given time out of teaching to join the sessions, where other staff members from the school had attended the follow-up support. Most of the delivery team support partners interviewed and the project manager observed a tendency for teachers to be more engaged, with TA attendance dropping over the sessions for some schools, and teachers being more likely to speak unless the TAs present were asked a question directly. However, one support partner reported arranging separate meetings with TA groups where they found the TAs to be engaged with the programme and contributing ideas. This was an option offered by all the support partners. One support partner noted that morning appointments were preferred and helped with attendance and that attendance was impacted by issues with staff being ill or overwhelmed—it is likely that this has been more of an issue during the Covid-19 pandemic. Some support partners observed that schools with one-form entry sometimes struggled to release staff and one support partner reported receiving more queries from staff at one-form entry schools. In one school, the support partner was only able to meet with the early years lead and not the class teacher, and staff at another school said after the second session that they did not need another meeting although the support partner was welcome to get in touch via email, and so forth.

Engagement with follow-up sessions

The survey also explored how teachers and TAs used these follow-up sessions by presenting them with a list of possible topics discussed in the sessions and then asking them to pick all the ways in which they had used the sessions. The most common way that staff used the sessions was to ask for clarification on aspects of the programme that they were

[29] The two-part Year 1 training was held on 14 and 15 July 2022 (3.30–5.00 pm). Overall, 29 of the 59 intervention schools took part, with 67 staff attending Part 1 and 58 staff attending Part 2.

unsure about (59%) followed by getting feedback on their ideas or plans for implementation (51%); others mentioned sharing approaches, resources, or plans that were working well in their school (46%). The least common ways that staff used the sessions were to access additional support or resources for a topic or aspect of the programme (17%) or to access support with planning their delivery (17%).

In case study interviews, staff described having broad discussions in the sessions about how delivery was progressing, with many receiving advice from UCL staff on how to deliver Flexible Phonics, including suggested activities. The follow-up sessions also acted as a means for schools to share resources, with the UCL support partners passing on resources and advice on good practice that schools had shared with them. For example, one school described sharing lesson plans with their support partner who then circulated these to other schools. Teaching staff also asked a range of clarifying questions around Flexible Phonics strategies including how to combine the suggested GPC order with existing phonics programmes and how to teach exception words and blending exercises.

Views on follow-up support sessions

Case study interviews found that most staff received these sessions positively and many reported that they were reassuring and supportive. Staff at a one-form entry school particularly appreciated when support partners told them about resources shared by other schools. Teachers and TAs at a few schools reported that they did not need much support or feedback and did not have many questions for the support partners, so sessions were used to talk over their approach to delivery with support staff asking questions and providing feedback. One teacher felt that the sessions should have focused on building skills rather than checking in, and another thought that the sessions could have been used to help prepare schools for delivering the programme in September in future.

Although support partners offered appointments during the school day to facilitate attendance from teachers and TAs, some schools chose for support sessions to take place outside of teaching hours. One teaching assistant reported that they had attended support sessions outside their working hours, meaning it was unpaid—TA contracts often cover teaching hours only. No other staff reported having to attend during unpaid hours but two TAs at one school reported that they were not given time to attend the sessions and, at another school, one nursery nurse also did not attend the sessions, which suggests that some schools may have struggled to make time for TAs or other staff to attend during contracted hours.

Views on Flexible Phonics support resources

School staff also had access to the UCLeXtend online platform, where schools could access resources and information to support delivery, interact with staff at other schools via message boards to discuss delivery, and share resources as a form of peer learning and support. Schools could also access ad hoc support from the delivery team via phone or email.

Engagement with support resources

In case study interviews, a few members of staff reported using UCLeXtend to access resources shared by other schools, including activities and slides, but none had uploaded resources to share with other schools. One reception teacher described how they had tried to access UCLeXtend but had found the user interface unintuitive and difficult to navigate, so they asked their allocated support partner to email the resources instead. Although there appeared to be a strong appetite for peer learning generally among staff interviewed and an interest in what the other Flexible Phonics schools were doing, most staff reported sharing resources via their allocated support partner or reading the Flexible Phonics newsletter rather than using the forum for discussion or posting their own resources to UCLeXtend. One school reported sharing learning in a school cluster group that they organised outside of the Flexible Phonics trial group. Several delivery team support partners noted that teachers and TAs would ask them to email them resources directly rather than accessing the platform themselves, and one support partner confessed to struggling to navigate it themselves. Just one staff member reported contacting the team via email for additional support regarding the direct mapping books not matching with key sounds. In interviews, several delivery team support partners commented that there was some email correspondence initially with schools earlier in the year, but that schools became less responsive once the partial school closures ended.

The endline survey also explored teacher and TA views on the support resources. The vast majority (95%) reported that the manual was helpful in supporting them to deliver the intervention; 53% said it was 'very helpful' (total N = 119). With regard to ad hoc support via email or phone, 71% felt that was helpful 'to some extent', one said the support was not helpful at all, and 22% had not accessed it (total N = 117). The UCLeXtend platform was found to be helpful 'to some extent' by 63% of respondents; 3% said it was not helpful at all and 22% had not used it (total N = 117). Finally, the regular newsletter was found helpful 'to some extent' by 71% although 5% felt they were not helpful at all (Total N = 117).

Fidelity

This section explores how the Flexible Phonics programme was delivered in the intervention group schools (fidelity), their views on the programme and course materials, and any barriers or enablers to delivery. This section focuses on the following research questions from the implementation and process evaluation.

Delivery

- IPE1. Are schools delivering the intervention and the trial as intended?
- IPE5. How often do participating teachers and TAs deliver Flexible Phonics strategies in phonics teaching?
- IPE8. How effectively do teachers/TAs use Flexible Phonics strategies?
- IPE13. Do all intended pupils receive Flexible Phonics teaching?
- IPE14. Do some pupils receive more Flexible Phonics teaching than others?
- IPE18. How does the intervention enhance/differ from existing phonics teaching
- IPE22. Have schools adapted the intervention—how and why?

Facilitators and barriers

- IPE9. What facilitates/hinders effective implementation?
- IPE10. Would teachers/TAs find additional support helpful in maintaining quality—what and from whom?
- IPE11. Are there unintended or negative effects of the intervention?
- IPE16. Is the intervention acceptable and practicable in schools' contexts?

Delivery

In the context of this study, the Flexible Phonics intervention was designed as an add-on component to be integrated with a school's existing phonics delivery, as opposed to functioning as a stand-alone phonics programme. The intervention comprised two key approaches or strategies for building phonics knowledge and confidence: direct mapping and set for variability. These two main strategies were incorporated into a five-strand approach which covered phonics knowledge necessary to be able to introduce the strategies, how to introduce the strategies, and some approaches to use with children who are struggling with phonics. The five strands of the Flexible Phonics intervention were:

1. GPCs and direct mapping—using high quality children's books to reenforce targeted phonics learning;
2. set for variability—introducing a variable consonant or vowel strategy and teaching the variability principle further through set for variability oral games such as 'Simon Says' where you deliberately mispronounce words so that children have to work out what they should be;
3. teaching vocabulary—teaching 66 key exception words as part of children's spoken vocabulary that help them make sense of texts so that they can start to use the set for variability approach when reading these;

4. mispronunciation correction—learning to read new exception words using all of the previous strategies and reflecting on what the word could be by thinking about words they know that sound sort of similar or working it out from sentence or picture context; and

5. support for the struggling readers—introducing strategies to help children who are struggling with earlier stages of phonics such as mapping sounds to letters (learning GPCs) or struggling with blending sounds after sounding out each letter.

A school would not necessarily be expected to deliver all five strands but to select those that are appropriate for the needs of the children being taught. For example, a class where children had a high level of language and reading may not need to use any strategies for the struggling readers. The delivery team specified that in order to be compliant with the intervention, schools needed to deliver direct mapping and one or both of set for variability or mispronunciation correction strategies. While schools were aware that the key aims of Flexible Phonics were to incorporate direct mapping and set for variability strategies into their phonics teaching, they were not explicitly told the compliance criteria set out above as these were agreed for the purpose of the study.

Planning

In case study interviews, staff were asked what planning they had done prior to starting to deliver the programme. As schools were partially closed with fewer pupils attending in person during the training period (January to February 2021), schools chose to start delivering the intervention in March, although some started trying out some activities while the training was ongoing. Schools that started delivering Flexible Phonics in March could potentially use the training period to plan how they would deliver the intervention once the schools fully opened in March. Around half the schools held planning meetings prior to delivery, while in the other half, staff planned their delivery individually. Many staff members said that the Flexible Phonics manual made planning simple as it provided a clear approach to delivery. SMT had varying involvement in planning: some were involved in delivery so took part in individual or group planning; of those that were not involved, a few reported spending time planning for staffing and timetabling, while two SMT staff said that they gave their teaching staff autonomy over planning and delivery so had little involvement.

During qualitative interviews with the delivery team, the project manager and some support partners reported that a few schools needed quite a lot of reassurance initially about whether they could use Flexible Phonics—as the DfE recommends using phonics schemes that have been validated by the DfE—and that they had to explain that Flexible Phonics was an add-on to their existing phonics programme and was just some additional strategies. Similarly, the project director reported that during the training there had been a need to reassure some teachers and TAs that you could use real children's books for teaching phonics as the DfE was promoting the use of fully decodable books or other texts where children should be able to decode all words for their stage of phonics teaching.

Which elements of Flexible Phonics did schools deliver?

Teachers and TAs were asked in the endline survey to specify which elements of the Flexible Phonics programme they had delivered (see Table 24). The 'GPCs and direct mapping' strand was a core element of the Flexible Phonics intervention where learning of phonics sounds (grapheme-phoneme correspondences) is reinforced through reading children's books containing the sounds. Three-quarters of staff surveyed reported that they had delivered this (77%) component. The other core element was set for variability, which was delivered through the mispronunciation correction and set for variability activities. The set for variability strategies build on basic general phonics knowledge, so these activities tended to be introduced later in the delivery period. Encouragingly, two-thirds of teachers and TAs (65%) reported teaching mispronunciation correction and one-third (35%) had delivered the set for variability activities. These findings suggest that there was wide take-up among intervention schools of the key Flexible Phonics elements. There was also evidence that schools were using other activities in the programme to support and enrich their phonics teaching practice. Around two-thirds of staff had taught exception word vocabulary (66%), which aimed to equip children with the background vocabulary to decode exception words using set for variability strategies, and around two-fifths had delivered support for struggling readers (44%).

Table 24: Aspects of Flexible Phonics delivered

Flexible Phonics activity	Frequency	Percentage (%)
GPCs and direct mapping	88	77
Teaching vocabulary: exception words	75	66
A strategy for reading key exception words—mispronunciation correction	74	65
Support for the struggling readers	50	44
Set for variability	40	35
N =	122	

Multiple options could be selected so percentages in the table may add up to more than 100%.

Where staff had reported delivering an element, they were then asked how easy or difficult they had found it to deliver on a five-point scale ranging from 'very easy' to 'very difficult'. For each element, the majority said that it was either very or quite easy to deliver, and no staff reported that any of the elements were very difficult to deliver. The proportion reporting that an element was easy to deliver was highest for the core elements, 'GPCs and direct mapping' (83%, total N = 87) and set for variability (92%, total N = 39), followed by mispronunciation correction (81%, total N = 73), teaching vocabulary (79%, total N = 75), and support for struggling readers (74%, total N = 49).

The delivery of Flexible Phonics elements was explored further in qualitative case study interviews. Among the case study schools, all eight reported delivering the core element direct mapping, and seven were delivering at least one of the set for variability elements, for example, set for variability oral games or mispronunciation correction. However, one school did not teach the set for variability elements or flipping sounds because it was teaching sounds in the order specified by the Read Write Inc phonics programme and so would not teach alternate sounds for letters until Year 1. Staff at one school described delivering all elements of the project, with most staff members interviewed from other schools delivering some, but not all, elements. In line with the survey findings, direct mapping was the most commonly used strategy, with many staff members describing delivering direct mapping sessions, often on a daily basis. Some delivery team support partners interviewed felt this element was particularly valued where schools used scheme books such as Read Write Inc. However, one support partner felt that a lot of schools were already doing direct mapping as part of their usual practice but did not realise. Set for variability was also commonly used, with many staff interviewed discussing embedding the 'flipping sounds' strategy into the school day using activities like Simon Says, deliberately mispronouncing children's names, and playing the robot game. Similarly, some described embedding mispronunciation strategies for reading exception words into the school day. Teaching vocabulary and delivering support for struggling readers were used less frequently. Just one school described teaching the meaning of key exception words, which they delivered using pictures and sign language. One teacher at another school said that they would have liked to deliver teaching around vocabulary but felt that children were already familiar with the key exception words. Again, just one member of staff described delivering support for struggling readers to a lower ability group of pupils and did not elaborate on how this was delivered.

Most delivery team support partners interviewed commented that, for the most part, schools did not start implementing Flexible Phonics until schools reopened in March 2021, when the second follow-up sessions were starting. One support partner thought that some schools might have been choosing one or two elements to engage with first. Another observed that schools varied in how much guidance they needed, with some needing quite a lot at first. However, most support partners commented that confidence and awareness around delivery increased over time. In line with the survey findings, the project director noted that in a meeting towards the end of the school year, the delivery team found that approximately 60% to 70% of schools were delivering both the direct mapping element and one of the set for variability components of Flexible Phonics. It is possible that this proportion might have been slightly higher than this in a normal school year as some schools felt unable to deliver set for variability aspects because children were still learning general phonics due to the disruption to their learning with the Covid-19 pandemic. Very few schools were identified by the support partners as not engaging with the programme at all or not accepting any guidance or support with delivering the programme that was offered by the support partners.

Support with delivering Flexible Phonics

During case study interviews, staff were asked whether they would have liked any additional support with delivering Flexible Phonics. Just one member of staff interviewed had contacted the delivery team for additional support: a TA had emailed with a concern that the direct mapping books did not always match the designated sounds well but the delivery team were able to reassure them that the sounds were present, although their frequency in the text may be less than they might be used to with reading scheme books. However, in interviews, delivery team support partners reported needing to remind schools of the different elements of the programme to make sure that they were incorporating all the relevant elements and recognising what they were already doing.

Similarly, most teaching staff interviewed as part of the case studies reported receiving good support from their SMT. A few staff members highlighted that they were provided with additional resources for delivery. Other means of support mentioned included the SMT showing an interest in the programme, making time for Flexible Phonics in their timetable, and handling communication with UCL. However, a few TAs felt that they were not adequately supported to deliver the programme. One TA said that they received some guidance at an initial planning session but since then had been planning and delivering the programme by themselves and would have liked additional informal support from their reception teacher to improve their understanding of and confidence delivering the programme:

> *'I'm always by myself in a room with these children, so I had no idea if I was doing it right or wrong ... no one told me anything' (TA, School 14R).*

Two other TAs who were not given time to attend the training also felt that being given time to attend the training sessions, follow-up support, and internal planning sessions would have helped them to deliver the programme. A few members of SMT suggested that staff would have benefitted from practical examples of delivery activities, such as video examples in a classroom setting and examples from other schools of successful integration and delivery. One SMT member felt that after the training there was still a lot for staff to determine in order to apply the programme. One delivery team support partner commented that it was more difficult for teachers and TAs in one-form entry schools to implement Flexible Phonics as they had a lot of responsibilities, whereas teachers and TAs at multiple-form entry schools could work as a team sharing ideas and responsibilities.

How were schools delivering Flexible Phonics?

Qualitative interviews also explored how often schools delivered Flexible Phonics activities and how these fitted into the school day. Most case study schools reported that Flexible Phonics was delivered by a mix of reception teachers and TAs. The approaches to delivery described in case study interviews tended to differ on a school-by-school basis. In the Flexible Phonics ToC model, the delivery team specified that schools would be expected to deliver the programme three to four times a week. However, most schools described delivering Flexible Phonics on a daily basis, with two schools delivering on a weekly basis and one on an ad hoc basis, such as delivering direct mapping when they had time at the end of the day. Many schools discussed embedding Flexible Phonics activities throughout the school day, particularly set for variability activities such as the Simon Says game. Staff at these schools described using these activities in maths lessons, at breaktime, and when giving instructions and during transitions between lessons or activities. Games and activities seemed to be more commonly used than more formalised elements of the programme. Just one teacher interviewed reported introducing a daily exception word, and another teacher also described putting print-outs of the strategy for reading exception words on every table to encourage pupils to use the strategy outside of phonics lessons. One reception teacher explained:

> *'You needn't do every single thing in the programme in one lesson. It's more about embedding all of those little things for phonics into what we do' (reception teacher, School 56).*

A few of the case study schools delivered Flexible Phonics as a discrete daily lesson, with pupils streamed into ability groups, although some of these schools also embedded activities throughout the day. A small number of case study schools added Flexible Phonics activities into their usual phonics sessions, for example, using direct mapping at the end of the session focusing on the sound they had been working on as part of their usual phonics curriculum.
As part of the programme, the delivery team also encouraged schools to develop and share resources to create a community of practice. In qualitative interviews, most support partners were able to give examples of schools they worked with that had developed activities or resources that had then been shared more widely among the schools in the intervention group. Some support partners felt that teachers and TAs were quite modest and would not think to share a resource or idea until the support partner suggested that it could be useful for other schools too. One support partner

described how a school identified nursery rhymes that contained relevant sounds from the Flexible Phonics programme and created a list of these which could be shared more widely among schools as nursery rhymes are out of copyright, in contrast to children's books.

Remote delivery of Flexible Phonics during school partial closures

As the initial delivery period took place during lockdown when schools were closed, the survey also explored whether schools had delivered any Flexible Phonics activities remotely. The majority of respondents (61%) had only delivered Flexible Phonics activities face to face, just under a quarter had delivered activities using video conferencing software, 17% pre-recorded videos for parents or children to watch, and 1% live streamed activities. In the qualitative interviews, however, most case study schools reported delivering, or attempting to deliver, some activities remotely through video conferencing software as part of their remote teaching. Around half of case study schools that had attempted remote delivery undertook remote direct mapping with small groups of pupils. This included making digital copies of books in some cases. Staff had varying views on the effectiveness of remote delivery of direct mapping. One reception teacher felt that it had worked well, with pupils finding it engaging and enjoyable, and that it helped to familiarise pupils with the format of direct mapping:

> *'So we did sort of start it online, which worked fine actually, we sort of scanned each other's books, whilst some of us were in school, and we used them. And then … we used to do a sort of a 10 minute PSED calling for children to join in, and we'd use some of the, sort of the, active activities on that ... So that worked well' (SLT, School 77).*

Another teacher said that while the reading element worked well, it was harder to deliver the sentence writing element of direct mapping as this required more one to one support. However, one headteacher reported that direct mapping worked better in person as parents could disrupt the online sessions by not letting pupils work out answers for themselves. Similarly, half of the case study schools that delivered elements of the programme remotely described delivering set for variability and blending activities as part of their remote teaching. However, a couple of staff reported that it was hard to engage pupils in Flexible Phonics remotely, and that pupils were too distracted to engage.

Integrating Flexible Phonics with existing phonics teaching

Evidence from the endline survey and case study interviews suggested that, for the most part, it was relatively straightforward to integrate Flexible Phonics with existing phonics teaching. In the endline survey, the majority of teachers and TAs (58%) reported that this was easy, however, 17% felt that it had been difficult to integrate it (total N = 117). Further statistical analysis explored whether ease of integration was affected by the phonics programme schools were delivering and by whether schools were also participating in the Nuffield Early Language Intervention (NELI).[30] The analysis compared the answers of schools delivering Letters and Sounds, the most common phonics programme in the sample, and other phonics programmes. Mann-Whitney tests found no significant difference between those whose school used Letters and Sounds and those that did not with regard to ease of integrating with Flexible Phonics. Similarly, there was no significant difference between survey respondents whose school was also delivering the NELI intervention and those whose school was not with regard to ease of integrating with Flexible Phonics.

In case study interviews, staff at several schools reported that they did not have any issues with integrating the programme with their usual phonics approach. Two of these schools reported that they had already finished delivering their usual phonics programmes by the time they started delivering Flexible Phonics so there was little integration needed, and one EYFS lead said that as Flexible Phonics is designed to be embedded into normal teaching practice, it was easy to integrate with their usual approach:

> *'We've integrated it rather than created all new subjects or new topics … It is always part of our literacy or our phonics. So, we haven't kind of reinvented things to make space for it. We just immersed it within what we do already … We've just tweaked what we do to make space to fit it in' (EYFS Lead, School 13).*

[30] The Nuffield Early Language Intervention (NELI) is a spoken language intervention targeted at children in reception class who are struggling with spoken language.

However, most staff interviewed reported experiencing minor challenges with integrating Flexible Phonics. The following section explores how staff adapted the programme to the needs of their students and to fit with their existing phonics approach.

Adapting the Flexible Phonics programme

As previously discussed, the Flexible Phonics programme was designed to be flexible—to fit with the needs of children and school curricula. Both the endline survey and qualitative case study interviews explored whether schools had adapted the programme and whether this affected the fidelity of the intervention. The endline survey presented respondents with a list of possible adaptations and asked them to indicate all that they had used. Results suggest that many teachers adapted the programme to fit with their existing phonics programme or retained approaches from their existing phonics programme. The most common adaptations reported by around half or more of teachers and TAs surveyed were teaching sounds in a different order than suggested (58%), continuing to teach sight words with some pupils (50%), and using text other than that provided by UCL to deliver direct mapping (48%). Seven staff reported adapting the programme in a way that was not listed and were asked to provide detail on how they had done so. Among those who provided further detail, most reported using selected elements of Flexible Phonics alongside their usual phonics approach and one explained that they used Flexible Phonics strategies to teach vocabulary as well as phonics.

Table 25: Adaptations made to Flexible Phonics

Adaptation	Frequency	Percentage (%)
Taught the sounds in a different order than suggested, for example, used the order in your usual phonics programme such as Read Write Inc.	62	58
Continued teaching some sight words with some or all of your learners, for example, using the traditional flashcard ‘whole word’ method.	54	50
Used texts other than the books provided by UCL to undertake the direct mapping element of the programme.	52	48
Taught additional sounds which are not part of the suggested Flexible Phonics list.	43	40
Taught a different list of exception words than suggested, for example, used the exception words list from your usual phonics programme.	42	39
Taught Flexible Phonics alongside your existing phonics programme with a specific group of children and continued to use only your usual approach with the rest of the children.	31	29
Adapted it in another way.	7	7
N =	122	

Multiple options could be selected so percentages in the table may add up to more than 100%.

The case study interviews and further survey analysis provided insights into how and why schools were making these adaptations. The following sections explore this further.

Adapting the order of GPCs

In qualitative interviews, most case study schools reported that the grapheme-phoneme correspondence (GPC) order suggested by Flexible Phonics did not match the order followed in their existing phonics programme although some schools worked around this by adapting their Flexible Phonics teaching to use the GPC order specified by their existing programme. The case study schools that reported experiencing this issue were delivering Letter and Sounds, Ruth Miskin, and Read Write Inc phonics programmes. This issue was compounded by the fact that Flexible Phonics delivery began later in the academic year once schools had reopened fully and children’s phonics learning had been disrupted by the pandemic. Staff at one school reported introducing new sounds from the Flexible Phonics GPC order to prepare pupils to engage with the programme. However, many staff interviewed said that they continued to follow their normal GPC order, and staff at one school cited government guidance that ‘any resources used should exactly match the Grapheme Phoneme Correspondence (GPC) progression of their chosen [phonics] approach’ (DfE, 2022a). One SMT member explained:

> *'We're pretty passionate about our phonics schemes [being] taught in a systematic manner … and we feel really passionately that we should be sticking to the principles of the Rose Review. However, when staff went on training about Flexible Phonics, they were told that they should change the order that the sounds are taught in, which really goes against the principles of our phonics scheme. We feel really, really strongly that we should follow the order of teaching … we're so unlikely to change the order ... and that's quite a major part of the Flexible Phonics' (SMT staff, School 85).*

Staff at this school clarified that as they usually would not introduce sound variations for GPCs in reception, they were therefore unable to deliver set for variability. However, there were no other examples of issues with integrating the GPC order causing major disruption to delivery of Flexible Phonics in the case study settings. The delivery team advised schools that if they taught GPCs in the order specified by their existing phonics programme but used Flexible Phonics approaches to teach these, then this would be considered compliant with Flexible Phonics. It may be helpful in future roll-out to clarify this when schools are signing up and to reiterate it explicitly during training to help alleviate concerns and avoid confusion.

Tailoring which students received Flexible Phonics

Although the Flexible Phonics programme was intended to be a universal intervention suitable for all students, a little under a third of teachers and TAs surveyed (29%) reported that they taught Flexible Phonics alongside their existing phonics programme with a specific group of children and continued to use only their usual phonics approach with the rest of the children (see Table 25). While 29% of staff surveyed does not necessarily equate to 29% of schools being non-compliant in this regard, this still suggests that a sizeable minority of schools were not complying with a key aspect of delivery that was not necessarily captured by the agreed compliance measure for the study as this focused predominantly of which parts of the programme were being delivered rather than capturing to whom it was delivered. Only a few of these respondents provided further detail on which groups of pupils they delivered Flexible Phonics to: three said they delivered Flexible Phonics to higher ability students, two to lower ability pupils such as those struggling to blend, and two said they delivered it to SEND students. This suggests that, while not all children received the Flexible Phonics intervention at all schools, there did not seem to be a systematic bias where one particular group of pupils did not receive Flexible Phonics teaching across the programme as a whole. However, in contrast to the survey findings, none of the case study schools sampled described using this approach and they all taught Flexible Phonics to all children in reception.

Most delivery team support partners interviewed reported that some schools initially had concerns about teaching Flexible Phonics to certain groups of children. Most had encountered schools with concerns about using Flexible Phonics with children with English as an additional language (EAL) and one support partner reported hearing from a school with concerns about using the programme with children with special educational needs and disabilities (SEND). Flexible Phonics included an optional element for supporting children who were struggling with aspects of phonics, so teachers and TAs may not have felt the need to raise it as separate issue. However, none of the support partners reported any cases where they were unable to reassure a school or work with them to support the needs of their group of children. Suggestions and resources for working with EAL and SEND pupils were shared on the UCLeXtend platform and via the regular newsletter to schools. In the majority of cases, the interviewed teachers and TAs reported that all pupils received the same amount of Flexible Phonics teaching through whole-class activities or grouped Flexible Phonics sessions. In further contrast to the survey findings, staff at two case study schools described how a small number of pupils received additional Flexible Phonics teaching: in one school, struggling pupils were given additional Flexible Phonics teaching on an ad hoc basis while in another, Flexible Phonics was delivered to a small number of pupils who attended school in person during lockdown, but not to pupils who were learning remotely.

Tailoring delivery to pupils' needs

Teaching staff were also asked in qualitative interviews whether they had adapted the Flexible Phonics programme to meet the needs of pupils, and this provides useful insights into how and why teachers and TAs made the adaptions reported in the endline survey. Staff at one case study school adapted the programme for SEND students, including using speech therapy books for direct mapping exercises and adapting teaching to individual student needs (such as using peer learning for a pupil who struggled to take instruction from the class teacher). Teaching staff at another school described tailoring the programme to fit the needs of the class as a whole according to their phonics level. Adaptations to support less able readers included delivering support for struggling readers and using resources like flash cards to

support blending. Tailoring for higher ability pupils included using the stretching exercises from the Flexible Phonics manual and, in one case, asking these pupils to read aloud to the class. While the Flexible Phonics programme was designed to be used flexibly to meet the needs of the class or children being taught, some of the adaptations described contradicted the approaches used in Flexible Phonics, for example, learning exception words by sight instead of using set for variability strategies to decode the word.

Tailoring delivery by existing phonics teaching approach

Further survey analysis explored whether there was a difference in adaptations made to the programme between schools that used Letters and Sounds as their main phonics programme and schools that delivered other phonics programmes. Letters and Sounds was chosen as the comparison programme as this was the most common phonics programme used by staff in our sample: 49% of the 241 teachers and TAs who responded to the endline survey reported that this was the main phonics programme used by their school. Mann-Whitney tests found no statistically significant differences by phonics programme in how likely they were to have made the adaptations listed previously in Table 25. One delivery team support partner commented that there had not been a pattern where certain phonics programmes were less compatible with Flexible Phonics but that it depended more on the school's approach or set-up. Another observed that where schemes were very 'rigid' there could be challenges with integrating Flexible Phonics, but she also observed that one scheme, Sounds Write, appeared to have a similar approach to Flexible Phonics in terms of not teaching exception words through sight-learning but by teaching that tricky words had tricky sounds in them.

Challenges to delivery

Challenges and barriers to delivering Flexible Phonics were explored through the endline survey and qualitative interviews with staff members. Teachers and TAs were asked in the survey whether they had experienced any challenges or barriers to delivery and were presented with a list of possible barriers and asked to select all that applied (see Table 26). The pandemic was a key challenge: the most commonly reported barrier was that Covid-19 restrictions had affected phonics teaching (45%). Other challenges caused by the pandemic included disruptions caused by pupil or staff isolation (27%) and Covid-19 restrictions impacting how staff could share or use direct mapping books (18%). In case study interviews, a few staff reported that lockdown and school closures had meant they started delivering the programme late or that some pupils had not reached a level with their general phonics where they were ready to engage with it. Staff absences, and in one case the entire reception class having to isolate, also caused disruptions.

Time pressures were also a key challenge. The next most commonly reported barriers were difficulty fitting Flexible Phonics into the normal phonics teaching schedule (40%), and not having enough time to prepare Flexible Phonics (34%). During case study interviews, many staff members reported that it was difficult to make time in their schedules to deliver Flexible Phonics, and some also said it was difficult to make time for Flexible Phonics in pupils' timetables or existing phonics sessions.

One school also struggled to make time for TAs to attend the training or plan for delivery. TAs, teaching staff, and members of the SMT at this school all highlighted this as a major challenge because support staff lacked knowledge and confidence in delivering the programme. For another school, lack of staffing was a key barrier as its assistant head for KS1 and reception left and was not replaced during the delivery period. Further information about time needed for planning and delivering the programme is included in the Costs section.

Table 26: Challenges and barriers to delivery

Barrier or challenge	Frequency	Percent (%)
Covid-19 restrictions affecting how you can teach phonics.	44	45
Difficulty fitting it into your normal phonics teaching schedule.	40	40
Not enough time to prepare Flexible Phonics activities.	33	34
Difficulty integrating it with your normal phonics approach.	27	27
Pupils have not progressed to the point in the phonics curriculum where you can start teaching Flexible Phonics approaches.	27	27
Ensuring consistency of approach across all staff teaching phonics.	26	27
Disruption to phonics teaching caused by pupils or staff testing positive for covid-19 and needing to self-isolate for two weeks.	26	27
Not enough time to deliver Flexible Phonics activities.	21	21
Covid-19 restrictions mean that the books provided can't be shared and used as intended	18	18
Having the resources to teach Flexible Phonics.	9	9
Needing more support from senior staff to be able to embed Flexible Phonics in phonics teaching for reception.	9	9
N =	122	

Multiple options could be recorded per respondent so percentages in the table may add up to more than 100%.

Another possible barrier to delivering Flexible Phonics was where schools were delivering other language interventions concurrently. In the endline survey, just under two-thirds of teachers and TAs in the intervention group (65%) reported participating in another project or programme focusing on literacy, phonics, or language outside of the school's normal approach, with two-fifths participating in NELI (total N = 245). Small proportions of staff described participating in other programmes, which are described in the Business as Usual section of this chapter. Teachers and TAs at schools that were participating in other programme(s) focusing on literacy or phonics or language outside of their schools' normal approach were asked in the endline survey whether this had affected their ability to deliver Flexible Phonics: 52% reported that it did not, 29% said that the approaches of Flexible Phonics and the other programme(s) were complementary and worked well together, 23% said that there were no conflicts between the programmes' approaches, and 16% said that the other programme did not affect Flexible Phonics delivery as they were focused on different areas of language (total N = 55). However, some staff members did report that participating in other programme(s) affected their delivery of Flexible Phonics: 18% reported that this resulted in spending less time on Flexible Phonics and 10% said that some of the programmes' suggested approaches were incompatible or in conflict with Flexible Phonics (total N = 55).

Three case study schools were delivering the NELI intervention and no staff interviewed at these schools reported experiencing any difficulty delivering Flexible Phonics because of the NELI intervention. Staff at one school commented that the two programmes worked well together and described delivering a general phonics session with the whole class, including elements of Flexible Phonics, before breaking into groups where a TA would deliver NELI with the target group and the other groups would read another book with the target sound (that is, direct mapping).

Barriers to engaging with the programme were also explored in interviews with staff at the two schools that withdrew from the trial. Both withdrew for reasons external to Flexible Phonics: in one case the school was being shut down, in another the Flexible Phonics Lead experienced a health issue that meant they could not deliver the programme. Staff at one of these schools did express concerns following the training that it would be hard to integrate the programme with their usual Letters and Sounds approach.

However, it is also important to note that a few staff members from case studies reported experiencing no barriers to delivery at all, with one member of staff from a multi-form entry school saying this was because they were well staffed,

which enabled them to embed the programme into their teaching. Also, no staff interviewed mentioned cost as a barrier to participation. More information about any costs to schools can be found in the Costs section.

School views on Flexible Phonics

This section utilises evidence from the endline survey and case study interviews to explore staff experiences of delivering Flexible Phonics, including their views on whether the programme was effective, and the suitability of the direct mapping books.

Views on the effectiveness of Flexible Phonics

In the endline survey, teachers and TAs were asked for their views on how effective elements of Flexible Phonics were. Broadly, staff felt that all elements of the programme were effective. The proportion of staff reporting that an element was quite effective or very effective was highest for the two elements teaching set for variability strategies—mispronunciation correction (87%, total N = 74) and set for variability (88%, total N = 40)—followed by 'GPCs and direct mapping' (79%, total N = 87), teaching vocabulary (78%, total N = 74), and support for struggling readers (78%, total N = 50).

In qualitative interviews, direct mapping was identified by many staff members as an effective element of the programme, and one TA commented that direct mapping was easy to integrate with their existing practice. Several members of staff highlighted the use of real books as particularly effective, with one specifying that the texts were more interesting than phonics texts which helped children to engage in the exercise, including SEND and EAL pupils. One SMT member described how children also started using the direct mapping technique of identifying specific sounds outside of the direct mapping sessions.

> *'Trying to find the sounds, has been really, really good. And it just means that the children are now searching for the sounds in anything you put on the board … which is really nice' (SMT member, School 77).*

Some staff also reported that pupils had engaged well with the set for variability and mispronunciation activities. Two staff members at one school described how mispronunciation correction games engaged their pupils and improved their confidence in reading as it taught them that it was okay to make mistakes:

> *'The children really like [mispronunciation games] because they feel like they're coming along the journey with us' (reception teacher and SMT member, School 56).*

However, a few members of staff also highlighted elements of the programme that worked less well. Staff at a school with a high proportion of less able readers, including SEND students, found set for variability hard to deliver as many of their pupils were still struggling with segmenting and blending. Staff at another school reported that they were already using similar approaches to Flexible Phonics so felt that all elements of the programme, especially direct mapping, did not add much to their practice. One senior leadership team member said that the support for struggling readers was harder to deliver as staff felt less confident with this area of the programme. Finally, one member of staff said that they did not deliver direct mapping regularly as it required a separate session and was harder to integrate into daily teaching than set for variability and mispronunciation correction. None of the staff interviewed reported any unintended or negative consequences as a result of delivering the Flexible Phonics programme.

Views on direct mapping books

In the endline survey, teachers and TAs were asked for their views on the books provided for direct mapping. Most respondents agreed or strongly agreed that the books were high quality (92%, total N = 114), appropriate for children in reception (92%, total N = 118), fun to read (94%, total N = 117), and useful for teaching Flexible Phonics (82%, total N = 118).

Staff views on the direct mapping books were explored further in qualitative case study interviews. In line with the survey results, most staff interviewed reported that the books were high quality. Some highlighted specific aspects of the books, such as having appropriate page layouts and illustrations and a good range in terms of the number of words on a page. Teachers also reported that the books worked well for teaching Flexible Phonics:

> *'I liked the range of books. Some were new to me, some not. They were all a good quality. The children enjoyed them all. Where the children were asked to listen out for a particular sound, they were able to do this well' (reception teacher School 85).*

Staff at one school felt there were too many common titles that pupils were very familiar with, and that some were pitched at too low a level. However, they caveated this by saying that they were based in an affluent area where pupils have good access to real books. One issue raised by several staff interviewed across three case study schools was the perception that some of the books were poorly matched to the sounds, which could leave pupils disappointed during direct mapping sessions. The delivery team checked the mapping of sounds to books when this issue was raised and confirmed that the sounds were present but that the sounds may occur less frequently than in programme scheme reading books.

Rolling out the intervention

This section explores whether the Flexible Phonics intervention could be delivered to a larger number of schools and whether any elements of the programme might need to be adapted if delivered at scale. First, we present the views of teachers and TAs interviewed in case studies with regard to whether the programme was suitable to be rolled out to other schools. Then we explore the delivery team's views on what had worked well, what might be changed in future delivery, and how the programme might be scaled-up in future to identify any potential issues or considerations.

This section focuses on the following research questions from the implementation and process evaluation.

- IPE2. Could the intervention be rolled-out on a larger scale so that the intervention is delivered as intended?
- IPE3. What adaptations would be required to roll-out the intervention on a larger scale and how might these affect the integrity of how the intervention is delivered?

Teachers' and teaching assistants' views

When asked in qualitative interviews whether they would recommend that the intervention be rolled out to other schools, the majority of staff at case study schools thought that it should, typically because they thought it was a good quality programme. One TA highlighted children's' enjoyment of, and engagement with, the programme:

> *'I've definitely seen a massive improvement in my children. They look forward to it every day and I think it's a really good programme. I definitely think other children should get a chance to do it' (TA, School 14R).*

Other reasons given for wider roll-out included that it is evidence based and that it is clear and easy to use.
Some staff interviewed suggested certain conditions that they felt were necessary for Flexible Phonics to be implemented at a school. These included requirements that schools did not already have a similar programme in place, that they had a well implemented main phonics programme for Flexible Phonics to run alongside, and that schools delivering the intervention were engaged and relevant staff received training. The only adaptation suggested by some teachers and TAs was to review the list which mapped direct mapping books to specified GPCs as some schools found that some books had very few instances of the specified GPC.

Delivery team's views

Interviews with the project director and project manager explored their views on how delivery had worked during the trial and how the programme could be scaled-up in future. It is worth noting that the model of delivery changed substantially from the pilot stage with the onset of the Covid-19 pandemic, so that a model of predominantly face to face delivery and support was adapted to online delivery within a relatively short space of time. For this reason, the discussion of future delivery includes consideration of how the mode of delivery might change for a context where the pandemic has passed.

Both the project director and project manager highlighted the high quality of the support partners who delivered the follow-up support sessions with schools and provided ad hoc support where they answered queries, provided advice or clarification, and ensured that schools were delivering relevant elements of the programme, as well as helping to share ideas and resources created by the schools. The eight support partners were recruited from the postgraduate population at the Institute of Education UCL, were qualified in a relevant subject, and often also had direct experience of working in the education sector.

> *'If it is a success in the future, I think a key part of it would be because they really understood the idea, they came with degrees in psychology and education, PhD one of them, another one working on a PhD, former headteacher. Some people had been involved in professional training for teachers in schools' (project director, delivery team).*

The project director and manager stated that they would want to employ similarly qualified individuals for any future roll-out. This could present some challenges for scale-up as the delivery organisation would need to recruit individuals with relevant qualifications or experience in large numbers and across different regions or areas. A model of recruiting postgraduate students would be reliant on local institutions offering postgraduate qualifications in relevant fields, such as education or psychology. If the delivery and support modes were to change to face to face, this might add a further requirement for either individuals with access to a car or good transport networks in the area.

There was also a substantial amount of training and support provided by the project director and manager for the support partners. The initial training and induction for the support partners was delivered over three and a half days. In addition to this, once the follow-up sessions with schools were being delivered, the project manager would speak or email with the support partners every day, quality assuring their assessments of the schools' activities and compliance. Initially, there were eight support partners delivering to eight schools each, so 13 support partners would be needed to deliver to 100 schools. Depending on how large the scale-up was, there may not need to be additional training sessions for the support partners, but the time spent by the project manager responding to queries and quality assuring school monitoring forms would need to increase proportionately. The project manager role in general was acknowledged by the project director and the project manager to have been challenging in terms of the time needed to carry out the role. The project manager suggested that a second part-time admin role would be beneficial to enable them to focus on the main tasks of supporting and quality assuring delivery. The project manager and support partners also created further resources on an ad hoc basis during the delivery period as well. Several support partners described working beyond their contracted hours at times, although they noted that there were quieter and busier periods. If, for future delivery, the delivery team decided to hire additional staff then this could increase delivery costs by a quite a significant amount. If the programme team were to charge a fee for delivering Flexible Phonics that was dependent on their delivery costs, then this would increase the cost to schools. However, our cost analysis modelling with the current staff arrangements suggests that fees for this intervention would be very low (Appendix P: Detailed costs and alternate cost model). So it is likely that even if costs increased quite a lot for the delivery, the overall cost to school would still be relatively low.

The project director and project manager proposed ideas for how the delivery of Flexible Phonics might be scaled up in future. They suggested using regional hubs where training could be delivered separately for each region if they were using face to face training. Costs could possibly be reduced if a participating school could host the training, which might also enable them to demonstrate Flexible Phonics with reception-aged children at the school. If training were delivered online again, then cohorts from different regions could potentially be trained together, while follow-up sessions and ad hoc support could be delivered by a team of support partners recruited for each participating region or area. The project director was keen to continue to deliver the training himself if the programme were scaled up. For the current trial, the training sessions were delivered online in an intensive couple of weeks and potentially a similar approach could be taken. However, physically travelling to venues would probably result in a longer delivery time compared to using online sessions. The project director is a professor teaching at a higher education institution and would be delivering Flexible Phonics in additional to this role, so would need to take time out from this to deliver the training. This could become quite challenging if scale-up is very large.

While the project director and manager both expressed a preference for face to face delivery generally, neither had a strong view on what elements of future delivery should be online or face to face at this point and wanted to see findings from the trial before making that decision.

> *'I like face to face, I like meeting people and having the coffee and having the authentic professional conversations in the spaces in between. I do think those are really important because you feel the credibility, you learn about people's concerns and you can reference those when you pick them up in the teaching' (project director, delivery team).*

The project manager had also received feedback from some teachers and TAs that they would have liked a member of the delivery team to come to their school to observe and confirm that they were delivering Flexible Phonics correctly, which had been part of the initial plan for support partners before the pandemic. However, both the project director and project manager felt that they had been able to deliver the intervention at a good quality in the current adapted online

format and acknowledged that online delivery could have benefits for delivering at scale and for delivering some elements remotely. They suggested that a hybrid regional hub delivery model could give teachers and TAs the opportunity to meet with, and share practice with, schools in their area. The project director, project manager, and support partners all commented that the UCLeXtend online portal had not been used much by teachers and TAs and had not created an online community of practice as hoped. They proposed replacing the portal with a website hosting the resources so the barrier of needing to log-in was removed. During the trial, the online portal had required a log-in in order to avoid teachers and TAs from control schools being able to access the materials, whereas the delivery team stated that, for future delivery outside of a randomised controlled trial, they would probably host the materials on a publicly-accessible website.

Another suggested change for future delivery was adapting delivery to facilitate participation by TAs. It was observed by many of the support partners interviewed and the project manager that TA attendance at follow-up support sessions dropped off in some cases, and that TAs tended to speak up less than teachers in training and follow-up sessions. Both the project manager and some of the support partners emphasised the role of TAs in delivering phonics teaching in the classroom or with small groups and their broader contributions as part of the reception teaching team. The project manager commented that there was wide variation among TAs in terms of qualifications and experience in schools so any targeted delivery would need to take this into account.

Usual practice

This section explores the business as usual approaches to phonics teaching used by schools in the study, any differences between Flexible Phonics and existing approaches to teaching phonics, and staff views on any perceived outcomes or benefits of Flexible Phonics. This section focuses on the following research questions from the implementation and process evaluation.

Usual practice

- IPE19. Before Flexible Phonics implementation, what was 'business as usual' and how was this embedded in wider approaches to reading?
- IPE20. What phonics teaching and wider reading strategies are used in control schools?
- IPE21. Does the behaviour of control schools change during the trial

Outcomes

- IPE12. What are TAs' and teachers' perceived benefits and outcomes of the intervention?
- IPE17. Do senior managers perceive the intervention as worthwhile and cost-effective?

Business as usual approaches

The Flexible Phonics programme is intended to work as an add-on to existing phonics teaching by introducing new strategies for teaching and embedding phonics and reading. This section explores the approaches to teaching phonics that the intervention and control schools were using at the start of the research period (2020/2021 academic year). Teachers and TAs from the intervention and control schools answered questions about their current phonics practice and experience with reception class in the baseline and endline surveys.[31] The following section draws on information collected during the endline survey about their pre-existing approach to teaching general phonics at the start of the academic year.

[31] Unfortunately, due to a technical issue with the baseline survey, it was not possible to identify which respondents went on to belong to the intervention or control groups (as it was prior to randomisation, this information had to be matched later) so data from the endline survey has been used to enable comparison between the intervention and control groups.

Pre-existing approaches to teaching phonics

Schools in England are encouraged by the DfE to adopt a complete systematic synthetic phonics (SSP) programme and the DfE provides a list of suitable programmes on its website (DfE, 2022a). As may be expected, the majority of intervention and control schools participating in the study use one of these programmes for their general phonics teaching. In the endline survey, teachers and TAs from both intervention and control schools were asked whether they used one of the programmes listed on the DfE website for their general phonics teaching. The most commonly used phonics programme was Letters and Sounds (intervention: 43%, control: 55%), followed by Read Write Inc (intervention: 33%, control: 26%), and Jolly Phonics (intervention: 12%, control: 6%). A few teachers and TAs also reported using Phonics Bug (intervention: 2%, control: 1%) and Floppy's Phonics (intervention: 1%, control: 0%), (total Ns = intervention: 120 respondents, control: 121 respondents). Around a tenth of teachers and TAs surveyed taught using a mix of programmes (intervention: 8%, control: 11%) and a minority used another phonics programme not on the DfE list (intervention: 3%, control: 1%) such as Espresso Phonics, LCP Phonics, or Storytime Phonics (total Ns: intervention: 120 respondents, control: 121 respondents). The majority of teachers and TAs who used a blend of phonics programmes were combining Letters and Sounds with either Jolly Phonics or Read Write Inc and often with a third programme such as Floppy's Phonics, Espresso Phonics, PhonicsPlay, Mr Thorne, or with both Jolly Phonics and Read Write Inc. A couple of teachers and TAs reported adding in resources and programme elements that their school had developed.

Case study interviews with teachers and TAs in the intervention group provided some examples of their usual phonics delivery approach. Several schools in the case studies reported mainly using one phonics programme and then adding in elements of one or two other programmes. For example, teachers and TAs at one school described a centrally planned approach to phonics with daily 30-minute sessions using Letters and Sounds as their main programme, with additional songs and activities from Jolly Phonics. A few schools described following just one programme: either Letters and Sounds or Read Write Inc. One of these schools taught Read Write Inc exclusively four times a week in 20-minute sessions and streamed children into ability groups. However, during this project, children were being taught in class groups because of the need to keep children in class bubbles because of the Covid-19 pandemic. Another school usually streamed reception and Year 1 classes into six groups but they were only streamed into three groups during the research period because of Covid-19 restrictions. Both the interview and survey data suggest that a number of schools were quite pragmatic in their use of phonics programmes and would add or adapt elements or activities as needed.

Other language interventions

In the endline survey, teachers and TAs from both intervention and control schools were asked whether their school was taking part in any other projects or programmes focusing on literacy, phonics, or language outside of the schools' normal approach during the research period. Around two-fifths of teachers and TAs reported that their school was participating in the EEF trial for the Nuffield Early Language Intervention (intervention: 40%, control: 41%). Other language or communication programmes that were identified included Destination Writer/Reader (intervention: 4%, control: 5%), Power of Reading (intervention: 5%, control: 3%), Early Talk Boost (intervention: 4%, control: 3%), Cornerstones' curriculum planning (intervention: 3%, control: 1%), Hooked by Books (intervention: 3%, control: 2%), Early Words (intervention: 2%, control: 0%), and one intervention school described using Language Links and Five Minute Box with small targeted groups of children but not with whole classes (total Ns: intervention: 122 respondents, control: 120 respondents). The NELI programme targets children who are struggling with language and is designed to improve spoken language ability.[32] Teaching assistants are trained to deliver three 30-minute sessions per week to groups of five children during the spring and summer terms (20 weeks), so this overlapped with the Flexible Phonics delivery and research period (January to June). Participating children also attend two 15-minute individual sessions per week. The sessions focus on listening, narrative, and vocabulary skills and phonological awareness is introduced in the final ten weeks. As the NELI intervention was new and similar proportions of intervention and control schools were participating in the trial, we anticipated that this might affect business as usual phonics delivery for both intervention and control schools to some extent. However, it is possible that the Flexible Phonics and NELI interventions interact so that they either complement and reenforce each other, or they reduce the effectiveness of the other intervention. For example, NELI targets a small group of struggling children from each class whereas Flexible Phonics is a universal intervention for all children in reception, so we would hope to see a difference across the wider population between intervention and

[32] https://www.elklan.co.uk/NELI/

control groups. However, it is possible that the biggest changes in performance would be among struggling readers, so the NELI intervention might dampen down this effect a bit. In addition, any school trying to do more than one intervention may have less time and resources to spend on each, so it is possible that this could impact on the implementation of Flexible Phonics in intervention schools. Three case study schools were delivering the NELI intervention and no staff interviewed at these schools reported experiencing any difficulty delivering Flexible Phonics because of the NELI intervention (see Challenges to Delivery section for further detail).

Reasons for incorporating Flexible Phonics into existing phonics teaching

The Flexible Phonics programme aims to introduce two novel strategies, direct mapping and set for variability, which can be used in addition to a school's general phonics programme to help embed learning, improve children's abilities to decode exception words, and improve reading skills and confidence in reading (see ToC model in the Introduction). During case study interviews, staff were asked about their motivations for taking part in the Flexible Phonics trial and what they hoped to gain from participating in the programme.

Reasons for taking part typically related to the school context and development priorities. Most SMT staff interviewed had heard about the programme via email mailing lists, including EEF alerts, and a smaller number by local authorities and word of mouth. A few schools highlighted staff CPD as a key reason for taking part, including a school which delivered its own initial teacher training and wanted to ensure their staff were up to date on phonics practice. One EYFS lead at another school explained that they participated in order to introduce new ideas into their phonics teaching and improve practice.

> *'I always think schools should have a freshen up on what they're doing. We can't always rely on what we've always been doing, although we are a school that's quite strong [in] reading, in terms of results, but I think there's always room for improvement and there's always room to have new ideas, and new research pumped into what we're doing. So, it did appeal to me to have a bit of a shakeup' (EYFS lead, School 13).*

SMT staff at a few schools commented that the main reason they decided to take part in the programme was to improve reading outcomes. For one of these schools this was a priority in their development cycle, while another had a high proportion of EAL pupils and had low baseline levels of achievement at the start of reception. Staff at two schools, including a designated research school, highlighted the fact that Flexible Phonics is an evidence-based intervention as a key reason for taking part. Finally, one member of SMT staff commented that they were drawn to accessing free children's books. A few schools liked that Flexible Phonics acts as an add-on to existing phonics programmes as they liked their current phonics programme or did not want to make a major change to it, so appreciated that Flexible Phonics would allow them to improve their phonics practice without changing their fundamental approach.

Perceived outcomes

This section explores any perceived benefits and outcomes of the intervention as identified by teachers, TAs, and managers at the settings. These include perceived changes in teaching approaches and strategies, children's engagement with phonics and reading, and children's progression in phonics during the study. This section draws upon information collected in case study interviews and the teacher and TA surveys.

Did teaching practice and behaviour change during the trial?

This section explores changes in behaviour or practice while teaching phonics to children in reception and investigates whether there were any differences between the behaviour and practice in the control group and the intervention group. Change in behaviour was measured by asking teachers and TAs in a survey to self-report on whether they engaged in certain activities as part of phonics teaching and, if so, how often. These activities were drawn from the 2017 Ofsted report 'Bold Beginnings: The Reception Curriculum in a Sample of Good and Outstanding Primary Schools' as examples of good practice in reception and included activities regarding phonics, reading comprehension, and spelling and handwriting. A full list of these activities can be found in Appendix N Endline Survey.

Teachers and TAs were asked in the endline survey to report how often they engaged in the activities and whether the frequency of this had changed since the start of the year.[33] Results for control and intervention groups were compared to identify any statistically significant differences in the behaviour and change over time reported by the two groups. Mann-Whitney tests identified no statistically significant differences between teachers and TAs in the intervention and control groups with regards to either self-reported engagement with the activities or self-reported change in engagement over time. Below, we present examples of some key activities that were most relevant to the strategies and activities used in Flexible Phonics. Tables presenting findings for all the behavioural and practice change items are included in Appendix O.

The Flexible Phonics programme encouraged staff to teach phonics on a daily basis and incorporate it throughout the school day. In the endline survey, the vast majority of teachers and TAs in both control and intervention schools reported undertaking basic phonics activities daily, such as using phonics to decode words (intervention: 94% of 105 respondents, control: 97% of 114) and making the correct sound for a letter or group of letters (intervention: 93% of 104, control: 96% of 112). The majority stated that, since the start of the year, they had not changed how often they were using phonics to decode words (intervention: 60% of 74, control: 64% of 97) or making the correct sound for a letter or group of letters (intervention: 61% of 75, control: 66% of 96). However, around a third had increased the frequency, for example, using phonics to decode words (intervention: 37% of 74, control: 32% of 97) and making the correct sound (intervention: 33% of 75, control: 32% of 96). Overall, staff in the intervention and control groups behaved similarly in how they delivered these activities over time.

The set for variability strategies aimed to support children with learning exception words. This was not reflected in a difference in levels of activity involving exceptions words. A large majority of teachers and TAs in both intervention and control schools undertook activities around reading exception words daily, such as reading by blending sounds in unfamiliar words (intervention: 83% of 104, control: 81% of 113) and highlighting unusual sound and letter combinations while reading exception words (intervention: 72% of 104, control: 68% of 113). For both exception word activities, the largest proportion of teachers and TAs reported no change in how often they carried out these activities (intervention: 41–58% of 73–74, control: 60% of 95–96). Similarly, there was also no difference between intervention and control schools in the levels of engagement with writing common exception words. Around half of teachers and TAs incorporated activities which involved writing common exception words that had been learned for reading on a daily basis (intervention: 55% of 96, control: 47% of 106). The majority of teachers and TAs reported that the frequency of this activity had not changed since the start of the academic year (intervention: 60%, control: 59%) (total Ns: intervention: 78 respondents, control: 86 respondents).

The direct mapping approach involved reading high quality children's books to reinforce phonics learning, but this did not lead to a difference between the intervention and control group in the frequency of activities involving reading books. Around half of teachers and TAs in both intervention and control schools described reading books to reinforce existing phonics knowledge (intervention: 59% of 102, control: 55% of 112) or re-reading books to build up confidence and fluency (intervention: 44% of 102, control: 42% of 113) on a daily basis. Around half of teachers and TAs reported no change in how often they engaged in these activities since the start of the year (intervention: 43–49% of 73–74, control: 55–59% of 93–94). In general, the Flexible Phonics programme aimed to build children's confidence when attempting to read new words. However, there was no difference between the intervention and control schools in how often activities involving reading unfamiliar words occurred. The majority of teachers and TAs in both intervention and control schools undertook daily activities incorporating reading unfamiliar words, such as using phonics knowledge to decode unfamiliar words (intervention: 71% of 103, control: 72% of 112) and using pictures and context to help understand unfamiliar vocabulary (intervention: 63% of 104, control: 60% of 111). The majority of teachers and TAs reported that they had not changed the frequency of these activities since the start of the year (intervention: 68–69% of 80, control: 62–63% of 98–99).

[33] In the original design of the study, the intention was to ask teachers and TAs at the start and end of the academic year how often they engaged in the specified activities and then to compare whether this was different between the intervention and control schools at each timepoint to identify whether there had been a change in practice. However, due to a technical issue with the baseline survey, it was not possible to identify which respondents belonged to the control or intervention groups, so it has not been possible to track the groups over time in this way.

However, most teachers interviewed for the case studies felt that the programme had made changes to their phonics teaching, with many planning to incorporate elements of the programme into their personal teaching practice. Flipping sounds and mispronunciation strategies for dealing with exception words were the most common elements that staff said they would continue using as these had been embedded into their daily teaching practice and communication with pupils. To a lesser extent, a few members of staff planned to continue using direct mapping and reported that they were more aware of how to use real books in phonics teaching. Staff were less likely to say they would continue using other elements of the programme, such as support for struggling readers and teaching vocabulary for exception words, with just one interviewee saying they would continue using the strategy of teaching 'schwa' sounds to struggling readers, and another saying they would teach variable GPCs from the start of reception in future.

The teachers and nursery nurse at one case study school reported that Flexible Phonics had not impacted their phonics practice as they felt that it was quite similar to their existing approach, Read Write Inc. However, this school was not teaching the set for variability elements of the intervention because they felt it did not fit with the order of teaching sounds in Read Write Inc so were omitting one of the core, novel elements of Flexible Phonics. One teacher at another school commented that Flexible Phonics had not impacted their practice because they were already confident teaching phonics. It is worth noting, however, that for all eight case study schools, at least one member of staff described learning and using new activities, strategies, or skills from the programme into their practice. Staff at one school felt it had made a really positive impact on their practice and another school felt it had given them the opportunity to question and completely reform their practice. A further two schools felt it had impacted their practice or given them skills.

Perceived changes to confidence and phonics knowledge

In the endline survey, teachers and TAs in both the intervention and control groups were asked whether they felt more or less confident teaching phonics compared to the start of the 2020/2021 academic year. There was no statistically significant difference between responses from the two groups: a little over a third of staff in both groups (34% and 25%, respectively) reported feeling a lot more confident, and a third of intervention group staff (33%) and a fifth of control group staff surveyed (20%) felt a bit more confident, while a third of intervention group staff (33%) and over two-fifths of control group staff (45%) reported that their confidence was about the same; only 1% of respondents in the control group felt a bit less confident and none in the intervention group reported this (total Ns = intervention: 120 staff, control: 121). There was no difference in perceived change in confidence between teachers and TAs responding to the survey.

However, across all teachers and TAs who responded to the endline survey, there was a significant difference in reported changes in confidence by the length of time that teachers or TAs had been teaching or supporting reception-aged children (Mann-Whitney test). We classified teachers and TAs into two categories: 'newer' teachers and TAs were those who had been teaching for five years or fewer, and 'experienced' teachers and TAs were those who had been teaching for more than five years. Newer teachers and TAs were more likely to report increased confidence (median = 2.0, mean = 1.86) than experienced staff (median = 2.0, mean = 2.21), $U = 5672$, $p < 0.01$, $r = -0.20$. Over two-fifths of newer teachers and TAs (44%) reported feeling a lot more confident compared with just over a quarter of experienced staff (26%), around a quarter each of newer and experienced staff reported being a bit more confident (25% and 28% respectively); 45% of experienced staff felt their confidence was about the same compared with 31% of newer staff, and 1% of experienced teachers and TAs felt a bit less confident (total Ns = Newer teachers and TAs: 111 staff; Experienced teachers and TAs: 130 staff). It is worth noting that in the baseline survey at the start of the academic year, a large majority of teachers and TAs (89%) agreed or strongly agreed that they felt confident in their ability to teach phonics or support phonics teaching with reception-aged children (total N = 421 staff).

Phonics teaching in reception year is common practice in England so it might be expected that teachers and TAs were already quite confident in this area. Flexible Phonics provides additional strategies to build on existing phonics programmes so while the training aimed to reenforce and expand existing understanding of phonics, language processing, and language development in reception-aged children, it is perhaps understandable that the intervention did not significantly impact reported confidence when most respondents were already reporting high levels ('ceiling') of confidence. Participating in Flexible Phonics neither increased nor reduced levels of confidence among staff.

While there was no statistical difference between the intervention and control groups in terms of confidence teaching phonics, there were some teachers in both groups who reported that their confidence increased over the trial period, especially less experienced staff. It seems logical that, for some staff, the Flexible Phonics programme could contribute

to this increased phonics confidence and understanding. In the qualitative case study interviews with staff who were delivering Flexible Phonics, several members of staff reported that it had improved their confidence and phonics knowledge and a few highlighted improved confidence as the main personal impact of the programme. One EYFS lead commented:

> *'I think it has given us additional skills, and it I think it's given us the confidence to use the skills that I think we've always known are really important to teach children' (EYFS lead, School 13).*

One reception teacher also reported that Flexible Phonics had made them more confident in teaching reception pupils the complexity of the English language and that the set for variability strategies enabled her to support children in encountering tricky words or sounds and encourage them to have a go at reading them without worrying about getting it wrong.

A few staff members interviewed also felt that their phonics knowledge had improved. This ranged from gaining a better understanding of teaching phonics to more specific knowledge, with one TA highlighting that they had learned that stop consonants are harder for children to learn than continuous ones. A reception teacher also said that they were thinking more critically about phonics and one TA explained that:

> *'It just made me feel, like, a lot more aware of, you know, what the children should be doing in their writing, and how they should be doing their reading, and the phonics' (TA, School 14R).*

Impacts on the phonics practice, confidence, and knowledge of teaching staff were also observed by some members of SMT interviewed. While some felt unable to comment on the impact of the programme on their teaching staff, a few said that it had acted as meaningful CPD that led to staff being more skilled and confident and having new approaches to phonics teaching. Overall, three case study schools had staff who reported increased confidence in their phonics teaching through participating, staff at two case study schools described becoming more confident at their delivery of the programme itself, and at one further school staff felt they were already quite confident in their phonics teaching before participating.

Perceived benefits and outcomes for children

In the endline survey, teachers and TAs in both the intervention and control groups were asked whether they felt children were more or less engaged with their phonics teaching compared to the start of the 2020/2021 academic year on a five-point scale from 'a lot more engaged' to 'a lot less engaged'. There was no statistically significant difference between responses from the two groups (Mann-Whitney test): around two-thirds of staff in the intervention and control groups (68% and 65% respectively) felt that children were a bit or a lot more engaged, and around a third felt that children's level of engagement was the same (intervention: 32%, control: 35%; total Ns: intervention, 120 staff; control: 121).

Teachers and TAs were also asked in the endline survey whether they felt that the Covid-19 pandemic had affected the phonics skills of their reception class so we could assess the context that the project was working in that academic year. Just over half of those responding (52%) thought that some children were delayed but others were at the usual level or above; over a fifth felt that the children were delayed by two to three months (23%); 7% each felt that children were delayed by four to six months or a month, 2% by more than six months, and 4% were unsure (total N = 241). Some teachers and TAs who were interviewed as part of the case studies commented that the children they worked with were behind in the their phonics or had lost confidence and felt that this could affect how much impact the Flexible Phonics intervention might have. A few went on to note that while children have been at home, phonics progression has been dependent on parents teaching phonics. Several teachers and TAs felt that the delay with starting delivery of Flexible Phonics (late March instead of late January) because of school closures during the pandemic may affect the degree of impact that Flexible Phonics might have. A couple of teachers and TAs suggested that it might be difficult to compare 2020/2021 to a typical year.

During case study interviews, most teachers and TAs reported some positive impacts on phonics and reading although a couple reported no impacts, one was unsure, and another was unsure whether progress could be attributed to the Flexible Phonics intervention. Among those who identified impacts, most described improvements in reading and increased confidence or engagement with children using contextual information on the page, such as pictures or having a go at figuring out a new or difficult word by sounding it out.

> *'I have one of the children in my class who they, I would say, up to January weren't able to recall many of the basic sounds from the alphabet and just by improving and getting him to think that, yes, lots of these could be tricky, there can be quite a few different sounds, we've got to use our flexible mindset and up to now he's finally got that confidence in him to carry on attempting the sounds of the words, whereas at the beginning of the year he would just break down and cry because it was just a bit too overwhelming' (reception teacher, School 56).*

Other improvements that were observed include improved resilience among readers of all levels, higher ability children starting to read more complex books, more awareness and thinking about sounds while reading or listening to a story, and using flexibility strategies such as flipping sounds. For all eight case study schools, there was at least one member of staff who reported positive impacts on children's outcomes, increased confidence. or motivation when reading, or children effectively incorporating Flexible Phonics strategies when reading.

Which groups of students did Flexible Phonics work more or less well for?

In the endline survey, teachers and TAs were asked to what extent they agreed or disagreed that Flexible Phonics was working well for different groups of students. The majority agreed that the programme worked well for high performing readers (81%, total N = 112) and average readers (79%, total N = 109) but just over half agreed that it worked well for struggling readers (55%, total N = 112). Only around a quarter felt that Flexible Phonics worked well for EAL children (26%, total N = 110) or those with SEND (24%, total N = 113).

However, views were somewhat mixed among teachers and TAs interviewed as to whether the programme was more or less suited to certain groups of students: a couple felt that the programme was suitable for all students and a few reported that it worked well with lower ability children who found the direct mapping books motivating and that the programme gave them additional tools to use. However, a couple reported that their children were still learning sounds and phase 2 phonics, so they were not able to engage with Flexible Phonics yet. One TA was focusing on the blending and segmenting sounds element, another felt that they maybe needed to find more strategies to use with their lower ability children, and a couple of teachers reported that there were a couple of children in their classes who did not participate and received one to one sessions instead. The programme included strategies for basic phonics and children who were struggling so it was intended that support partners would be able to support staff with these activities. Indeed, a couple of staff reported using the approach successfully with EAL children and that it improved their vocabulary and helped them engage with peers, and another two described using Flexible Phonics strategies with children with special educational needs. Several staff described how Flexible Phonics worked well with high ability children and that it could be used to stretch them through more complex books and being less concerned about making mistakes. Overall, there were three case study schools where staff felt that Flexible Phonics could be used with all children, one school where staff interviewed were unsure, and four where staff highlighted that set for variability was challenging for children who had not yet learned or felt confident with their basic phonics.

It is worth noting that while the Flexible Phonics programme included elements to support children who were struggling with learning phonics, using set for variability strategies during reading requires children to have mastered certain basic phonics sounds and skills, for example, blending, and so could be considered a more advanced technique. Although schools may have used oral mispronunciation correction games such as 'Simon Says' with their whole class, they may not have started teaching set for variability during reading to learners who were struggling. This approach would be considered 'compliant' in terms of the trial but may have felt to schools like they were not teaching the 'full' Flexible Phonics programme because they had not been able to teach all aspects to all children in their reception year.

Intentions to continue using Flexible Phonics

The majority of staff interviewed in the case studies reported that their school planned to keep using the Flexible Phonics approach in the next academic year, with some teachers and TAs highlighting elements of the programme that they liked: many mentioned the direct mapping element, and a small number mentioned the flipping sounds strategy. One school particularly liked that Flexible Phonics included a mix of decoding strategies and reading. A few schools reported that they were intending to roll out the training to more staff and two case study schools planned to extend the approach to their Year 1 teachers as well so that the participating cohort could continue to be supported in this way if needed after the disruption to their education during their reception year. Some staff planned to use the programme with all their Year 1 students.

A few managers were unsure whether they would continue using Flexible Phonics and wanted to review with their staff how the programme had gone this year. One school was planning to change their overall phonics approach and would need to consider how Flexible Phonics fitted in with this and another was waiting to see how the programme had impacted pupil outcomes. Among the few schools that were not intending to deliver the full Flexible Phonics programme the following year, staff were still planning to incorporate some elements of the programme into their practice, typically direct mapping and decoding strategies.

Several senior managers said that they would recommend the scheme to other schools to deliver alongside their existing phonics programme, although a couple of managers felt they would need to know more about the final outcomes of the study before recommending it.

Cost

This section outlines the cost to schools of participating in the Flexible Phonics programme as it was delivered during the evaluation period. Changes to how the intervention was delivered due to the Covid-19 pandemic had a significant impact on the costs: the planned face to face training was delivered online (meaning no travel costs for school staff were incurred) and because remote learning was in place during January and February 2021 when the training was being delivered, teacher cover costs were not needed. In addition to this, all schools in England are expected to teach phonics as part of the national curriculum so they already have in place the materials and resources for phonics teaching. The Flexible Phonics intervention is an extension to existing phonics teaching so the programme does not require much in the way of additional materials. The main additional materials—a set of high quality children's books and Flexible Phonics manuals—were provided by the delivery team free of charge. Thus, in this particular instance, the cost to schools was minimal. Table 27 shows the resources that were needed to implement the Flexible Phonics programme, Table 28 presents the actual cost incurred to intervention schools in this trial, and Table 29 shows the estimated cumulative costs of delivery over three years.

Table 27: List of resources (ingredients) for Flexible Phonics

Category	Item
Personnel for preparation and delivery	One classroom teacher and one TA per class.
Personnel for training	Delivery team encourages all reception staff to attend training, which usually includes one teacher and one TA per class. We based our model on a two-form entry school, so costs are calculated for two teachers and two TAs per school.
Training and programme costs	UCL currently provides training free of charge.
	Training took place online due to the pandemic but usually would be face to face over two full days. Normally, supply teacher/TA cover would be needed, as well as travel costs, if the training was in person. Staff also attended three support sessions throughout the year, which usually took place after or before teaching hours and lasted up to 30 minutes.
Facilities, equipment and materials	Flexible phonics uses the materials that schools already have for their normal phonics teaching and so did not require further resources.
	UCL provided a set of children's books and Flexible Phonics manuals but this may be a cost in future.
Other programme inputs	N/A

In qualitative interviews, staff were asked whether there were any additional costs to delivering the intervention: the majority reported none and the few that did described these as optional purchases. One school purchased two additional sets of the direct mapping books, another printed posters of the five-step strategy for encountering tricky words. Two individual members of staff described purchasing additional resources to deliver the programme—magnetic letters and Lego blocks, which they attached letters to for blending and writing activities. However, a few members of SMT staff interviewed said that they would have had to pay to cover staff to attend the training had it not been delivered during school partial closures. One TA reported that they had attended a follow-up support session outside of their contracted hours and had not been paid for this. This issue was not, however, raised by any other staff interviewed and SMT staff interviewed did not report paying staff to work extra hours to attend training or follow-up sessions. It is difficult to estimate what proportion of staff in the wider intervention group had also worked outside of contracted hours without pay to participate in Flexible Phonics activities and whether there is a hidden cost here—as well as an ethical concern if staff are being expected to work extra unpaid hours to deliver the intervention.

Table 28: Cost of delivering Flexible Phonics

Item	Type of cost	Amount	Total over 3 years	Total per pupil per year over 3 years
Teacher cover	Start-up cost per school	£0 (not needed in pandemic year)	£0	£0
TA cover	Start-up cost per school	£0 (not needed in pandemic year)	£0	£0
Programme fee (normally covers cost of training and delivery teams—staff time)	Start-up cost per school	£0 (covered by UCL)	£0	£0
Travel and subsistence for training	Start-up cost per school	£0 (not needed in pandemic year)	£0	£0

Table 29: Cumulative costs of Flexible Phonics—assuming delivery over three years

	Year 1	Year 2	Year 3
Flexible Phonics	£0	£0	£0

We have provided a detailed breakdown of costs in Appendix P as well as an alternate model, which includes costs for a more usual year where all children would be taught face to face in the classroom as normal and using assumptions where:

- staff would attend face to face training;
- cover teachers and TAs would be needed while staff attended training;
- travel costs would be incurred;
- schools would be required to buy the set of children's books and the flexible phonics manual; and
- there would be a fee associated with the programme to cover the costs of training and UCL delivery time.

Even so, this alternate model shows very low costs per pupil so it can be concluded that the Flexible Phonics programme is a very low-cost intervention.

Time required to deliver Flexible Phonics

Staff at intervention group schools were asked in online surveys and case study interviews about the time needed for various aspects of the programme. Table 30 presents the time needed for training and teacher cover, and Table 31 shows time need to prepare and deliver Flexible Phonics.

Table 30: Total time devoted by personnel for training and teacher cover

		Year 1		Year 2		Year 3	
		Number of teachers	Total number of hours	Number of teachers	Mean number of hours (measure of dispersion)	Number of teachers	Mean number of hours (measure of dispersion)
Training and support sessions	Teacher training	1 per class	3 x 3-hour sessions; 3 half-days if online; (2 full days if in person).	0	0	0	0
	TA training	1 per class	3 x 3-hour sessions; 3 half-days if online; (2 full days if in person).	0	0	0	0
	Teacher attends support sessions	1per class	3 x 30-minute sessions; 2 in-person visits to settings: observation of a phonics lesson and follow-up support (30–60 mins) as needed.	0	0	0	0
	TA attends support sessions	1per class	3 x 30-minute sessions; 2 in-person visits to settings: observation of a phonics lesson and follow-up support (30–60 mins) as needed.	0	0	0	0
Teacher cover	Teacher cover	1 per class	0 days during trial (2 days in a non-pandemic year).	0	0	0	0
	TA cover	1 per class	0 days during trial (2 days in a non-pandemic year).	0	0	0	0

In interviews, one senior manager commented that the partial school closures in January to March 2021 were quite convenient in terms of not needing to pay for cover for staff while they were attending training.

> *'The training came at a great time because it was just as the pandemic started and people were at home, so we could flexibly provide people doing that' (senior manager, School 85).*

In both the endline survey and qualitative interviews, staff were asked to describe how much time they needed to prepare and deliver the Flexible Phonics programme. In the endline survey, staff reported spending between zero and ten hours preparing each week, with a mean of 1.5 weekly preparation hours. In terms of delivery, staff reported spending between 0.25 and 35 hours each week, with a mean of 2.9 hours delivery time. Overall, staff spent between 0.7 and 35.5 hours in total each week preparing and delivering the programme, with a mean of 4.4 hours. The median (3 hours) and mode (3 hours) averages were slightly lower. It is worth noting that the delivery team encouraged schools to incorporate Flexible Phonics activities and strategies across the reception curriculum and teaching, which may account for the member of staff who reported spending 35 hours a week delivering Flexible Phonics. Only two other staff responding to the survey reported spending more than a day—seven hours—preparing and delivering the programme each week: one reported spending 30 hours, the other 18. These findings inform the table below.

Table 31: Total time devoted by personnel for preparation and delivery

		Year 1		Year 2		Year 3	
		Number of teachers	Mean number of hours per week (min–max) as indicated by survey data	Number of teachers	No data collected so estimates are provided* (hours as in year 1)	Number of teachers	No data collected estimates are provided*
Preparation	Teacher	1 per class	1.5 (0–4)	1 per class	1.5	1 per class	1.5
	TA	1 per class	1.6 (1–3)	1 per class	1.6	1 per class	1.6
Delivery	Teacher	1 per class	1.96 (0.45–5)	1 per class	1.96 (0.45–5)	1 per class	1.96 (0.45–5)
	TA	1 per class	1.94 (1–3)	1 per class	1.94 (1–3)	1 per class	1.94 (1–3)

* As the programme was delivered for one year only, we have estimated the number of hours for preparation in years 2 and 3 to be the mean hours indicated in year 1, but we believe this is likely to be less. We have assumed that the hours needed for delivering Flexible Phonics would be the same as in year 1.

In interviews, most staff members who discussed daily planning described spending only a short amount of time each day planning Flexible Phonics, including choosing a key sound or exception word, choosing direct mapping books, and arranging other resources such as whiteboards or scanning books for remote delivery. At one school, this planning did not add an additional time burden as it was incorporated into their usual daily planning sessions, while two other members of teaching staff at other schools described spending around five minutes a day on planning. However, two members of staff also described carrying out planning over lunch or in their own time. A few said that they spent more time planning towards the start of delivery, but this reduced to a small daily amount once delivery was well established. One TA described attending a three-hour initial planning meeting and a teacher and TA from one school described needing more planning time initially—25 minutes to two hours per week while setting up—but this had reduced over time. A few SMT members discussed attending initial planning meetings, which for one interviewee including spending time to plan staffing of the project. In some cases, early years leads or phonics leads described being quite involved in the initial set-up and preparation for Flexible Phonics delivery.

Teaching staff interviewed reported spending a small amount of time each day delivering Flexible Phonics as it was incorporated with their usual phonics teaching. For example, a couple described spending 20 to 25 minutes delivering phonics or reading activities and in one case an additional 15 minutes was spent delivering phonics as part of other activities, such as playing mispronunciation games when lining up in the playground. Just one case study school reported delivering separate, daily 20-minute sessions, which they had to account for in their timetable. A small number of SMT staff interviewed provided ongoing support with delivery such as observing sessions and providing informal support to teaching staff. Some senior managers reported needing time initially to support staff as they started to deliver the programme but that this had reduced over the delivery period.

Conclusion

Table 32: Key conclusions

Key conclusions	
1.	Pupils who participated in Flexible Phonics made the equivalent of one month less progress, on average, in early word recognition than pupils who did not receive the programme. This result has a moderate to high security rating.
2.	Pupils who participated in Flexible Phonics made the equivalent of zero months' progress, on average, in reading comprehension and correcting deliberately mispronounced words than children in other schools.
3.	Exploratory subgroup analyses found pupils who were eligible for free school meals who participated in Flexible Phonics made the equivalent of no months' additional progress in word recognition compared to similar children who did not receive the programme. There was marginal evidence that in Flexible Phonics schools that also received the Nuffield Early Language Intervention (NELI), pupils made more progress in word recognition than in Flexible Phonics schools that did not register for NELI.
4.	Teachers and TAs in Flexible Phonics schools reported that it was relatively straightforward to integrate the programme into existing phonics practice. However, a minority of educators were unclear about which elements of the programme were compulsory to deliver, so future delivery could seek to emphasise these aspects.
5.	Around 100 teachers and TAs surveyed in Flexible Phonics schools suggested that there was no change in their confidence or overall practice regarding phonics teaching, although confidence was already high at the start of delivery. They indicated that children engaged in activities well and approached reading with confidence and increased resilience.

Impact evaluation and IPE integration

Evidence to support the theory of change

The theory of change (Figure 1 in the Introduction) was developed over the course of the evaluation as discussed in the Introduction. These updates were mainly due to the changes made necessary to the delivery due to the Covid-19 pandemic, but also as more detail was available about the ongoing support that would be given to schools and the resources they would receive. The ToC still appears to be accurate up until the 'outputs' box (the rationale, theory of change, and inputs have been updated accordingly throughout the project). However, the rest of the model (outputs, short term outcomes, long term outcomes, and enabling factors) are now all outdated based on the analysis of this evaluation. Although the intervention was delivered as intended in the ToC in approximately 70% of the schools (as discussed in the Compliance section of the Impact section), the first element of the outputs, 'teachers/TAs delivering Flexible Phonics as part of their standard practice', seems to have been at least largely met. But the second element, 'pupils use these strategies for everyday reading …', does not seem to have been met as teachers did not consistently perceive changes to children's reading. The short term outcomes and mediators have also not been met as our teacher and TA survey did not find changes in teacher or TA confidence in delivering phonics lessons, and we also did not find that pupils had greater word reading abilities in the primary analysis of the impact evaluation. There was also little evidence of the long term outcomes as overall literacy outcomes have not been improved and neither has phonological awareness (which were measured by the secondary outcome measures in the impact evaluation). Although we know teachers were using direct mapping and set for variability strategies with the pupils, it is unclear how often the pupils were using these strategies independently and therefore the third element of the long term outcomes is only partially met. Moving on to the enabling factors, the FSM analysis does show potentially a differential impact on FSM pupils, so the first element of this could be partially met, but as the credibility intervals cross zero this is not established. The examination of low versus high ability pupils did not show any differential impact. The ToC posited that schools taking part in NELI may have less resources to take part in Flexible Phonics, but in fact the impact evaluation shows that those taking part in both were at an advantage, so this has not been supported by the evidence of this evaluation. We did not measure differences within geographical regions in this evaluation, or if there were differences by one-form versus multiple-form entry or teaching in small groups versus whole-class delivery, so these could be explored in a future evaluation.

Interpretation

Child outcomes

The impact analysis shows that there was no clear positive effect of the intervention on the primary and secondary outcomes. The absence of any positive impact on these outcomes implies that the set for variability and direct mapping approaches are not effective. This result held regardless of whether prior attainment was taken into account.

There was no differential impact by student ability, nor eligibility for free school meals, indicating there is no clear evidence that Flexible Phonics was more effective for these groups. However, there was some evidence that it was more effective in schools participating in NELI (see the earlier discussion on Tables 18 and 19 for more information on this). The Covid-19 pandemic disrupted the trial to the extent that some schools had remote-testing for the pre-test. This did appear to have an impact: Flexible Phonics was less effective in schools that were tested face to face rather than remotely.

The finding that Flexible Phonics was more effective in schools that were participating in the EEF effectiveness trial for the Nuffield Early Language Intervention possibly suggests that the Flexible Phonics programme might have had greater impact in a normal year when children were less delayed in their general phonics learning. As the NELI intervention targets children struggling with spoken language in reception, it is possible that children in schools that were in the intervention group would have caught up with general phonics faster, therefore enabling the set for variability and mispronunciation correction strategies to be taught earlier and have more effect. Alternately, it is possible that schools that were participating in both the Flexible Phonics and NELI trials had greater motivation or resource for supporting language development in reception than schools there were not participating in both trials. It may also be the case that some activities or aspects of the two programmes are similar and so these strategies are doubly re-enforced, or that strategies from both interventions are being used alongside each other with the whole class or with the groups of struggling readers. However, in case study interviews with schools implementing the NELI, staff did not mention using NELI strategies with the whole class and one school explained that the teacher taught Flexible Phonics and the TA taught NELI so neither was familiar with the other programme.

With regard to the finding that Flexible Phonics was more effective at schools that had remote testing rather than face to face testing, it is possible that schools that opted for remote testing may have been more cautious in their Covid-19 restrictions and so were able to maintain more continuity of teaching, or it is also possible that they may have been more affluent or better resourced. Remote testing required a member of reception staff to sit with the children during testing, a laptop, and a reliable wi-fi connection. Schools that had more resource may have been better placed to support children's language recovery after the partial school closures or may have more children from affluent areas or families where children's language development and the communities themselves were less impacted by the pandemic. However, it is also possible that children who were tested remotely may have somehow been cued by the member of reception staff present in the room during testing, although measures were taken to reduce this risk.

In contrast to the impact analysis findings, most reception teachers and TAs interviewed as part of the case studies felt that there had been improvements in children's reading, increased confidence, and increased engagement with phonics, such as children using contextual information on the page or sounding out new or difficult words. However, in the endline survey, when teachers and TAs were asked to report whether they felt that children's engagement with phonics had increased over the year, there was no difference between control and intervention groups. As the aim of the Flexible Phonics programme is to build on existing phonics teaching and further extend phonics skills, it is possible that the programme enriches children's phonics skills but that the benefits are not substantial enough to create a significant difference when compared with general phonics teaching. In England, teaching phonics in reception class and, most recently, using a validated systematic synthetic phonics (SSP) programme has been recommended by the government as good practice for several years, and phonics teaching is well established in English schools. Previous studies that have found significant effects of the direct mapping and set for variability approaches took place in other English-speaking countries where phonics teaching may not be as firmly established in the curriculum.

It is also worth noting that one of the key elements of Flexible Phonics, set for variability, introduces advanced strategies for reading words with an irregular phoneme to grapheme mapping and so requires a good understanding of general phonics as a basis for this. However, the results of the Mispronunciation Correction Test (MCT) suggests that children

had not reached proficiency using this strategy by the end of the delivery period as no difference in performance was found between children in the intervention and control groups. It is possible that children in the intervention group may continue to develop this skill and that this could possibly lead to increased reading performance in the future by facilitating the reading of new irregular words through strategies and increased confidence to attempt new words.

It is also possible that the context of the Covid-19 pandemic may have affected the level of impact that the Flexible Phonics programme was able to have. Some teachers and TAs interviewed from the case studies reported that they were delayed in their phonics teaching as children in their classes were delayed in their general phonics because of disruptions to teaching. They felt that this might have affected the potential impact of Flexible Phonics as set for variability requires children to have mastered basic phonics skills first and some teachers and TAs were still catching up with general phonics at the start of the Flexible Phonics delivery period. In the endline survey, teachers and TAs reported mixed impacts of the pandemic on children's phonics skills: just over half (52%) thought that some children were delayed but others were at the usual level or above; over a fifth felt that the children were delayed by two to three months (23%; total N = 241). It is difficult to say whether and to what extent the impact of Flexible Phonics on children's outcomes might have changed in a non-pandemic year.

The impact analysis found no differences between different groups of children, such as FSM or children with a low reading score, in terms of how effective the Flexible Phonics programme was in improving their reading performance. Findings from the IPE research suggest that while some staff had concerns that the programme may not work as well with certain groups, there were examples of schools using Flexible Phonics successfully with all the groups identified in these concerns. In the endline survey, most staff felt that Flexible Phonics worked well for advanced or average readers and about half felt that it worked well for struggling readers, but fewer staff felt that it worked well for EAL or SEND pupils. Staff noted that where children were behind in phonics, they had not been able to engage with the set for variability element. However, others described using the strategies for struggling readers and finding these useful. Although Flexible Phonics was intended as a universal intervention, some staff described using it exclusively to target specific groups of learners including both higher performing readers or with children who were struggling with phonics.

Compliance

Two-thirds (67%) of schools delivered the Flexible Phonics programme to the required standard. The compliance impact analysis found that even when Flexible Phonics was implemented as intended (according to assessments made by the delivery team) it did not have a notable effect on the primary outcome, EWR score. However, as the delivery team was unable to visit schools to support and observe delivery, the compliance measure was based on teachers' self-report so there may have been issues around accuracy. In the endline survey, 50% of teachers and TAs surveyed reported that they were using flashcards to teach exception words by sight memorisation, which is in direct contradiction of the Flexible Phonics approach.

The Flexible Phonics training was offered to all staff at schools in the intervention group who taught phonics to children in reception. The vast majority of staff attended all three online training sessions or were sent catch-up videos to watch (98%; total N = 311). Most felt that the training was high quality and interesting and that they felt able to teach the Flexible Phonics programme with their class. However, a few felt that the training included quite a lot of theory and several would have liked to have had more practical demonstrations, preferably with children. As the training took place during the partial closure of schools, most schools were unable to try out delivering elements of the programme with their class between training as the delivery team originally intended. Some support partners from the delivery team observed that sometimes TAs had not attended the training or that their attendance at support sessions had dropped-off over time. Furthermore, a couple of support partners noted that TAs sometimes spoke or engaged less than teachers in the support sessions. However, one support partner was able to organise specific sessions for TAs only to facilitate engagement. For future delivery of Flexible Phonics, the delivery team may wish to consider this or other possible ways of engaging with schools to facilitate TA voice and engagement in training and support sessions.

Fidelity

The Flexible Phonics programme consists of five strands of work and schools are expected to teach only those strands that are appropriate to the needs of the children being taught. However, in order to be compliant, schools were expected to deliver the 'GPCs and direct mapping' component and one or both of the 'set for variability' or 'mispronunciation correction' strands. The delivery team found that the majority of schools in the intervention group (60% to 70%) were

delivering both the direct mapping element and one of the set for variability components. This compliance rate may have been affected by delays in children's language and communication development because of the Covid-19 pandemic. Most teachers and TAs observed that at least some of the children in their class had delays in language when they returned to school after the partial closures. Some schools described focusing on general phonics and preferring to wait before introducing the set for variability or mispronunciation correction strands of the programme, and one school expected to be catching up on general phonics until the end of the summer term. It is possible that compliance may have been higher in a non-pandemic year. However, it may still be helpful in future delivery to emphasise that the direct mapping and set for variability and mispronunciation correction strands are core elements of the programme that should be delivered, if possible, as some school staff interviewed were unsure which aspects of the programme were required or most important.

Flexible Phonics was intended to be a universal intervention appropriate for children of all phonics abilities and most schools used it in this way. However, some schools reported using the intervention only with a target group such as high-ability learners, those with lower phonics abilities, or SEND students. Although the training and materials state that the programme is suitable for all children and describe approaches for working with children at different levels of phonics development, more work may be needed to ensure that the programme is taught to all children in reception. It may be helpful to present examples during the training of how schools have incorporated Flexible Phonics to work with all children in reception across different groups. For example, some schools reported using it as part of their usual phonics sessions and then giving extra Flexible Phonics teaching to specific groups such as high- or low-performing pupils or those with SEND and EAL.

Implementation

In the IPE, the majority of intervention schools indicated that they found it fairly straightforward to incorporate the Flexible Phonics activities and strategies with their usual phonics approach. The programme was designed to be fairly flexible, so schools were able to make small adaptations, such as changing the suggested order that sounds were taught to the order used in their existing phonics teaching, to facilitate integration with their existing phonics teaching. Only one school chose not to engage with the Flexible Phonics programme and preferred to continue delivering only its usual phonics approach as it found that to be very effective and did not want to risk reducing this effectiveness. However, another school needed to be reassured that teaching Flexible Phonics alongside its usual approach was consistent with government guidance for schools in England around using a validated systematic synthetic phonics (SSP) programme. In future, it may be helpful to make explicit during recruitment and training that Flexible Phonics is compatible with this government guidance and that GPCs can be taught in the order of their existing programme. With regard to existing phonics teaching approaches, intervention schools used a range of programmes for their general phonics teaching and there was no systematic evidence of Flexible Phonics being incompatible with any particular programme. In addition to this, some school staff and some delivery team support partners observed that the Sounds Write phonics programme had some similarity to Flexible Phonics in that it also did not teach exception words through sight-learning but by teaching that tricky words had tricky sounds in them.

As may be expected, disruptions to education and restrictions associated with the Covid-19 pandemic caused challenges for implementing Flexible Phonics. Some schools reported that some pupils were behind with their general phonics compared to a typical year, which meant that some aspects of Flexible Phonics, such as set for variability, could not be introduced until later in the school year. Restrictions also made sharing resources between classes, such as books, more difficult and there was disruption to teaching as pupils and staff had to isolate if they tested positive for Covid-19. With respect to other challenges for the Flexible Phonics programme itself, the main challenge cited was the initial planning and time needed to work out how best to integrate Flexible Phonics into existing teaching or to add in a daily direct mapping reading session if needed. However, the amount of time varied by school and no school indicated that it felt that this was impossible or too unreasonable. Most schools indicated that after some initial discussions among the reception class teams, they were able to implement the programme without requiring large amounts of time each week, for example, up to two hours. One further challenge identified by several staff during interviews was that for some of the direct mapping books, there were not many instances of the target phoneme it was being used for. For future delivery, it may be useful to review the direct mapping books and revise the book list and mapping document accordingly. However, overall, school staff were mostly positive about the books provided and the way children in their class engaged with them, although one teacher noted that some children may already be familiar with some of them.

The costs for the Flexible Phonics programme were also relatively low with direct mapping books, a training manual, and access to an online portal where they could share resources among schools all provided for free. For the most part, schools described using or adapting existing phonics teaching resources to teach the Flexible Phonics strategies. A couple of staff members described voluntarily purchasing resources such as Lego bricks so that they could label these with letters and use them to practically demonstrate how sounds fit together but they noted that pieces of card or paper could be used for this instead. In the current study, the cost of books, the online portal, and printing manuals were covered by either the EEF grant or Institution of Education UCL resources; in a future scale-up there would need to be a programme fee charged to schools to pay for this. However, our estimated costs for these suggest that even with a programme fee, Flexible Phonics would remain a low-cost intervention.

Changes to teacher and TA practice

As part of the IPE, the endline survey with teachers and TAs explored any changes to practice using activities regarding phonics, reading comprehension, and spelling and handwriting drawn from the 2017 Ofsted report 'Bold Beginnings: The Reception Curriculum in a Sample of Good and Outstanding Primary Schools'. Statistical tests found no difference between practice in the intervention and control groups with regard to these activities. The case studies, however, did find that most schools were incorporating aspects of Flexible Phonics in their daily phonics teaching. Focusing on the core elements of Flexible Phonics for compliance, most staff reported that they were implementing the direct mapping element and some schools were using texts other than the books provided as part of their delivery of this element. The majority were also using mispronunciation correction activities to introduce children to the idea of trying to correct a deliberately mispronounced word. Less than half of case study schools had introduced set for variability when reading exception words. Staff at some case study schools also described teaching other strategies such as 'flipping sounds' where a grapheme mapped to more than one sound, and several case study schools reported that they now embedded phonics activities throughout the day such as playing phonics-based games when lining up and when writing on the whiteboard or reading with the class. Most teachers and TAs interviewed intended to continue with the Flexible Phonics activities they were using as they felt that they worked well and that children engaged with them well. It is worth noting that the practices and behaviours explored in the endline survey were quite general phonics activities, such as how often they undertook activities involving reading exception words, so while the Flexible Phonics programme did not result in high level, broad changes to phonics teaching, qualitative evidence from the case study schools suggests that teachers and TAs did incorporate activities and strategies from the programme into their phonics teaching.

With respect to staff confidence, there was no difference between the intervention and control groups in the endline survey as to whether they felt more or less confident teaching phonics compared to the start of the year. The Flexible Phonics intervention aimed to enrich the understanding of staff and their confidence in phonics by introducing two novel strategies—direct mapping and set for variability—as well as other approaches and advice for teaching general phonics, which incorporated the latest findings from the research literature. These additional strategies were intended to facilitate teachers' and TAs' ability to be flexible and adaptive when teaching phonics. In the baseline survey at the start of the year, the vast majority of teachers and TAs participating in the trial agreed that they felt confident in their ability to teach phonics or support phonics teaching with reception-aged children suggesting that teachers and TAs in reception were already quite confident at teaching phonics and that the Flexible Phonics programme neither increased nor decreased this. The government encourages schools in England to teach systematic synthetic phonics programmes; phonics teaching is thus well established and so it is possible that confidence in teaching phonics is at ceiling. In this case, it is possible that the Flexible Phonics training did positively impact confidence but could not be measured as scores were already at ceiling.

Covid-19

The context of the Covid-19 pandemic affected delivery of the Flexible Phonics training, the length and timing of the period when Flexible Phonics was implemented in schools, how phonic was taught in schools (in some cases), staff and student absences, and the needs of children in reception class.

Schools in England were partially closed from 5 January to 8 March 2021 because of the pandemic and only delivered in-person teaching to the children of key workers or vulnerable children during this time. In response to this, the Flexible Phonics training sessions were changed from two days of in-person training to three half-days online, which reduced the need and, therefore, costs for staff cover and travel expenses. While some schools attempted to teach some Flexible Phonics strategies while teaching remotely in January and February, many did not start teaching Flexible Phonics until

8 March when face to face teaching resumed with whole classes again. For this reason, the intervention delivery time was adapted to run until mid-June. Overall, this reduced the delivery time by approximately one and a half months, but the delivery team felt that alongside the remote teaching and planning undertaken in schools, this should still have been a sufficient time period for the intervention to elicit an effect. In the endline survey of teachers and TAs, the majority felt that at least some children in their class were behind in phonics to some extent and a minority were unsure whether there had been an impact. It is possible that this may have affected the impact of Flexible Phonics on children's reading outcomes. In qualitative interviews with case study schools, some staff reported being unable to teach 'mispronunciation correction' or 'set for variability' strategies as children in their class were still catching up with general phonics. In addition to this, there were disruptions to teaching phonics such as changing from teaching phonics in streamed ability groups across reception year to a mixed-ability, whole-class teaching approach because of the need to maintain staff and student 'bubbles' and children or staff needing to isolate for two weeks if they tested positive for Covid-19.

Existing evidence

This study did not find evidence that teaching direct mapping and set for variability, that is, Flexible Phonics, were more effective at supporting reading skills than best practice phonics teaching when taught to children aged four to five in reception class in England. This study, therefore, did not replicate the findings of Savage et al. (2018) who found that this approach taught to struggling readers aged five to seven in Canada had resulted in improved reading skills, with additional positive outcomes for spelling and reading. As phonics teaching in Canada starts later, Professor Savage, the study author and project director for the delivery team, felt that children in England aged four to five would be at a similar developmental phase of reading but it is possible that the age difference may have affected the level of impact. Previous studies have shown that explicitly linking phonics learning with a relevant reading task (Hatcher et al., 1994, 2004, 2006; Shapiro and Solity, 2008) or direct mapping (Chen and Savage, 2014) was more effective than regular phonics teaching or a vocabulary learning task. While there was no significant impact from the intervention, some teachers and TAs commented that for some books, the linked sound did not actually occur very frequently in the text, and a few teachers and TAs commented that they already read books with children as part of their existing phonics programme. Either of these factors could have reduced the potential effect of Flexible Phonics.

Further evidence from the literature suggests that teaching set for variability is more effective than standard phonics teaching for reading irregular words (Dyson at al., 2017; Zipke, 2016) or reading all words including those with regular pronunciation (Elbro et al., 2012; Elbro and de Jong, 2017; Kearns at al., 2016; Steacy et al, 2019). Although the intervention had no impact, some teachers and TAs at case study schools described children being more willing to attempt reading new words, trying different approaches, and being less concerned about getting the word 'wrong'. It is also worth noting that set for variability is an advanced strategy that is taught after children have learned general phonics. The majority of teachers and TAs reported that either some or most of the children in their class were delayed in their language and communication development and one school reported that they were still focusing on catching up on general phonics until the end of the academic year whereas, typically, schools would aim for most general phonics teaching to be completed by January. It may be the case that there was a delay in learning set for variability strategies for a large number of children in the trial, which will have limited the potential impact of a core element of the programme.

Finally, as the trial took place during 2020/2021, the context of the Covid-19 pandemic may also have affected the programme's level of impact: partial school closures, for example, reducing the delivery period from five to three and a half months, may have contributed to this. In addition, research on the impact of the pandemic on primary school children has consistently found pupil delays across the board by several months (Rose et al., 2021). Particularly relevant here, research by Bowyer-Crane et al. (2021) found that children starting school in autumn 2020 were reported by approximately 90% of schools to have delays in communication and language, literacy, and personal and social development, which aligns with the delays in language and communication development for the pupils in this evaluation reported by teachers and TAs. Most recently, the 2021/2022 Year 1 phonics screening check confirmed an impact on language skills for this cohort at the end of Year 1 as well: the check found that the percentage of children meeting the expected standard had reduced from 82% in 2019 to 75% in 2022 (DfE, 2022b). Therefore, it would be reasonable to expect that children in this cohort may struggle with some of the more sophisticated techniques in Flexible Phonics and hence not see the full benefit of the intervention. In addition, the wider context of the pandemic may have impacted on the personal, social, and emotional development of the children, which could also relate to lower concentration, lower attendance, and lower attainment outcomes.

Limitations and lessons learned

Limitations of the trial

As the trial ran during the 2020/2021 academic year, the Covid-19 pandemic had a substantial impact on the day to day running of schools, the everyday lives of families, the training, support, and delivery of the Flexible Phonics programme, and the delivery of the evaluation. Findings should be interpreted with some caution as the disruption to children's education and development, as well as to the delivery of the Flexible Phonics programme, may have affected the level of impact achieved and it is possible that findings may have been different during a normal year. In addition to reducing the delivery period from five months to three and a half and switching training from face to face to online, Covid-19 restrictions that limited access to schools also meant that the delivery support team and also the evaluation team were unable to enter schools to observe the delivery of Flexible Phonics in practice. This meant that both delivery and evaluation teams were reliant on educator's self-reported descriptions of how they delivered the intervention and how they perceived children's engagement with the approach, which may have limited the delivery team's ability to fully support the needs of the teachers and schools delivering Flexible Phonics or accurately assess the degree of compliance in delivering the programme. It also provided the evaluation team with a slightly limited and less objective picture of how schools were delivering Flexible Phonics as well as missing out on contextual information from the school and local environment. However, it is reasonable to believe randomisation would balance the impact of the Covid-19 pandemic across intervention and control groups. It is unlikely that the trial arms were affected by the pandemic in a differential way that would change the result.

The Covid-19 pandemic also impacted the collection of child assessment data to some extent. The YARC and MCT assessments with children at pre- and post-intervention were mostly able to happen in person but there were a small number of cases where schools requested that assessments take place online. The need for classes (teachers and pupils) to isolate at home when someone in the class tested positive for Covid-19 did mean that some pupil data was not able to be collected as it was not possible to return to test the class within the assessment period. To mitigate this, the assessment team planned in extra time towards the end of the assessment period to return to schools or allow for extra visits if lots of children were missing during previous visits. This minimised the potential impact of Covid-19 related cancellations on the level of attrition in the study. However, the impact analysis identified that Flexible Phonics was more effective at schools that had remote testing, which was an unexpected result. Some possible reasons for this are discussed in the earlier Child Outcomes section but it is also possible that children may have been influenced by reception staff being present in the room during online testing, although measures were taken to reduce this risk. Staff were instructed to sit out of view behind the child and refrain from helping the child during testing, and assessors were instructed to look out for, and address, any instances of this.

As discussed earlier in the Child Outcomes section, the impact analysis found an effect based on whether schools had participated in the EEF effectiveness trial for the Nuffield Early Language Intervention that was running concurrently with the Flexible Phonics trial. The sample for the NELI effectiveness trial focused on children struggling with spoken language in reception class and the geographical area overlapped with the Flexible Phonics trial in Greater London. Around two fifths of schools in the trial (48) were also participating in the NELI intervention and the proportion of schools affected was roughly similar for both the intervention group (42%, 25 of 59) and the control group (38%, 23 of 61) so any effects of participating in the NELI trial should have affected both groups more or less equally. Without further research, it is not possible to say whether the increased effect of Flexible Phonics for schools also participating in the NELI programme was because of the focus on struggling learners, the strategies and activities taught in the NELI programme, or because schools that participated in both interventions may have had greater motivation or resources for language support. If the effect was not driven by school motivation or resource, it would be helpful to identify which aspects of the NELI programme worked to complement and support the delivery of Flexible Phonics.

In terms of the impact analysis, attrition was 20%, which despite not being as large as some recent studies was slightly larger than anticipated. This is likely due to the Covid-19 pandemic causing pupil absenteeism. Reassuringly, there were no differences in the rate of attrition between the intervention and control groups. However, it is important to note, a strong assumption is made that this missingness is random. The missing data analysis shows this might not be the case. Though the trial arm has no correlation with a missing post-test score, there is a statistically significant association found between eligibility for free school meals and a greater likelihood of having a missing post-test score. Although FSM students may be at higher risk of attrition in other studies, the Covid-19 pandemic is likely to have exacerbated this. If

this group of children stand to benefit the most from the intervention, then the results may be understated. To attempt to address the concern that missing values may have a bearing on the results, a sensitivity analysis, in which the missing scores are imputed, finds that the main result does not change.

There were some concerns regarding the suitability of the YARC as a pre-test measure for this age group after pre-test data collection was completed. The pre- and post-test score histograms showed considerable 'floor effects'. The concentration of scores around zero and smaller variability in pre-test scores may have made it a poor measure of students' ability at baseline. This could be a reason for the similar results found when performing a sensitivity test that re-estimated the main analysis without controlling for the pre-test score. Some of these low scores could possibly be attributed to delays in language observed during the pandemic. However, it was hoped that the test would have been more resilient to this phenomenon.

Considerations for future delivery of the programme

While the Flexible Phonics programme was delivered in exceptional circumstances, which involved significant changes to how the programme was delivered as well as children's learning and home environments, the are some general points that could be considered for any future delivery.

There was evidence that some TAs were not able to attend some of the training and support sessions so the delivery team may need to continue to find ways to communicate to, and engage with, schools regarding the importance of TAs being able to attend and of their role in effective delivery. Further to this, sometimes TAs engaged less during support sessions where teachers present tended to speak up more. The delivery team could consider ways to support TA engagement by, for example, offering separate support sessions for TAs.

While most teachers and TAs reported that they understood how the Flexible Phonics intervention worked and the underlying language processes, some felt that the language used in the training could have included less technical language or theory. Several teachers and TAs suggested that practical demonstrations with children or videos of elements being delivered with children would have been very helpful. This was not possible during the trial due to Covid-19 and most schools were unable to try out strategies with children in their class between training sessions because of school partial closures but practical demonstrations would be a valuable addition to future delivery. A few staff also felt that the training was quite long. The delivery team may wish to consider the balance of theory, practical elements, and length for future delivery.

During the trial, some schools needed reassurance on how Flexible Phonics fitted in with existing phonics programmes or requirements for phonics teaching in England. When recruiting schools in future, the delivery team could highlight that Flexible Phonics is a supplementary intervention that is delivered alongside a school's usual phonics programme and is compatible with government requirements to use a systematic synthetic phonics programme. They could also consider reassuring schools at this point that they may continue to teach sounds in the order specified in their current phonics programme.

A minority of staff and schools continued to teach some aspects of phonics in a way that conflicted with the Flexible Phonics approach, for example, teaching some pupils exception words by sight, and a few staff expressed uncertainty around which elements of the programme were compulsory. Future training and follow-up support could provide more emphasis and clarity around what aspects of the programme are mandatory. Also, some teachers and TAs would have liked a member of the delivery team observe their delivery of Flexible Phonics in person so they could receive feedback on whether they were delivering the activities correctly. This had not been possible due to pandemic restrictions, but the delivery team could consider reinstating this aspect in future delivery.

Finally, the delivery team could review some of the Flexible Phonics resources to increase their effectiveness. Some staff felt that some of the books provided for direct mapping did not closely match the target sounds listed. The delivery team checked the mapping following queries and the sounds were present but may not have appeared as frequently as staff might be used to in reading scheme books. The delivery team could consider reviewing the book list and including additional books with higher frequencies of some sounds. There was also little engagement from teachers and TAs with the UCLeXtend portal so future delivery could focus on ways of sharing peer learning and resources that were more effective, for example, sharing via support partners or newsletters. The UCLeXtend portal was used during the trial to

prevent control schools from accessing the intervention materials and to ensure privacy on discussion boards but in future, materials could be hosted on an open website, which would remove the barrier of needing to log-in.

Future research and publications

The next step for the Flexible Phonics efficacy trial was planned to be longitudinal analysis using Year 1 phonics screening check data for the children that participated in the trial in order to test for any effect of the Flexible Phonics intervention over time. However, it was decided in collaboration with the EEF not to continue with this aspect of the research after reviewing findings from the main impact analysis. The phonics screening check would have been an imperfect measure of the added benefit of the intervention as it measures the decoding of regular words, whereas the intervention focuses on the words that break the phonic rules; however, it is the only freely available phonics measure that can be collected systematically across all schools, which is why it was originally chosen.

As the Covid-19 pandemic impacted on the delivery of the Flexible Phonics programme, it may be worth exploring either the programme as originally intended (face to face delivery over five months) or a hybrid version with some online and some face to face training and support as considered by the delivery team for future delivery (also over five months). Ideally, a future study would take place at a time less affected by the Covid-19 pandemic when children's language and communication development has been less impacted by disruption to their education and when their personal, social, and emotional development has been less impacted by disruption to their home and family lives, but this is unlikely to be possible for many years. There would be little value in exploring the Flexible Phonics online delivery model used during this trial in a non-pandemic time as the delivery team did not intend to use a completely online model of delivery in the future. However, testing the online model with the original dosage of five months would enable the evaluation team to establish whether the reduced dosage impacted the effect of the programme as it is currently not possible to speculate from the current study given the context of the Covid-19 pandemic.

Similarly, as around two fifths of schools in the trial (48) were also participating in the EEF Nuffield Early Language Intervention and the intervention targeted children in reception who were struggling with their spoken language, this may also have affected the results of the current trial so it may be beneficial to evaluate the Flexible Phonics programme without this possible competing effect. However, the Department for Education has made £8 million available for state-funded schools in England with reception pupils to access NELI in 2021/2022 (DfE and Ford, 2021) so it may not realistically be possible to do this.

One further challenge to a future evaluation of Flexible Phonics may be the updated essential core criteria for systematic synthetic phonics (SSP) programmes that were published in April 2021 and subsequently updated (DfE, 2023). The updated criteria include elements which are similar to key features of the Flexible Phonics programme. Previously, the set for variability strategies in Flexible Phonics stood in contrast to some existing approaches such as learning exception words 'by sight' whereas now, SSP programmes must ensure that children are taught to 'decode and spell common exception words'. Further to this, SSP programmes are expected to use decodable reading material closely matched to the level of childrens' phonics attainment so that they can practise. While the direct mapping approach specifies that when children learn a new grapheme-phoneme correspondence they should read a text containing that GPC on the same day, it is possible that validated SSP programmes will lead to similar practice. As schools in England are expected to use the new essential core criteria SSP programme for teaching reading, it is possible that schools across England will be using some approaches that are similar to some aspects of the Flexible Phonics programme in future, which would make it harder to detect an effect.

It is the intention of the project and evaluation teams to publish these findings, alongside accounts of delivering and evaluating the Flexible Phonics programme, as an educational book aiming to demonstrate how a randomised controlled trial in education works in practice and explore the complexities of this for a mixed readership of researchers, students, and educators and practitioners. The delivery and evaluation teams may explore specific aspects of language processing for joint publication in future.

References

Bowyer-Crane, C., Bonetti, S., Compton, S., Nielsen, D., D'Apice, K. and Tracey, L. (2021) 'The Impact of Covid-19 on School Starters: Interim Briefing 1. Parent and School Concerns About Children Starting School', London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Impact_of_Covid19_on_School_Starters_-_Interim_Briefing_1_-_April_2021_-_Final.pdf

Camilli, G., Vargas, S., Ryan, S. and Barnett, W. S. (2008) 'Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development', *Teachers College Record*, 112 (3), pp. 579–620.

Chen, V. and Savage, R. S. (2014) 'Evidence for a Simplicity Principle: Teaching Common Complex Grapheme-Phonemes Improves Reading and Motivation in At-Risk Readers', *Research in Reading*, 37, pp. 196–214. doi:10.1111/1467-9817.12022

DfE (2010) 'The Importance of Teaching: The Schools White Paper 2010', White Paper, November 2010, Department for Education.

DfE (2011) *New phonics check will identify thousands of children needing extra reading help*, Department for Education. https://www.gov.uk/government/news/new-phonics-check-will-identify-thousands-of-children-needing-extra-reading-help

DfE (2014) Guidance: *Phonics: Choosing a Programme*, Department for Education: **https://www.gov.uk/government/collections/phonics-choosing-a-programme**

DfE (2017) 'Statutory Framework for the Early Years Foundation Stage' (March 2017), Department for Education.

DfE (2018) 'Schools, Pupils and Their Characteristics: January 2018', *Statistical Publication*, 28 June 2018, Department for Education.

DfE (2019) 'Phonics Screening Check and Key Stage 1 Assessments in England, 2019' (26 September 2019), Department for Education: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/851296/Phonics_screening_check_and_key_stage_1_assessments_in_England_2019.pdf

DfE (2019) National Pupil Database. [data series]. 2nd Release. UK Data Service. SN: 2000108, DOI: http://doi.org/10.5255/UKDA-Series-2000108

DfE (2021) Guidance: *Development Matters*, Department for Education: https://www.gov.uk/government/publications/development-matters--2

DfE (2022a) Guidance: *Choosing a Phonics Teaching Programme*, Department for Education: https://www.gov.uk/government/publications/choosing-a-phonics-teaching-programme/list-of-phonics-teaching-programmes

DfE (2022b) 'Academic Year 2021/22: Key Stage 1 and Phonics Screening Check Attainment', Department for Education: https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-1-and-phonics-screening-check-attainment/2021-22

DfE (2023) Guidance: *Validation of Systematic Synthetic Phonics Programmes: Supporting Documentation*, Department for Education: https://www.gov.uk/government/publications/phonics-teaching-materials-core-criteria-and-self-assessment/validation-of-systematic-synthetic-phonics-programmes-supporting-documentation

DfE and Ford, V. (2021) *Every School with Reception Class Offered Early Language Training*, Department for Education press release: https://www.gov.uk/government/news/every-school-with-reception-class-offered-early-language-training

Dyson, H., Best, W., Solity, J. and Hulme, C. (2017) 'Training Mispronunciation Correction and Word Meanings Improves Children's Ability to Learn to Read Words', *Scientific Studies of Reading*, 21, pp. 392–407. doi:10.1080/10888438.2017.1315424

EEF (2017) 'Improving Literacy in Key Stage 1', London: Education Endowment Foundation.

EEF (2019a) 'Implementation and Process Evaluation Guidance for EEF Evaluations', London: Education Endowment Foundation:

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

EEF (2019b) 'Cost Evaluation Guidance for EEF Evaluations', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/Cost_Evaluation_Guidance_2019.12.11.pdf

Elbro, C., de Jong, P. F., Houter, D. and Nielsen, A.-M. (2012) 'From Spelling Pronunciation to Lexical Access: A Second Step in Word Decoding', *Scientific Studies of Reading*, 16, pp. 341–359. doi:10.1080/10888438.2013.8119356

Elbro, C. and de Jong, P. F. (2017) 'Orthographic Learning is Verbal Learning: The Role of Spelling Pronunciations', in K. Cain, D. Compton, and R. Parrila (eds), *Theories of Reading Development*, Amsterdam: John Benjamins (pp. 169–190).

Galuschka, K., Ise, E., Krick, K. and Schulte-Körne, G. (2014) 'Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities: A Meta-Analysis of Randomized Controlled Trials', *PLoS ONE*, 9 (2): e89900.

Gibson, E. J. (1965) 'Learning to Read: Experimental Psychologists Examine the Process by Which a Fundamental Intellectual Skill is Acquired', *Science*, 148, pp. 1066–072. doi:10.1126/science.148.3673.1066

Hatcher, P. J., Hulme, C. and Ellis, A. W. (1994) 'Ameliorating Early Reading Failure by Integrating the Teaching of Reading and Phonological Skills: The Phonological Linkage Hypothesis', *Child Development*, 65, pp. 41–57. doi:10.1111/j.1467-8624.1994.tb00733.x

Hatcher, P. J., Hulme, C., Miles, J. N. V., Carroll, J. M., Hatcher, J., Gibbs, S., Smith, G., Bowyer-Crane, C. and Snowling, M. J. (2006) 'Efficacy of Small Group Reading Intervention for Beginning Readers with Reading-Delay: A Randomised Trial', *Child Psychology and Psychiatry*, 47, pp. 820–827. doi:10.1111/j.1469-7610.2005.01559.x

Hatcher, P. J., Hulme, C. and Snowling, M. J. (2004) 'Explicit Phoneme Training with Phonic Reading Instruction Helps Young Children at Risk of Reading Failure', *Child Psychology and Psychiatry*, 45, pp. 338–358. doi:10.1111/j.1469-7610.2004.00225.x

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K. (2016) 'Implementation and Process Evaluation (IPE) for Interventions in Education Settings: An Introductory Handbook', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Guidance_Final.pdf

Husain, F., Wishart, R., Marshall, L., Frankenberg, S., Bussard, L., Chidley, S., Hudson, R., Votjkova, M. and Morris, S. (2018) 'Family Skills: Evaluation Report and Executive Summary', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Family_Skills.pdf

Kearns, D., Rogers, H. J., Koriakin, T. and Al Ghanem, R. (2016) 'Semantic and Phonological Ability to Adjust Decoding: A Unique Correlate of Word Reading Skill?', *Scientific Studies of Reading*, 20, pp. 455–470. doi:10.1080/10888438.2016.1217865

Kohnen, S., Banales, E. and McArthur, G. (2020) 'Videoconferencing Interventions for Children with Reading and Spelling Difficulties: A Pilot Study', *Telemedicine and e-health*. doi:10.1089/tmj.2020.0061

Lovett, M. W., Lacerenza, L., Steinbach, K. A. and De Palma, M. (2014) 'Development and Evaluation of a Research Based Intervention Program for Children and Adolescents with Reading Disabilities', *Perspectives on Language and Literacy*, 40, pp. 21–29.

Ofsted (2017) 'Bold Beginnings: The Reception Curriculum in a Sample of Good and Outstanding Primary Schools', Report no. 170045, November 2017.

Robinson Smith, L., Menzies, V., Cramman, H., Wang, Y., Fairhurst, C., Hallett, S., Beckmann, N., Merrell, C., Torgerson, C., Stothard, S. and Siddiqui, N. (2019) 'Easypeasy: Learning Through Play, Evaluation Report', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EasyPeasy.pdf

Rose, S., Twist, L., Lord, P., Rutt, S., Badr, K., Hope, C. and Styles, B (2021) 'Impact of School Closures and Subsequent Support Strategies on Attainment and Socio-Emotional Wellbeing in Key Stage 1: Interim Paper

2, London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/nfer-impact-of-school-closures-and-subsequent-support-strategies-on-attainment-and-socioemotional-wellbeing-in-key-stage-1

Savage, R. S. (2019) 'Literacy, Basic Processes, and Interventions for Struggling Readers', in *Oxford Research Encyclopedia of Education*, Oxford University Press: http://dx.doi.org/10.1093/acrefore/9780190264093.013.925

Savage, R. S., Carless, S. and Ferraro, V. (2007) 'Predicting Curriculum and Test Performance at Age 11 Years from Pupil Background, Baseline Skills and Phonological Awareness at Age 5 Years', *Child Psychology and Psychiatry*, 48, pp. 732–739. doi:10.1111/j.1469- 7610.2007.01746.x

Savage, R., Georgiou, G., Parrila, R. and Maiorino, K. (2018) 'Preventative Reading Interventions Teaching Direct Mapping of Graphemes in Texts and Set-for-Variability Aid At-Risk Learners', *Scientific Studies of Reading*, 22 (3), pp. 225–247. doi:10.1080/10888438.2018.1427753

Savage, R. S., Georgiou, G., Parrila, R., Maiorino, K., Dunn, K. and Burgos, G. (2020) 'The Effects of Teaching Complex Grapheme-Phoneme Correspondences: Evidence from a Dual Site Cluster Trial with At-Risk Grade 2 Students', *Scientific Studies of Reading*, 24, pp. 321-337. doi:10.1080/10888438.2019.1669607

Savage, R. S. and Stuart, M. (2001) 'Orthographic Analogies and Early Reading: Explorations of Performance and Variation in Two Transfer Tasks', *Reading and Writing*, 14, pp. 571–598. doi:10.1023/A:1012052631557

Savage, R. S. and Stuart, M. (2006) 'A Developmental Model of Reading Acquisition Based Upon Early Scaffolding Errors and Subsequent Vowel Inferences', *Educational Psychology*, 26, pp. 33–53. doi:10.1080/01443410500340983

Shapiro, L. R. and Solity, J. (2008) 'Delivering Phonological and Phonics Training Within Whole-Class Teaching', *Educational Psychology*, 78, pp. 597–620.

Sibieta, L., Kotecha, M. and Skipp, A. (2016) 'Nuffield Early Language Intervention. Evaluation Report and Executive Summary', London: Education Endowment Foundation.

Steacy, L. M., Elleman, A. M., Lovett, M. W. and Compton, D. L. (2016) 'Exploring Differential Effects Across Two Decoding Treatments on Item-Level Transfer in Children with Significant Word Reading Difficulties: A New Approach for Testing Intervention Elements', *Scientific Studies of Reading*, 20, pp. 283–295. doi:10.1080/10888438.2016.1178267

Stokes, L., Dawson, A., Williams, C., Alexander, K., Akehurst, G., Manzoni, C., Runge, J. and Xu, L. (2022) 'Tips by Text: Evaluation Report', London: Education Endowment Foundation: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/Tips-by-Text-Final.pdf?v=1651212334

Torgerson, C., Brooks, G., Gascoine, L. and Higgins, S. (2019). 'Phonics: Reading Policy and the Evidence of Effectiveness from a Systematic "Tertiary" Review', *Research Papers in Education*, 34 (2), pp. 208–238. http://www.tandfonline.com/doi/full/10.1080/02671522.2017.1420816

Torgerson, C., Brooks, G. and Hall, J. (2006) 'A Systematic Review of the Research Literature on the Use of Phonics in the Teaching of Reading and Spelling', Department for Education and Skills.

Tunmer, W. E. and Chapman, J. W. (2012) 'Does Set for Variability Mediate the Influence of Vocabulary Knowledge on the Development of Word Recognition Skills?', *Scientific Studies of Reading*, 16, pp. 122–140. doi:10.1080/10888438.2010.542527

Venezky, R. L. (1999) *The American Way of Spelling: The Structure and Origins of American English Orthography*, New York: Guilford.

Yeung, S. and Savage, R. S. (2020) 'Teaching Grapheme-Phoneme Correspondences Using a Direct Mapping Approach for At-Risk Second Language Learners: A Randomized Controlled Trial', *Learning Disabilities*, 53 (2), pp. 131–144. doi: 10.1177/0022219419894563

Zipke, M. (2016) 'The Importance of Flexibility of Pronunciation in Learning to Decode: A Training Study in Set for Variability', *First Language*, 36, pp. 71–86.

Appendix A: EEF cost rating

Appendix table 1: Cost Rating

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

Appendix table 2: Padlock rating

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	MDES	Attrition					
5	Randomised design	<= 0.2	0-10%					
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	0.21 - 0.29	11-20%					
3	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	0.30 - 0.39	21-30%			Adjustment for threats to internal validity [-1]		3
2	Design for comparison that considers selection only on some relevant confounders	0.40 - 0.49	31-40%					
1	Design for comparison that does not consider selection on any relevant confounders	0.50 - 0.59	41-50%					
0	No comparator	>=0.6	>50%					

Threats to validity	Threat to internal validity?	Comments
Threat 1: Confounding	Moderate	0.33 difference in pre-test means, which is 0.05 of an SD (as far as I can calculate from the table, as this isn't clearly reported
Threat 2: Concurrent Interventions	High	Concurrent intervention of NELI highly reported, and evidence of it impacting intervention success
Threat 3: Experimental effects	Low	Low risk of contamination/crossover of intervention
Threat 4: Implementation fidelity	Moderate	Fidelity not well defined, and where definitions are offered, it appears at most 70%. Considerable instance – 29% of delivery as targeted, which is not intended programme, and so risk that programme delivered does not align with ToC
Threat 5: Missing Data	Moderate	Total missing data is moderate (20%), but well accounted for in analyses
Threat 6: Measurement of Outcomes	Low	Low risk of problems due to blinding/marking. Strong validity of chosen outcomes
Threat 7: Selective reporting	Low	No evidence of selective reporting

- **Initial padlock score:** 4 Padlocks – According to the security rating criteria the evaluation fits the 4 padlocks as MDES is 0.21-0.29 and attrition is 20%
- **Reason for adjustment for threats to validity**: 3 Padlocks – score has been adjusted due to threats to validity as the risk of a concurrent intervention (NELI) might have imposed a high risk on the impact of the intervention.
- **Final padlock score:** initial score adjusted for threats to validity = 3 Padlocks

Appendix C: Effect size estimation

Appendix table 3: Effect size estimation

			Intervention group		Control group			
Outcome	Unadjusted differences in means	Adjusted differences in means	n (missing)	Variance of outcome	n (missing)	Variance of outcome	Pooled variance	Population variance (if applicable)
Early Word Recognition raw score	-0.52	-0.05	1,256	80.88	1,283	77.89	79.37	-
YARC	-0.22	-0.02	1,256	11.37	1,283	9.73	10.54	-
MCT	-0.62	-0.04	684	82.53	718	82.93	82.74	-

Appendix D Flexible phonics logic model December 2019 for report 21.04.22

Rationale/need for intervention

Systemic phonics now has a lot of evidence, but there are still ways to refine it further. Recent evidence suggests combining Direct Mapping and Set-for-Variability strategies will help to do this.

1. The first strategy, Direct Mapping, requires children to read texts that include several examples of the GPCs that they have just learned on the same day. In the first instance, these will be carefully selected pre-existing decodable texts, or specifically crafted controlled texts, before real books are introduced slowly and strategically.
2. The second strategy, Set-for-Variability, explicitly teaches pupils to add in another step after they have blended phonemes to graphemes where pupils 'set-for-variability'. This is a metacognitive step, where pupils consider what the word may be, given both the distance between these blended sounds and known words, and potential spelling to sound inconsistencies. This enables children to better recognise all words but can also be especially useful when learning to recognise exception words.

Long term outputs/Impacts

- Improved overall literacy outcomes for pupils
- Improved phonological awareness for pupils
- Increased use of Direct Mapping and set- for variability strategies by pupils

Theory of change

To provide two strategies that Teachers and TAs can use in all Reception phonics teaching through suggested lesson plans to add value to their standard practice which can improve children's literacy outcomes.

Inputs

- Selection of 115 schools from the Greater London
- Pupil details collected (by evaluator for trial only)
- Time needed to attend a one-day initial training, one half-day follow-up training and then for research assistants to attend two follow-up sessions
- Time needed in standard phonics lessons for new strategies
- £400 of specifically chosen books are given to the schools to enable the Direct Mapping strategy

Activities

- All Teachers and TAs attend a one-day initial training and a half-day follow-up training at regional locations
- Teachers and TAs deliver Flexible Phonics strategies across all normal phonics lessons from January–May 2021
- Teachers and TAs receive a manual and additional resources (developed during the pilot phase) to support the lessons
- Ongoing support is provided by email/phone response

Outputs

- Teachers and TAs deliver the Flexible Phonics strategies approximately 3–4 times a week as part of their standard practice
- Pupils use these strategies for everyday reading in school and at home

Short term outcomes/ Mediators

- Teachers/ TAs report more confidence in delivering phonics lessons
- Pupils have greater word reading abilities

Enabling factors/conditions for success

- Flexible phonics strategies may be more or less effective with pupils from low-income families (evaluated by free school meals sub-group analysis)
- Children's baseline attainment may act as a moderator and make the intervention more or less effective for some children (evaluated by examination of low and high baseline assessment sub-group analysis)
- Spill over-effect should not be a problem for this trial as it is cluster randomised but can be confirmed through the IPE
- Geographical region may also have an impact on intervention effectiveness due to different pupil demographics, but this will not be examined in the current trial

Appendix E: School Information Sheet

Flexible Phonics

This project is a collaboration between the Education Endowment Foundation (EEF), the Institute for Employment Studies (IES) and the Institute of Education University College London.

We are seeking state-funded primary schools (reception classes) in the Greater London area to take part in an innovative project testing a programme aiming to improve language and literacy amongst all reception age children. A version of the programme has already been tested in Canada and has been shown to **improve literacy by the equivalent of 2-3 additional months' progress**.

What, how and when?

We will work with reception class teachers and teaching assistants in the classroom delivery of new strategies designed to optimise the teaching of reading to all children. The work fits well around existing phonics programmes that can be delivered broadly as usual. A novel aspect of *Flexible Phonics* is that it gives children more strategies to flexibly read all words, and could be particularly powerful in enabling children to independently read novel exception words (words that break phonic rules such as 'the', 'two', 'between', 'above', etc). Children learn how to use phonics in close conjunction with authentic children's texts to become confident, motivated, readers.

In this project, we are testing the impact of Flexible Phonics on pupil outcomes using a randomised controlled trial. This means that the curriculum will be delivered to half of the schools in the study. These schools will be chosen at random in Winter 2020. Those who are allocated to receive the intervention will receive two half-days and then a further half-day (2 weeks later) of professional development training, a copy of a Teacher Manual and associated resources. They will also receive two in-class follow up visits where we can provide further feedback and guidance and (with teacher's consent) observe the classroom context. Those schools that are allocated to the control condition will not receive professional development training in Flexible Phonics and teachers will carry out their regular teaching as usual. Children will be assessed at the beginning and end of the study, irrespective of their school's allocation. Control condition schools will receive £1000 on completion of the study to thank them for taking part.

We will prepare an information sheet for participating schools to send to parents. Parents will also have the opportunity to withdraw their child from the study at any point.

What are the benefits to my school?

By participating in this study your school will if allocated to the Flexible Phonics intervention, receive:

- **Free** children's books to the value of £400 per school
- **Free professional development training** from Professor Robert Savage, who has run successful reading interventions in the UK, Canada, Australia and Hong Kong. This will be delivered online as three half-day sessions.
- A programme that has the potential to improve pupils' reading attainment and prevent reading difficulties.
- The opportunity for your school to **work with the EEF to build ground-breaking evidence** on what works to improve educational outcomes for all children.

By participating in this study your school will if allocated to the *control* condition, receive:

- £1000 on completion of the study to thank you for taking part.
- The opportunity for your school to **work with the EEF to build ground-breaking evidence** on what works to improve educational outcomes for all children.

What will it cost my school?

Participation is free. All costs will be covered by the EEF.

How much time will it take my school?

The programme itself will be implemented entirely by your regular school staff who will receive professional training. Schools will be required to:

- Send out a parent information letter
- Provide pupil level data
- Enable all reception staff (teacher(s) and TA(s)) to attend 3 half-day sessions of online training (multiple training dates will be made available)
- Facilitate pre- and post-testing in schools, delivered by trained researchers; and the two in-class visits to observe classrooms and discuss the programme with RAs
- Participate in two short online surveys
- The evaluation team may request to visit your school to better understand your experience

How will data sharing work?

Participating schools will be required to provide pupil data (including pupil name, date of birth, gender, unique pupil number) in order to facilitate the evaluation and enable linkage to the National Pupil Database. All personal data collected as part of the study will be treated with the strictest confidence by the project team and processed only in accordance with the requirements of the GDPR and the Data Protection Act 2018. IES have run research studies in more than 100 schools nationwide since the introduction of GDPR in May 2018. Further details about the precise roles of each member of the project team in relation to the processing of personal data has been shared in Memorandum of Understanding signed by schools to formalise their involvement in the trial. It will also be shared in information sheet for parents.

Project team

- The UCL Institute of Education is a world-leading centre for research and teaching in social science and education, ranked number one for education worldwide in the 2020 QS World University Rankings for the seventh year running. In the most recent Research Excellence Framework assessment of university research, 94% of the Institute's research was judged to be 'world class'.
- The Institute for Employment Studies (IES): IES is an independent, international centre of research and consultancy in education, public employment policy and organisational human resource management. They will independently evaluate the impact of the Flexible Phonics programme.
- Education Endowment Foundation (EEF): The EEF is an independent charity dedicated to breaking the link between family income and educational achievement. They run projects which test the efficacy of education interventions to generate new evidence of what works.

How will it be evaluated?

We will use short reading and related language assessment tests and qualitative feedback from teachers and parents to evaluate impact. At least one of your reception classes will receive a pre- and post-test (a standard reading assessment which will take no longer than 10 minutes at pre-test and 20–30 minutes at post-test per child). You will be informed of which class(es) have been chosen for testing in October 2020. External researchers will come into your school to deliver these assessments. **Schools will not have to administer these assessments.**

Key dates/timeline

Month	Activity
Nov 2019–July 2020	Schools sign up to the project with a formal Memorandum of Agreement outlining commitments (first come, first served!)
Autumn 2020	Schools send the information sheet to parents and submit pupil data
Oct/Nov 2020	In school pre-testing of chosen classes (IES)
Winter 2020/2021	Delivery of training to half of the schools begins (UCL)
June 2021	School's assessed delivery of Flexible Phonics intervention ends
June/July 2021	In school post-testing of chosen classes (IES)
Summer 2022	Results of RCT published (EEF)

Next steps

We have limited spots for this project. If you are interested in participating or finding out more, please email: flexiblephonics@ucl.ac.uk

We look forward to hearing from you soon!

Appendix F: Memorandum of Understanding in relation to participation in the Flexible Phonics study

Please sign both copies and complete the information on phonics practice, retaining one for your own records and returning the second copy to **flexiblephonics@ucl.ac.uk** or post to IOE Project Manager Flexible Phonics at: Psychology & Human Development, UCL Institute of Education, 25 Woburn Square, London, W1H0AA.

School name:

_____

School address:

_____

_____

_____

This Memorandum of Understanding (MOU) sets out the roles and responsibilities of schools participating in, and the parties involved in delivering and evaluating, the Flexible Phonics project.

This document is being sent to your school because you have indicated interest in participating in the project.

1. The project team and their roles

The Flexible Phonics project is a collaboration between the Institute of Education at UCL (IOE) and the Institute for Employment Studies (IES) funded by the Education Endowment Foundation (EEF). Together, these parties (with the exception of the EEF) are referred to as the "project team" in this MOU. This MOU will also refer to the Flexible Phonics to be used with children in reception class as the "Intervention".

The EEF are the funders of the project.

IOE will be responsible for designing the Flexible Phonics intervention and training staff in how to deliver Flexible Phonics curriculum reception classes.

IES will evaluate the programme's impact on pupil outcomes. They will also assess whether the programme could be implemented at a larger scale.

2. Communication with parents

All participating schools will be required to distribute an information letter and adjoining privacy notice to parents of pupils entering reception in September 2020. This letter and privacy notice will be provided by the project team and shared with schools shortly.

The letter and privacy notice will inform parents of the nature of the project, the personal data that will be collected about their child and how this data will be processed. It will also give them the opportunity to withdraw their child's data from the project if they wish. Providing parents/carers with the opportunity to withdraw their child's data if they so wish will provide due consideration for their privacy and rights in relation to their data.

Schools should allow parents an initial two weeks to respond if they wish to withdraw from the study. In cases where parents inform the school that they do not wish for their child's data to be used in the project, schools should remove data of these pupils from the data they share with the project team.

Beyond the initial two-week period, parents/carers will be able to withdraw their children's data from the study at any point during the project by contacting IES at flexible.phonics@ies.ac.uk.

Parents/carers will be provided with full details on their rights under data protection laws and contact details for the project team in the information letter (which IOE will provide to schools).

3. The evaluation

The evaluation of the programme is being conducted by IES. The project involves the evaluation of the Flexible Phonics intervention through a randomised controlled trial (RCT), along with an implementation and process evaluation. Schools will be randomly chosen to either receive or not receive Flexible Phonics for their reception classes (roughly a 50:50 split). Using randomly assigned groups is the best way of evaluating if a programme has an impact on pupil outcomes, as the two groups can then be assumed to be the same, therefore any differences in outcomes found can be attributed to the programme.

As part of the evaluation, reception teachers/ TAs will be asked to:

- Complete a short, online survey prior to receiving training on Flexible Phonics and materials, covering existing approaches to teaching phonics as well as broader approaches to teaching reading with reception class children.
- a short online survey towards the end of the programme, focusing on teachers' experiences of the programme and any perceived impacts on pupils and the wider school
- Provide selected information about participating pupils (a template will be provided by the evaluation team)
- Facilitate the reading assessments of pupils prior to the programme (October-December 2020) and towards the end of the programme (June-July 2021)
- Allow researchers in to schools to observe follow- up activities from the training across two of the schools
- Participate in brief, face- to- face interviews at the end of the programme. Staff will be selected from a small number of schools (eight in total from across the region) to participate in this stage of the research

4. Data sharing and data protection

- For the purposes of conducting the evaluation to assess the impact of Flexible Phonics, IES and IOE will both become data controllers of personal data of school staff and pupils obtained from schools and other sources such as the National Pupil Database. They may share personal data with trusted processors such as academics, test administrators, transcribers and research assistants solely for the purposes of proper delivery, management and evaluation of the project. At the end of the project,

data will be submitted to the EEF's data archive. At this point, EEF will become a data controller and the archive manager will be a data processor

- The legal basis for processing data for this project is legitimate interests (IES) and public interest (IOE).
- IES and IOE will securely delete all personal data within six months of the project finishing.
- The Privacy Notice for this project is available to download at **https://www.employment-studies.co.uk/sites/default/files/resources/files/Flexible_Phonics_Privacy_Notice.pdf**
- Pupils will be asked to complete very short reading assessments in October-December 2020 of approx. 10 minutes and short assessments in June-July 2021 of approx. 20-30 minutes. The responses will be collected by an independent test administrator and sent to IES and then shared with IOE, the Department for Education, the EEF's archive manager and, in an anonymised form, with the Office for National Statistics and potentially other research teams
- Further matching to NPD and other administrative data may take place during subsequent research.
- Your school's data will be treated with the strictest confidence and will be transferred securely and saved in secure locations only accessible by the project team in line with GDPR and the Data Protection Act 2018
- We will not use pupil names, teacher/ TA names or the name of your school in any report arising from this project.

5. Responsibilities

a. The project team as a whole will:

- Produce an information sheet and privacy notice for parents which schools will distribute to parents. This will set out the full details of the project and the anticipated personal data processing
- Ensure that any parental withdrawals (once the programme is live) are attended to as quickly as possible

b. IOE will:

- Deliver training and support on Flexible Phonics to staff teaching reception classes that have been selected at random to participate in the intervention as part of the RCT
- Act as the main point of contact for schools and parents for anything to do with the Flexible Phonics intervention.

c. IES will:

- Share a data collection template with schools and assist schools with data collection where necessary
- Conduct the randomisation of schools to the treatment (being taught Flexible Phonics) and control groups (continuing their usual phonics approach)
- Act as the main point of contact for schools regarding data collection
- Facilitate data sharing with other members of the project team as necessary for the implementation and evaluation of the project

- Coordinate with the independent test administrator and schools to help arrange testing in October-December 2020 and June-July 2021
- Carry out online surveys of school staff in autumn 2020 and summer 2021
- Provide an information letter for school staff regarding the project, explaining what will be required of them, providing information on data security and specifying a main point of contact for any questions
- Visit a sample of schools and carry out observations of follow- up visits and case study interviews with staff
- Collect and analyse the data from the project and write up the findings
- Disseminate findings from the study - the final summary report for this project will be shared with all participating schools and will be available online on the EEF's website

d. The schools will:

- Name a 'Project Champion' to serve as the main point of contact for the school with the project team
- Send parents of pupils entering reception in September 2020 the parent information letter and privacy notice (which is being prepared by the project team and will be sent to schools shortly)
- Provide the evaluation team with the data required to evaluate the project (ensuring accuracy of the data and removal of all pupils who have withdrawn from the study)
- Commit time for staff to complete the online surveys in September-October 2020 and June-July 2021
- Commit time for each assessment phase (October- December 2020 and June- July 2021) and liaise with the independent test administrator to find appropriate dates and times for assessments to take place
- Commit all Reception Teachers and TAs time off to attend a 1-day training session and a half-day training session in a local location in early 2021 (multiple training dates will be provided so that not all reception staff are out of school on a given day).
- Allow research assistants from IOE to visit the school at two time points in the month following the training to observe phonics lessons and meet to discuss the programme
- Liaise with the evaluation team and assist in the arrangement of case study and observation visits, if selected to take part, enabling short face to face interviews with relevant staff
- Ensure staff are briefed about the programme and their role in it through distributing a teacher/ TA information letter (which will be provided by the research team) and support them to complete the surveys at the beginning and end of the project
- Inform IOE if the school is taking part in another EEF funded project
- Inform IOE of current phonics practice
- Ensure the shared understanding and support of all staff for the project and personnel involved

e. All parties will:

- Provide such assistance to each other as is reasonably required to enable all parties to comply with requests from parents and pupils who are involved in the project to exercise their rights under data protection legislation

- Comply with EU data protection laws including the General Data Protection Regulation and the data protection laws of the UK including the Data Protection Act 2018
- Use all reasonable endeavours to work together collaboratively and productively, in particular in relation to meeting key dates and timeframes set out in the School Information sheet

6. No partnership or agency

Nothing in this agreement is intended to, or shall be deemed to, establish any legal partnership or joint venture between any of the parties, constitute any party the agent of another party, or authorise any party to make or enter into any commitments for or on behalf of any other party.

7. Binding Terms

Nothing in this document will constitute or evidence a legally binding contract to create legal relations between the Parties.

We commit to participating in the Flexible Phonics study as detailed above

Signature of authorised officer of the School:

_____

Date: _____

Full name: _____

Position: _____

Contact email: _____

Project Champion(s) (if different to authorised officer named above):

Name: _____

Job title/role: _____

Email address: _____

Name: _____

Job title/role: _____

Email address: _____

8. Information on phonics practice

Do you use a specific phonics programme? Is it followed closely?:

_____

_____

Who teaches your phonics lessons? _____

What is the frequency of phonics lessons and how long do they

last?_____

Is streaming used for phonics teaching in Reception and if so with whole class/ small

groups?:_____

Current number of reception classes expected for 20/21 year: _____

Current number of reception pupils expected for 20/21 year: _____

Appendix G: Information for Parents/Guardians

Dear parent/guardian,

[School Name] is taking part in a project called 'Flexible Phonics'. As part of the project, staff at the school may be trained to use the Flexible Phonics approach to teaching reading with children in reception year. The Flexible Phonics approach aims to teach children in reception year a set of strategies to help them with learning to read new words by themselves.

In this project, we will test the effectiveness of the Flexible Phonics approach using a randomised controlled trial (RCT). This means that some schools will be taught using the Flexible Phonics approach and some will be taught phonics in the usual way. Which schools are taught Flexible Phonics and which are taught using the usual phonics approach will be decided at random. Using random selection is the best way to see if a programme has made a difference to pupil outcomes.

The Flexible Phonics project is a collaboration between the Institute of Education at UCL (IOE) and the Institute for Employment Studies (IES) funded by the Education Endowment Foundation (EEF).

What is involved?

- Your child may be asked to complete a short literacy assessment, carried out by trained researchers, in October-December 2020 and June-July 2021. The assessment is designed to be fun and engaging, and we expect that your child will enjoy taking part

What will happen to your child's data?

- In order to run the programme and evaluate its impact, [School Name] will share data on your child with the Research Team
- We take your child's data security very seriously and will be working closely with [School Name] to ensure that data is handled appropriately and securely
- Data will be shared and processed in accordance with GDPR and the Data Protection Act 2018

For the purpose of research, the data provided by schools along with the data from the assessments will be linked with pupil information from the National Pupil Database (NPD), and shared with the Department for Education, the EEF's archive manager and, in an anonymised form, with the Office for National Statistics and potentially other research teams. Further matching to NPD and other administrative data may take place during subsequent research. Anonymised data will be retained by us for 5 years then destroyed by the collaborating organisations.

- If you have any questions about [School Name] sharing data with the Research Team, please contact [Relevant person at school]
- A final report is expected to be published in summer 2022. No individual pupil or school will be identifiable in this report - all details will be fully anonymised

- Please read the Privacy Notice for full details which can be downloaded at **https://www.employment-studies.co.uk/sites/default/files/resources/files/Flexible_Phonics_Privacy_Notice.pdf**

Participating in the study

If you do not want your child's data to be included in the project, please return the form below within two weeks. Your child will still receive the same phonics teaching as everyone else in their class. If you decide you would like to withdraw your child's data after that please email flexible.phonics@ies.ac.uk.

If you would like any further information about the project, you can call email flexible.phonics@ies.ac.uk

The evaluation has been approved by the Ethics Committees of the Institute for Employment Studies. If you require clarification of the ethical approval or have any concerns during the course of the research, please contact Clare Huxley on flexible.phonics@ies.ac.uk

If you DO NOT want your child to take part, please sign and return the attached form to [School Name] by [Date].

If you are happy for your child to participate, you do not need to do anything but please keep this letter for your information.

Yours sincerely,

Dr Anneka Dawson

Co-Principle Investigator, IES

Email: flexible.phonics@ies.ac.uk

Telephone: 01273 763400

Withdrawal Form

Please only return this form to the class teacher if you are NOT willing for your child's data to be used as part of the Flexible Phonics study.

I **DO NOT** give permission for my child's data to be included in the Flexible Phonics study.

Your Child's Name ……..……

Your Name ……..

Your Signature ……..……

Parent / Guardian (delete as appropriate)

Date ……

Appendix H: Privacy notice to parents

Introduction

Your child's school is participating in a research project testing the impact of an early learning programme called 'Flexible Phonics'.

The Flexible Phonics project is a collaboration between the Institute of Education at UCL (IOE) and the Institute for Employment Studies (IES) funded by the Education Endowment Foundation (EEF).Together, these parties (with the exception of the EEF) are referred to as the 'Research Team' in this Privacy Notice. The roles of each party named above are as follows:

- The EEF are the funders of the project
- IOE will be responsible for designing the Flexible Phonics manual and training staff in how to deliver Flexible Phonics with reception classes
- IES will evaluate the impact of Flexible Phonics on pupil outcomes. They will also explore whether Flexible Phonics could be implemented at a larger scale

This privacy notice sets out how the Research Team will collect and use your personal data.

After your child's school has shared personal data of pupils taking part in the project with the Research Team, each member of the Research Team will become a separate data controller of this personal data in order to perform the roles set out above.

If you have any questions about this privacy notice, including any requests to exercise your legal rights in relation to your personal data, please contact IES in the first instance.

IES

- **Post:** Institute for Employment Studies, City Gate, 185 Dyke Road, Brighton, BN3 1TL
- **Email:** **Suzanne.anderson@employment-studies.co.uk**

You also have the right to make a complaint at any time to the Information Commissioner's Office (ICO), the UK supervisory authority for data protection issues (**www.ico.org.uk**). We would, however, appreciate the chance to deal with your concerns before you approach the ICO, so please contact IES in the first instance.

We promise to respect any of your personal information which is under our control and to keep it safe. We aim to be clear when we collect your information about what we will do with it and let you know of any material changes to this notice.

The Research Team deals with and shares your personal data pursuant to a data sharing agreement between the Research Team members. The agreement sets out the purposes for which we may process and share your personal data and our agreement to cooperate to protect your personal data and deal with any requests you may have.

What kind of information do we collect?

Your child's school will share the following information with IES (who will then share it with other members of the Research Team as necessary for them to fulfil their roles):

Pupils:

- Name
- Date of birth
- Gender
- Unique Pupil Number (UPN)
- Class name
- School name

IES will access and link this pupil data to background and school data held on the National Pupil Database (NPD). The NPD data to be requested will include whether or not the pupil is eligible for Free School Meals (FSM) as well as their gender and their phonics score in Year 1. Specifically, we are processing data on FSM eligibility to determine if the programme has a different impact on this group of pupils. EEF was established with a remit to break the link between family background and educational attainment, and analysis of impact for FSM pupils is carried out as part of all EEF evaluations. We are processing data on gender in order to account for gender in the evaluation, given anticipated relationships between gender and outcomes.

IES will match all the above pupil data (both the data collected directly from schools and the data requested from the NPD) to data on pupil outcomes. This will include data from questionnaires and assessments administered as part of the project including a standard assessment of literacy skills and a measure of mispronunciation correction as well as data on outcomes available through the NPD.

What do we do with information we collect?

Each organisation will use the data for different purposes.

IOE

- To deliver the programme to teachers
- To manage communications with teachers
- To act as the main point of contact for schools for anything to do with the intervention

IES

- To conduct the randomisation (i.e., to randomly allocate schools to receive or not receive the intervention)
- To match data received from schools to NPD data and outcome data
- To contact teachers about participating in interviews and surveys
- To instruct and liaise with independent test administrators
- To evaluate the impact and effectiveness of the programme and prepare a report about the project

EEF

- At the end of the project, data will be submitted to the EEF's data archive. At this point, EEF will become a data controller and the archive manager will be a data processor

What is the lawful basis for processing your personal data?

We will only use your personal data where we have a lawful basis for doing so.

- The IOE's lawful basis for processing research data is 'public interest task.' This research will build on existing work by the Department of Psychology and Human Development on phonics and reading development.
- IES' legal basis for processing personal data is 'legitimate interests'. The evaluation of Flexible Phonics fulfils one of IES' core business purposes (undertaking research, evaluation and information activities) and is therefore in our legitimate interest, that processing personal information is necessary for the conduct of the

evaluation. Our condition for processing special category personal data (gender) is that this is necessary for scientific research purposes and is in the public interest.

- When the EEF becomes data controller at the end of the project as the data is submitted to the data archive, the legal basis for processing personal data is 'legitimate interests'. These legitimate interests include gather data about what educational interventions work best, under what conditions, for which pupils, with a view to increasing attainment and reducing educational disadvantage. Special category data is processed for the purpose of scientific research and archiving as permitted under GDPR Article 9 (j).

Who else has access to your information?

The Research Team may disclose your information to third parties in connection with the purposes of processing your personal data set out in this notice. These third parties may include suppliers, research assistants, trusted academic partners and subcontractors who may process information on behalf of the Research Team to carry out such work as administering tests, undertaking interviews, transcribing interviews and archiving data. In all cases, the Research Team will ensure that these third parties enter into appropriate data processing agreements with us and that they keep your personal data secure and confidential.

We may also disclose your personal information if required by law, or to protect or defend ourselves or others against illegal or harmful activities, or as part of a reorganisation or restructuring of our organisations.

International Transfers

Your personal information will not be transferred outside of the EEA.

Security

We take all reasonable steps to protect your personal information and follow procedures designed to minimise unauthorised access, alteration, loss or disclosure of your information.

We have put in place procedures to deal with any suspected personal data breach and will notify you and any applicable regulator of a breach where we are legally required to do so.

Data Retention

We will only retain your personal data for as long as necessary to fulfil the purposes we collected it for, including for the purposes of satisfying any legal, accounting, or reporting requirements. When it is no longer necessary to retain your personal data, it will be securely deleted.

To determine the appropriate retention period for personal data, we consider the amount, nature, and sensitivity of the personal data, the potential risk of harm from unauthorised use or disclosure of your personal data, the purposes for which we process your personal data and whether we can achieve those purposes through other means, and the applicable legal requirements.

- IES and IOE will delete any personal data six months after the completion of the project
- The IOE will retain the anonymised data until March 2025 to permit further analysis
- IES will send the data to EEF's archive hosted by the Office for National Statistics (ONS) Secure Research Service within three months of project completion. The Fischer Family Trust will manage the data on behalf of EEF, at which point EEF take responsibility for Data Protection Compliance

Please note that, under Data Protection legislation, and in compliance with the relevant data processing conditions, personal data can be kept for longer periods of time when processed purely for archiving purposes in the public interest, scientific or historical research, and statistical purposes.

Your legal rights

Under certain circumstances, you have rights under data protection laws in relation to your personal data, including rights:

- To request access to your personal data: this enables you to receive a copy of the personal data we hold about you and to check we are lawfully processing it
- To request correction of your personal data: this enables you to have any incomplete or inaccurate data we hold about you corrected
- To request erasure of your personal data: this enables you to ask us to delete or remove personal data where there is no good reason for us continuing to process it
- To object to processing of your personal data: you can object where we are relying on a legitimate interest (or those of a third party) and there is something about your particular situation which makes you want to object to processing on this ground as you feel it impacts on your fundamental rights and freedoms
- To request restriction of processing your personal data: This enables you to ask us to suspend the processing of your personal data
- To request transfer of your personal data
- To object to direct marketing (including profiling) and processing for the purposes of scientific/historical research and statistics
- Not to be subject to decisions based purely on automated processing where it produces a legal or similarly significant effect on you

If you wish to exercise any of the rights set out above in connection with this research project, please contact **Suzanne.anderson@employment-studies.co.uk**

You will not have to pay a fee to access your personal data (or to exercise any of the other rights). However, we may charge a reasonable fee if your request is clearly unfounded, repetitive or excessive. Alternatively, we may refuse to comply with your request in these circumstances.

We may need to request specific information from you to help us confirm your identity and ensure your right to access your personal data (or to exercise any of your other rights). This is a security measure to ensure that personal data is not disclosed to any person who has no right to receive it. We may also contact you to ask you for further information in relation to your request to speed up our response.

We try to respond to all legitimate requests within one month. Occasionally it may take us longer than a month if your request is particularly complex or you have made a number of requests. In this case, we will notify you and keep you updated.

Changes to this Notice

We may change this Privacy Notice from time to time. If we make any significant changes in the way we treat your personal information we will make this clear by contacting schools and ensuring they provide you with an updated version of this Privacy Notice.

Company details for the Research Team and Funder

- The Institute of Education (IOE) is part of University College London
    - University College London a body corporate established by Royal Charter with company number RC000631 whose principal place of business is at Gower Street, London, WC1E 6BT
- The Institute for Employment Studies is a charitable company limited by guarantee registered in England.
    - Registration number: 931547
    - Charity registration number: 258930
    - Registered office: Institute for Employment Studies, City Gate, 185 Dyke Road, Brighton, BN3 1 TL
    - Email: **suzanne.anderson@employment-studies.co.uk**
- The Education Endowment Foundation is a charity registered in England
    - Registered office: Millbank Tower, 21-24 Millbank, London SW1P 4QP

- Charity registration number: 1142111
- Company number: **07587909**
- Email: info@eefoundation.org.uk

Appendix I: Privacy notice for schools

Introduction

Your school is participating in a research project testing the impact of an early learning programme called 'Flexible Phonics'.

The Flexible Phonics project is a collaboration between the Institute of Education at UCL (IOE) and the Institute for Employment Studies (IES) funded by the Education Endowment Foundation (EEF).Together, these parties (with the exception of the EEF) are referred to as the 'Research Team' in this Privacy Notice. The roles of each party named above are as follows:

- The EEF are the funders of the project
- IOE will be responsible for designing the Flexible Phonics manual and training staff in how to deliver Flexible Phonics with reception classes
- IES will evaluate the impact of Flexible Phonics on pupil outcomes. They will also explore school staff's views on Flexible Phonics and whether it could be implemented at a larger scale

This privacy notice sets out how the Research Team will collect and use your personal data.

After your school has shared personal data of Teachers and TAs taking part in the project with the Research Team, each member of the Research Team will become a separate data controller of this personal data in order to perform the roles set out above.

If you have any questions about this privacy notice, including any requests to exercise your legal rights in relation to your personal data, please contact IES in the first instance.

IES

- **Post:** Institute for Employment Studies, City Gate, 185 Dyke Road, Brighton, BN3 1TL
- **Email:** **suzanne.anderson@employment-studies.co.uk**

You also have the right to make a complaint at any time to the Information Commissioner's Office (ICO), the UK supervisory authority for data protection issues (**www.ico.org.uk**). We would, however, appreciate the chance to deal with your concerns before you approach the ICO, so please contact IES in the first instance.

We promise to respect any of your personal information which is under our control and to keep it safe. We aim to be clear when we collect your information about what we will do with it and let you know of any material changes to this notice.

The Research Team deals with and shares your personal data pursuant to a data sharing agreement between the Research Team members. The agreement sets out the purposes for which we may process and share your personal data and our agreement to cooperate to protect your personal data and deal with any requests you may have.

What kind of information do we collect?

Your school will share the following information with IES (who will then share it with other members of the Research Team as necessary for them to fulfil their roles):

School staff:

- Name
- Role
- Email address
- Class name
- School name

- Whether the staff member teaches phonics to Reception

IOE will collect data on who attends the Flexible Phonics training and they will collate information about support requests for delivery the programme which they will pass on to IES. They will also collect information at the follow-up visits about how Flexible Phonics is being implemented in schools that received the training.

IES will carry out an online survey in autumn 2020 to capture teaching staff's views on teaching phonics and reading in their school and will repeat some of those questions in a further online survey in summer 2021. This survey will also ask teaching staff about their experiences of the Flexible Phonics project for those who attended the training. Finally, some Teachers and TAs will be asked to take part in telephone or face to face interviews which will ask in more detail about their experiences of Flexible Phonics.

IES will match the above data to data on pupil outcomes collected through a reading assessment and a measure of mispronunciation correction as well as data on outcomes available through the NPD.

What do we do with information we collect?

Each organisation will use the data for different purposes.

IOE

- To deliver the programme to teachers
- To manage communications with teachers
- To act as the main point of contact for schools for anything to do with the intervention

IES

- To conduct the randomisation (i.e., to randomly allocate schools to receive or not receive the intervention)
- To match data received from schools to NPD data and outcome data
- To contact teachers about participating in interviews and surveys
- To instruct and liaise with independent test administrators
- To evaluate the impact and effectiveness of the programme and prepare a report about the project

EEF

- At the end of the project, data will be submitted to the EEF's data archive hosted by the Office for National Statistics (ONS) Secure Research Service within three months of project completion. The Fischer Family Trust will manage the data on behalf of EEF, at which point EEF will become a data controller and the archive manager will be a data processor

What is the lawful basis for processing your personal data?

We will only use your personal data where we have a lawful basis for doing so.

- The IOE's lawful basis for processing research data is 'public interest task.' This research will build on existing work by the Department of Psychology and Human Development on phonics and reading development.
- IES' legal basis for processing personal data is 'legitimate interests'. The evaluation of Flexible Phonics fulfils one of IES' core business purposes (undertaking research, evaluation and information activities) and is therefore in our legitimate interest, that processing personal information is necessary for the conduct of the evaluation.
- When the EEF becomes data controller at the end of the project as the data is submitted to the data archive, the legal basis for processing personal data is 'legitimate interests'. These legitimate interests include gather data about what educational interventions work best, under what conditions, for which pupils, with a view to increasing attainment and reducing educational disadvantage.

Who else has access to your information?

The Research Team may disclose your information to third parties in connection with the purposes of processing your personal data set out in this notice. These third parties may include suppliers, research assistants, trusted academic partners and subcontractors who may process information on behalf of the Research Team to carry out such work as administering tests, undertaking interviews, transcribing interviews and archiving data. In all cases, the Research Team will ensure that these third parties enter into appropriate data processing agreements with us and that they keep your personal data secure and confidential.

We may also disclose your personal information if required by law, or to protect or defend ourselves or others against illegal or harmful activities, or as part of a reorganisation or restructuring of our organisations.

International Transfers

Your personal information will not be transferred outside of the EEA.

Security

We take all reasonable steps to protect your personal information and follow procedures designed to minimise unauthorised access, alteration, loss or disclosure of your information.

We have put in place procedures to deal with any suspected personal data breach and will notify you and any applicable regulator of a breach where we are legally required to do so.

Data Retention

We will only retain your personal data for as long as necessary to fulfil the purposes we collected it for, including for the purposes of satisfying any legal, accounting, or reporting requirements. When it is no longer necessary to retain your personal data, it will be securely deleted.

To determine the appropriate retention period for personal data, we consider the amount, nature, and sensitivity of the personal data, the potential risk of harm from unauthorised use or disclosure of your personal data, the purposes for which we process your personal data and whether we can achieve those purposes through other means, and the applicable legal requirements.

- IES and IOE will delete any personal data six months after the completion of the project
- The IOE will retain the anonymised data until March 2025 to permit further analysis
- IES will send the data to EEF's archive within three months of project completion. At this point EEF take responsibility for Data Protection Compliance

Please note that, under Data Protection legislation, and in compliance with the relevant data processing conditions, personal data can be kept for longer periods of time when processed purely for archiving purposes in the public interest, scientific or historical research, and statistical purposes.

Your legal rights

Under certain circumstances, you have rights under data protection laws in relation to your personal data, including rights:

- To request access to your personal data: this enables you to receive a copy of the personal data we hold about you and to check we are lawfully processing it
- To request correction of your personal data: this enables you to have any incomplete or inaccurate data we hold about you corrected
- To request erasure of your personal data: this enables you to ask us to delete or remove personal data where there is no good reason for us continuing to process it
- To object to processing of your personal data: you can object where we are relying on a legitimate interest (or those of a third party) and there is something about your particular situation which makes you want to object to processing on this ground as you feel it impacts on your fundamental rights and freedoms
- To request restriction of processing your personal data: This enables you to ask us to suspend the processing of your personal data
- To request transfer of your personal data
- To object to direct marketing (including profiling) and processing for the purposes of scientific/historical research and statistics
- Not to be subject to decisions based purely on automated processing where it produces a legal or similarly significant effect on you

If you wish to exercise any of the rights set out above in connection with this research project, please contact **Suzanne.anderson@employment-studies.co.uk**

You will not have to pay a fee to access your personal data (or to exercise any of the other rights). However, we may charge a reasonable fee if your request is clearly unfounded, repetitive or excessive. Alternatively, we may refuse to comply with your request in these circumstances.

We may need to request specific information from you to help us confirm your identity and ensure your right to access your personal data (or to exercise any of your other rights). This is a security measure to ensure that personal data is not disclosed to any person who has no right to receive it. We may also contact you to ask you for further information in relation to your request to speed up our response.

We try to respond to all legitimate requests within one month. Occasionally it may take us longer than a month if your request is particularly complex or you have made a number of requests. In this case, we will notify you and keep you updated.

Changes to this Notice

We may change this Privacy Notice from time to time. If we make any significant changes in the way we treat your personal information we will make this clear by contacting schools and ensuring they provide you with an updated version of this Privacy Notice.

Company details for the Research Team and Funder

- The Institute of Education (IOE) is part of University College London
    - University College London a body corporate established by Royal Charter with company number RC000631 whose principal place of business is at Gower Street, London, WC1E 6BT
- The Institute for Employment Studies is a charitable company limited by guarantee registered in England.
    - Registration number: 931547
    - Charity registration number: 258930
    - Registered office: Institute for Employment Studies, City Gate, 185 Dyke Road, Brighton, BN3 1 TL
    - Email: **suzanne.anderson@employment-studies.co.uk**
    - The Education Endowment Foundation is a charity registered in England
    - Registered office: Millbank Tower, 21-24 Millbank, London SW1P 4QP

- Charity registration number: 1142111
- Company number: **07587909**
- Email: **info@eefoundation.org.uk**

Appendix J: School randomisation code

```
set more off

use "$path_derived\schools_pre_test_manipulated", clear

gen sortorder=runiform()

display c(seed)
gen state=c(seed)

sort sortorder

bysort sortorder: gen t=_n
su t

sort sortorder t
gen treatid=_n
gen treated = mod(treatid,2)

la def treated 0"Comparison Group" 1"Treatment Group",replace
la val treated
lab var treated "School Selected for Treatment"

compress

save "$path_derived\selected_schools_first_round.dta", replace
```

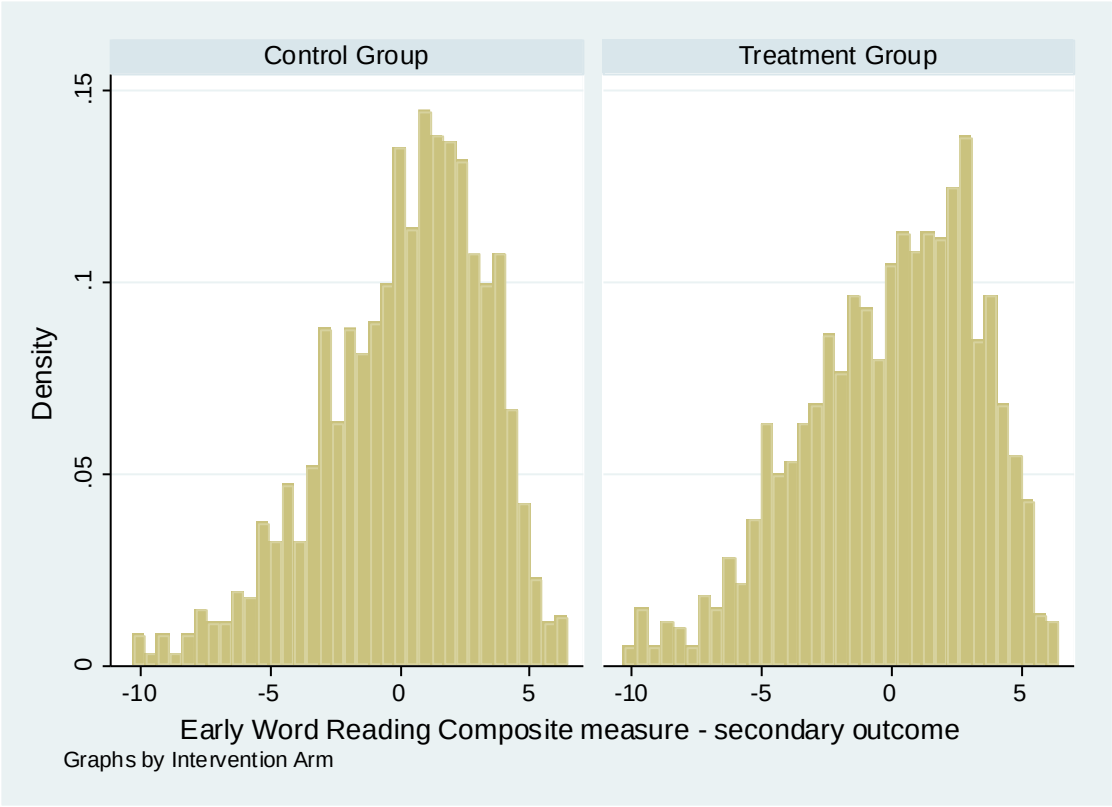
Appendix K: Differences in baseline characteristics at pupil-level for the randomisation sample

Appendix table 3: Absolute standardised differences in baseline characteristics at pupil-level – randomisation sample

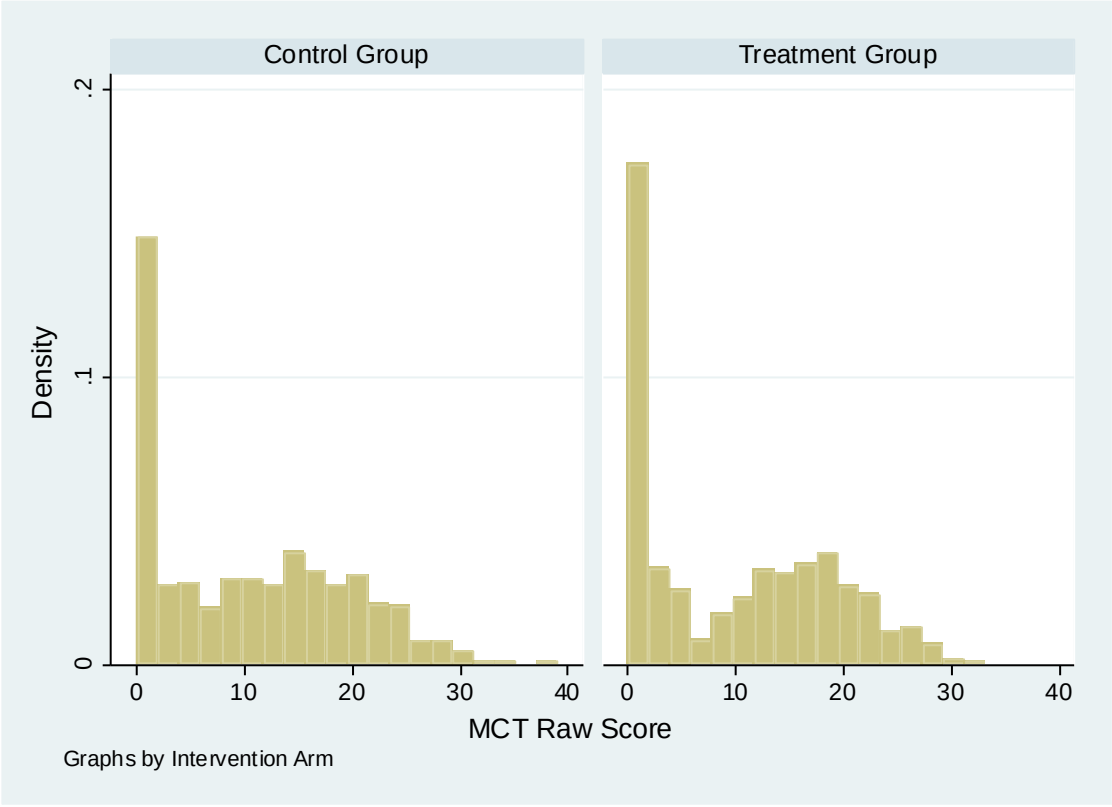
	Intervention group mean	Control group mean	Absolute standardised difference
Pupil-level categorical variables:			
Female	48.18%	48.34%	0.32
Low-ability pupils	52.22%	47.68%	9.10
Participating in NELI	11.73%	8.23%	11.72
Eligible for Free School Meals	18.6%	19.4%	1.80
Pupil-level continuous variables:			
Age in months at time of pre-test	56.20%	56.29%	2.32
Age in months at time of post-test	63.56%	63.58%	0.46
EWR rawscore at time of pre-test	3.65%	3.98%	5.03
Standardised pre-test for secondary outcomes	-0.08%	0.09%	9.51

Appendix L: Histograms of post-test scores on the secondary outcome measures

Appendix figure 1: Histogram of YARC Early Word Reading composite score, by trial arm



Appendix figure 2: Histogram of MCT raw score, by trial arm



Appendix M: Compliance analysis

Appendix table 4: First stage regression results

	Compliance
EWR pre-test score	0.001 (0.001)
Intervention group	0.673 (0.063)***
N	2,539

*Note: Models also includes school fixed effects. Standard errors in parentheses. Statistical significance is indicated as ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.*

Results of F-test: F(2, 2536)=59.83, Prob > F=0.000

Appendix table 5: Instrumental Variables (2-stage least squares) regression results

	EWR post-test score
Compliance	-0.367 (0.784)
EWR pre-test score	0.854 (0.025)***
Intervention group	0.673 (0.063)***
N	2,539

*Note: Models also includes school fixed effects. Standard errors in parentheses. Statistical significance is indicated as ***significant at the 1% level; **significant at the 5% level; *significant at the 10% level.*

Appendix N: Flexible Phonics: Teacher/ TA survey – endline

The endline survey will be sent to all teachers and TAs. In the case of schools with more than one form entry at Reception, the survey will be sent to all Reception class teachers and TAs.

[Hidden question]

1. [hidden question: answer seeded from contacts data file] Which experimental condition does your school belong to? {compulsory}
    a. Intervention group
    b. Control group

Introduction

Thank you for taking part in the Flexible Phonics project. The study is being funded by the Education Endowment Foundation (EEF) to assess the effectiveness of the Flexible Phonics programme, where Teachers and Teaching Assistants (TAs) receive extra training and support to use different techniques in their phonics teaching.

Before the Flexible Phonics project ends, we would like to learn a bit about your role, and how you, as a Teacher or TA, have been teaching phonics, reading comprehension, and writing skills this year. The survey should take approximately ten to fifteen minutes to complete.

All answers are entirely confidential. By taking part in the survey you are consenting for us to use your answers in our analysis. Your data will be held securely in accordance with the Data Protection Act 2018/GDPR and only accessed by the research team at the Institute for Employment Studies (IES). Your employer, or any other third parties, will not see your responses. This information will be used for research purposes only as part of the evaluation.

- The Privacy Notice for this project is available to download at https://www.employment-studies.co.uk/sites/default/files/resources/files/Flexible_Phonics_Privacy_Notice.pdf

If you have any questions about the survey or the evaluation, you can contact Clare Huxley (**clare.huxley@employment**-studies.co.uk) at IES.

If you need to change your answers to any page on the survey at any point, you can use the 'Reset' button at the bottom of each page. This will only reset that particular page of answers.

Your role

2. What is your job role? *Please select one*. {compulsory}
    a. Reception Teacher
    b. Reception TA
    c. Other, please specify
3. How many years of experience do you have teaching or supporting reception-aged children? *Please select one.* {compulsory}

    a. None
    b. Less than a year
    b. 1-5 years
    c. 5-10 years
    d. 10+ years

General Phonics Teaching

In this section, we would like you to focus on what you do in your current role.

4. Do you teach phonics or support phonics teaching in a Reception class? {compulsory}
    a. Yes (route to questions 5-7)
    b. Not sure (route to questions 5-7)
    c. No (route to question 8)

5. Compared to the start of the academic year 20/21 how confident are you in your ability to teach phonics or support phonics teaching with children in Reception? Please select one.
    a. A lot more confident
    b. A bit more confident
    c. About the same
    d. A bit less confident
    e. A lot less confident

6. Compared to the start of the academic year 20/21 how engaged do you think pupils are in lessons when you teach phonics or support phonics teaching with children in Reception? *Please select one.*
    a. A lot more engaged
    b. A bit more engaged
    c. About the same level of engagement
    d. A bit less engaged
    e. A lot less engaged

7. Does your school use one of the phonics programmes from the DfE website to teach phonics? *Please select the one they primarily use*.
    a. Floppy's Phonics Sounds and Letters
    b. Jolly Phonics
    c. Letters and Sounds
    d. Phonics Bug
    e. Phonics International
    f. Read Write Inc
    g. Sound Discovery
    h. Sounds-Write
    i. We use a different phonics programme? If yes, please specify [*Open text*]
    j. We use a mix of different phonics programmes. If yes, please specify [*Open Text*]

k. We do not use a phonics programme

8. Is your school currently taking part in any projects/ programmes focusing on literacy/phonics or language outside of the schools' normal approach? *Please select all that apply.*

   a. Yes, Flexible Phonics
   b. Yes, Nuffield Early Language Intervention (NELI)
   c. Yes, Destination Writer/Reader
   d. Yes, Early Talk Boost
   e. Yes, Early Words
   f. Yes, Hooked by Books
   g. Yes, Power of Reading
   h. Yes, Cornerstones' curriculum planning
   i. Yes, another project
      a. If 'yes', please specify briefly [open text]
   j. No

9. [if Q1=a 'intervention', and Q9='Yes', i.e., a-h] Has participating in another phonics/literacy/language programme alongside Flexible Phonics affected your ability to deliver Flexible Phonics? *Please tick all that apply.*

   a. Yes, I spend less time on Flexible Phonics because I am engaging with the other programme
   b. Yes, as some of their suggested approaches are incompatible or conflicting
   c. No, their approaches are complementary and work well together
   d. No, there are no conflicts between the two approaches
   e. No, they are focused on different areas of language
   f. Not sure

10. [if Q1=2 'control'] What are the costs involved with delivering your usual phonics programme? Please could you briefly describe the type of expense and give an idea of the approximate cost below, e.g., write £100 as 100, or £50.55 as 50.55.

    a. Type of expense [open text box] Cost in £s [open text numeric/sterling only]
    b. Type of expense [open text box] Cost in £s [open text numeric/sterling only]
    c. Type of expense [open text box] Cost in £s [open text numeric/sterling only]
    d. Type of expense [open text box] Cost in £s [open text numeric/sterling only]

11. Do you think that the Covid-19 pandemic has affected the phonics skills of this year's Reception class? *Please select one.*

    a. Yes, they are delayed by up to a month

    b. Yes, they are delayed by about two-three months

    c. Yes, they are delayed by about four-six months

    d. Yes, they are delayed by more than six months

    e. Some children are delayed, some are at the usual level or above

    f. No, they are about the same as usual

    g. Not sure.

Learning about Flexible Phonics [if Q1=a 'intervention']

12. [if Q1=a 'intervention'] Did you attend any of the Flexible Phonics training sessions with Prof. Rob Savage? {compulsory}

    a. Yes, I attended all three

    b. Yes, I attended some

    c. Yes, I attended one/two sessions and my school had a catch-up session with Prof. Savage/ support partners separately because we could not attend all the training

    d. No, but my school had a catch-up session with Prof. Savage/ support partners separately because we could not attend the scheduled training

    e. No, I did not attend any of the training

13. [If Q15='yes a,b,c' or 'no,catch-up, d'] Please indicate the extent to which you agree or disagree with the following statements.
    Strongly agree, agree, Neither agree not disagree, disagree, strongly disagree

    After taking part in the online training or catch-up sessions with Prof. Savage…

    a. I had a basic understanding of the elements involved in the Flexible Phonics programme.

    b. I had a basic understanding of language processing during reading and how the Flexible Phonics strategies aimed to support that.

    c. I had a basic understanding of the Direct Mapping approach and how to teach it.

    d. I had a basic understanding of the Set-for-Variability and mispronunciation correction strategies for reading new exception words and how to teach these approaches.

    e. I had a basic understanding of strategies that I could use with children struggling with general phonics sounds and blending that didn't involve learning sight words by memory.

    f. I felt ready to start planning ways to incorporate these strategies into my phonics teaching with colleagues or by myself.

    g. I felt ready to start teaching Flexible Phonics strategies to children in reception.

14. [if Q1=a 'intervention'] Did you attend any of the follow-up sessions with the Flexible Phonics team, e.g., a support partner? {compulsory}

    a. Yes

    b. No

c. Not sure

15. [If Q17='yes'] Did you use the follow-up sessions in any of the following ways:

    a. To ask for clarification on aspects of the Flexible Phonics programme that you were unsure about

    b. To ask for advice on integrating Flexible Phonics with your usual phonics approach

    c. To access support with planning your delivery of Flexible Phonics

    d. To get feedback on your ideas or plans for implementing the Flexible Phonics programme

    e. To ask for advice on using Flexible Phonics with a specific group of children or child, e.g., struggling readers, EAL, SEND

    f. To access additional support or resources for a topic or aspect of the programme

    g. To share an approach/resource/plan you had developed that was working well at your school

16. [if Q1=a 'intervention'] As part of the Direct Mapping approach, schools were sent a set of children's books to use when teaching Flexible Phonics. Please indicate the extent to which you agree with the following statements.
    Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly disagree.

    I felt that the children's books provided were…

    a. High quality

    b. Appropriate for children in Reception

    c. Enjoyable/fun to read

    d. Useful for teaching Flexible Phonics

17. [if Q1=a 'intervention'] Please could you provide feedback on the following resources provided to support the delivery of the Flexible Phonics programme. If you did not use or access a particular resource, please tick 'N.A.' to indicate not applicable.

    Very helpful, quite helpful, a bit helpful, not at all helpful, not sure, N.A.

    a. Flexible Phonics training manual

    b. Ad hoc help via email/phone from the Flexible Phonics delivery team

    c. UCLeXtend forum

    d. Monthly newsletters from the Flexible Phonics delivery team

Teaching Flexible Phonics [if Q1=a 'intervention']

18. [if Q1=a 'intervention'] How easy or difficult was it to integrate Flexible Phonics with your school's usual phonics teaching approach?

    a. Very easy

    b. Quite easy

    c. Neither easy nor difficult

    d. Quite difficult

    e. Very difficult

    f. Not sure.

19. [if Q1=a 'intervention'] Which aspects of Flexible Phonics have you used with your class? *Please tick all that apply.*

    a. GPCs and Direct Mapping: [info button: Teaching the grapheme phoneme correspondences (GPCs) (letter and sound pairings) and then reinforcing this through reading books that contain those letter/sound pairs.]

    b. Set-for-Variability: [info button: Introducing a variable consonant/vowel strategy and teaching the variability principle through Set for Variability oral games such as playing Simon Says with mispronounced words.]

    c. Teaching vocabulary: exception words: [info button: Teaching 66 key exception words that help children understand texts because children need to know what a word means when they hear it before they can try to read it.]

    d. A strategy for reading key exception words – mispronunciation correction: [info button: Teaching children to read exception words through the mispronunciation correction approach and using these strategies to read 'real books' and decode new words.]

    e. Support for the struggling readers: [info button: Less capable readers may need further work on strategies to overcome challenges of linking individual GPCs to blending such as being aware of 'schwa' sounds and continuous (stretchy) consonants or stop consonants.]

20. [if Q1=a 'intervention'] Please rate how easy or difficult you found it to deliver the different aspects of the Flexible Phonics programme. [*only show options that they have selected in Q22*].
    Very easy, quite easy, Neither easy nor difficult, quite difficult, very difficult.

    a. GPCs and Direct Mapping [repeat info button]

    b. Set-for-Variability [repeat info button]

    c. Teaching vocabulary: exception words [repeat info button]

    d. Reading key exception words – mispronunciation correction [repeat info button]

    e. Support for the struggling readers [repeat info button]

21. [if Q1=a 'intervention'] Please rate how effective you felt that that different aspects of the Flexible Phonics were with your learners. [*only show options that they have selected in Q22*]
    [Very effective, quite effective, a bit effective, not all effective, N.A.]

    a. GPCs and Direct Mapping [repeat info button]

    b. Set-for-Variability [repeat info button]

    c. Teaching vocabulary: exception words [repeat info button]

    d. Reading key exception words – mispronunciation correction [repeat info button]

    e. Support for the struggling readers [repeat info button]

22. [if Q1=a 'intervention'] Have you adapted any part of the Flexible Phonics programme or used other resources to embed and deliver the programme with your learners?

    a. Taught the sounds in a different order than suggested, e.g., used the order in your usual phonics programme such as Read Write Inc.

    b. Taught additional sounds which are not part of the suggested Flexible Phonics list.

    c. Used texts other than the books provided by UCL to undertake the Direct Mapping element of the programme.

    d. Taught a different list of exception words than suggested, e.g., used the exception words list from your usual phonics programme.

e. Continued teaching some sight words with some or all of your learners, e.g., using the traditional flashcard 'whole word' method.

f. Taught Flexible Phonics alongside your existing phonics programme with a specific group of children and continued to use only your usual approach with the rest of the children.

    i. If 'yes', please can you specify which group of learners you teach Flexible Phonics to: [*open text*]

g. Adapted it in another way

    i. If 'another way', please specify briefly: [open text].

23. [if Q1=a 'intervention'] Please indicate to what extent you agree with the statements below about using Flexible Phonics with different groups of pupils:
Strongly agree, Agree, Neither agree not disagree, Disagree, Strongly disagree.

    "I think that Flexible Phonics worked well for…"

    a. High performing readers

    b. Average readers

    c. Struggling readers

    d. Children with English as an Additional Language (EAL)

    e. Children with special educational needs and disabilities (SEND)

24. [if Q1=a 'intervention'] Did you deliver, or attempt to deliver, any Flexible Phonics teaching or activities remotely, e.g., over the internet? *Please tick all that apply*.

    a. Yes, I used video-conferencing software, e.g., Zoom, Teams, etc

    b. Yes, I pre-recorded videos for children and/or parents to watch

    c. Yes, I did live-streaming, e.g., YouTube, Twitch

    d. No, I have only delivered Flexible Phonics face-to-face

25. [if Q1=a 'intervention'] Have you experienced any challenges or barriers when trying to deliver the Flexible Phonics programme? *Please tick all that apply.*

    a. Not enough time to <u>prepare</u> Flexible Phonics activities.

    b. Not enough time to <u>deliver</u> Flexible Phonics activities.

    c. Difficulty fitting it into your normal phonics teaching schedule.

    d. Difficulty integrating it with your normal phonics approach.

    e. Ensuring consistency of approach across all staff teaching phonics

    f. Covid-19 restrictions affecting how you can teach phonics.

    g. Disruption to phonics teaching caused by pupils or staff testing positive for covid-19 and needing to self-isolate for two weeks.

    h. Having the resources to teach Flexible Phonics.

    i. Pupils have not progressed to the point in the phonics curriculum where you can start teaching Flexible Phonics approaches.

    j. Covid-19 restrictions mean that the books provided can't be shared and used as intended.

k. Needing more support from senior staff to be able to embed Flexible Phonics in phonics teaching for reception.

Resourcing, costs and support needs [if Q1=a 'intervention']

26. [if Q1=a 'intervention'] How much time have you needed to **prepare for** and **deliver** teaching the Flexible Phonics programme each week? Please give the average number of hours below, e.g., 2 for two hours, 3.5 for three and a half hours.

    a. [open text] numeric only.

27. [if Q1=a 'intervention'] Have there been any extra costs involved with delivering the Flexible Phonics programme? If so, please could you briefly describe the type of expense and give an idea of the approximate cost below, e.g., write £100 as 100, or £50.55 as 50.55.

    a. Type of expense [open text box] Cost in £s [open text numeric/sterling only]

    b. Type of expense [open text box] Cost in £s [open text numeric/sterling only]

    c. Type of expense [open text box] Cost in £s [open text numeric/sterling only]

    d. Type of expense [open text box] Cost in £s [open text numeric/sterling only]

28. [if Q1=a 'intervention'] Have you received any support from senior or specialist staff, e.g., literacy/early years lead, to deliver Flexible Phonics?

    a. Yes

        i. If 'yes', please could you estimate how many hours of support you have received, e.g., 2 for two hours, 3.5 for three and a half hours.

        [open text] numeric only.

    b. No

    c. Prefer not to say

Phonics activities in Reception

29. As part of your phonics teaching how often on average across the year do you undertake activities involving the following or that teach/ encourage children to do the following?

Applying phonic knowledge and skills in order to decode words
Responding with the correct sound to graphemes (letters or groups of letters) taught in the school's phonics programme
Reading by blending sounds in unfamiliar words that use only the grapheme-phoneme correspondences that have been taught, including words with adjacent consonant sounds and simple words with more than one syllable
Reading exception words, including common words and words of special interest to children, with the teacher/TA highlighting any unusual correspondences between spelling and sound and where these occur in the word
Responding to upper-case letters with the sound that has been taught for corresponding lower-case letters, once upper-case letters have been introduced
Reading words without overt sounding and blending once confident in their decoding, but not before
Reading aloud books that are consistent with children's developing phonic knowledge and that do not require them to use other strategies to work out words
Re-reading books to build up their fluency and confidence in word reading

[list as grid question and respondents can specify frequency for each option]

a. Daily

b. Weekly

c. Monthly

d. Half-termly

e. Termly

f. Yearly or less often

g. Never

30. Reflecting on your teaching since the start of the academic year, has there been a change in how often you undertake the following activities?

[list as grid question and respondents can specify frequency for each option]

a. I do this more frequently

b. I do this less frequently

c. I do this about the same amount

d. Not sure

e. I have never undertaken this activity

Reading comprehension activities in Reception

31. As part of teaching reading comprehension, how often on average across the year do you undertake activities involving the following or that teach/encourage children to do the following?:

[list as grid question and respondents can specify frequency for each option]

a. Daily

b. Weekly

c. Monthly

d. Half-termly

e. Termly

f. Yearly or less often

g. Never

32. Reflecting on your teaching since the start of the academic year, has there been a change in how often you undertake the following activities?

[list as grid question and respondents can specify frequency for each option]

a. I do this more frequently

b. I do this less frequently

c. I do this about the same amount

d. Not sure

e. I have never undertaken this activity

Applying phonic knowledge and skills to decode unfamiliar words accurately, before trying to understand them during reading.
Using pictures and context to help understanding of unfamiliar vocabulary during reading, once the word has been decoded
Asking for help when they do not understand a word or text if something they are reading doesn't make sense.
Discussing and answering questions about what they have read, to show they understand
Encouraging pupils to engage with the content of a text once they have mastered reading it, e.g., discussing how characters feel, or what is happening'
Talking about books they have read and say which ones they like

Spelling and handwriting activities in Reception

33. As part of teaching spelling and handwriting, how often on average across the year do you undertake activities involving the following or that teach/encourage children to do the following?":

[list as grid question and respondents answer can specify frequency for each option]

a. Daily

b. Weekly

c. Monthly

d. Half-termly

e. Termly

f. Yearly or less often

g. Never

34. Reflecting on your teaching since the start of the academic year, has there been a change in how often you undertake the following activities?

[list as grid question and respondents answer can specify frequency for each option]

a. I do this more frequently

b. I do this less frequently

c. I do this about the same amount

d. Not sure

e. I have never undertaken this activity

Listening to sounds and identifying out loud the correct corresponding graphemes (letters or groups of letters), according to those correspondences that have been taught in the school's phonics programme
Writing a corresponding grapheme (letter or group of letters) after listening to a sound
Forming capital letters, as prompted by the phonics programme being followed or at least by the end of Reception
Spelling words by identifying the sounds in the order in which they occur and using their phonic knowledge to represent those sounds with graphemes

Writing, from dictation, simple English words made up of the grapheme-phoneme correspondences learned
Writing some common exception words that have been learned for reading;
Writing stories which include letter- sound combinations they have learnt

Thank you

Thank you for taking time to complete the survey.

Please click ‘Submit’ to send your responses.

[Route to IES thank you page]

Appendix O: IPE outcome tables

Appendix table 6: Frequency of Phonics Activities by treatment group

	Treatment group	Daily		Weekly		Monthly		Half-termly		Termly		Yearly or less often		Never		Total	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Applying phonic knowledge and skills in order to decode words.	Intervention	99	94.3	6	5.7	0	0	0	0	0	0	0	0	0	0	105	100
	Control	111	97.4	3	2.6	0	0	0	0	0	0	0	0	0	0	114	100
Responding with the correct sound to graphemes (letters or groups of letters) taught in the school's phonics programme.	Intervention	97	93.3	5	4.8	0	0	2	1.9	0	0	0	0	0	0	104	100
	Control	107	95.5	5	4.5	0	0	0	0	0	0	0	0	0	0	112	100
Reading by blending sounds in unfamiliar words that use only the grapheme-phoneme correspondences that have been taught, including words with adjacent consonant sounds and simple words with more than one syllable.	Intervention	86	82.7	15	14.4	2	1.9	0	0	1	1	0	0	0	0	104	100
	Control	92	81.4	20	17.7	0	0	0	0	0	0	0	0	1	0.9	113	100
Reading exception words, including common words and words of special interest to children, with the teacher/TA highlighting any unusual correspondences between spelling and sound and where these occur in the word.	Intervention	75	72.1	23	22.1	3	2.9	1	1	1	1	0	0	1	1	104	100
	Control	77	68.1	32	28.3	2	1.8	1	0.9	0	0	0	0	1	0.9	113	100
Responding to upper-case letters with the sound that has been taught for corresponding lower-case letters, once upper-case letters have been introduced.	Intervention	45	45	43	43	6	6	3	3	0	0	0	0	3	3	100	100
	Control	52	46.4	35	31.3	12	10.7	3	2.7	1	0.9	5	4.5	4	3.6	112	100
Reading words without overt sounding and blending once confident in their decoding, but not before.	Intervention	57	57	34	34	4	4	1	1	0	0	2	2	2	2	100	100
	Control	70	62.5	34	30.4	4	3.6	1	0.9	0	0	0	0	3	2.7	112	100
Reading aloud books that are consistent with children's developing phonic knowledge and that do not require them to use other strategies to work out words.	Intervention	60	58.8	33	32.4	3	2.9	2	2	1	1	0	0	3	2.9	102	100
	Control	61	54.5	39	34.8	8	7.1	0	0	0	0	0	0	4	3.6	112	100
Re-reading books to build up their fluency and confidence in word reading.	Intervention	45	44.1	43	42.2	11	10.8	1	1	1	1	0	0	1	1	102	100
	Control	47	41.6	49	43.4	10	8.8	3	2.7	0	0	0	0	4	3.5	113	100

Source: Evaluation endline survey, 2022

Appendix table 7: Change in frequency of Phonics Activities by treatment group

		More frequently		Less frequently		About the same		Not sure		Never		Total	
	Treatment group	N	%	N	%	N	%	N	%	N	%	N	%
Applying phonic knowledge and skills in order to decode words.	Intervention	27	36.5	0	0	44	59.5	3	4.1	0	0	74	100
	Control	31	32	2	2.1	62	63.9	2	2.1	0	0	97	100
Responding with the correct sound to graphemes (letters or groups of letters) taught in the school's phonics programme.	Intervention	25	33.3	2	2.7	46	61.3	2	2.7	0	0	75	100
	Control	31	32.3	1	1	63	65.6	1	1	0	0	96	100
Reading by blending sounds in unfamiliar words that use only the grapheme-phoneme correspondences that have been taught, including words with adjacent consonant sounds and simple words with more than one syllable.	Intervention	27	37	1	1.4	42	57.5	3	4.1	0	0	73	100
	Control	34	35.4	1	1	58	60.4	2	2.1	1	1	96	100
Reading exception words, including common words and words of special interest to children, with the teacher/TA highlighting any unusual correspondences between spelling and sound and where these occur in the word.	Intervention	35	47.3	5	6.8	30	40.5	3	4.1	1	1.4	74	100
	Control	34	35.8	2	2.1	57	60	1	1.1	1	1.1	95	100
Responding to upper-case letters with the sound that has been taught for corresponding lower-case letters, once upper-case letters have been introduced.	Intervention	28	38.4	6	8.2	32	43.8	5	6.8	2	2.7	73	100
	Control	28	29.2	8	8.3	53	55.2	4	4.2	3	3.1	96	100
Reading words without overt sounding and blending once confident in their decoding, but not before.	Intervention	26	36.1	4	5.6	34	47.2	7	9.7	1	1.4	72	100
	Control	34	36.2	4	4.3	52	55.3	3	3.2	1	1.1	94	100
Reading aloud books that are consistent with children's developing phonic knowledge and that do not require them to use other strategies to work out words.	Intervention	30	41.1	7	9.6	31	42.5	4	5.5	1	1.4	73	100
	Control	33	35.1	3	3.2	55	58.5	2	2.1	1	1.1	94	100
Re-reading books to build up their fluency and confidence in word reading.	Intervention	30	40.5	5	6.8	36	48.6	2	2.7	1	1.4	74	100
	Control	35	37.6	3	3.2	51	54.8	3	3.2	1	1.1	93	100

Source: Evaluation endline survey, 2022

Appendix table 8: Frequency of Reading Comprehension Activities by treatment group

	Treatment group	Daily		Weekly		Monthly		Half-termly		Termly		Yearly or less often		Never		Total	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Applying phonic knowledge and skills to decode unfamiliar words accurately, before trying to understand them during reading.	Intervention	73	70.9	29	28.2	1	1	0	0	0	0	0	0	0	0	103	100
	Control	81	72.3	29	25.9	1	0.9	0	0	0	0	0	0	1	0.9	112	100
Using pictures and context to help understanding of unfamiliar vocabulary during reading, once the word has been decoded.	Intervention	65	62.5	38	36.5	0	0	0	0	0	0	0	0	1	1	104	100
	Control	66	59.5	44	39.6	0	0	0	0	0	0	0	0	1	0.9	111	100
Asking for help when they do not understand a word or text if something they are reading doesn't make sense.	Intervention	61	59.2	38	36.9	1	1	2	1.9	0	0	0	0	1	1	103	100
	Control	69	62.7	40	36.4	0	0	0	0	0	0	0	0	1	0.9	110	100
Discussing and answering questions about what they have read, to show they understand.	Intervention	70	68.6	31	30.4	0	0	0	0	0	0	0	0	1	1	102	100
	Control	65	58.6	44	39.6	0	0	0	0	0	0	1	0.9	1	0.9	111	100
Encouraging pupils to engage with the content of a text once they have mastered reading it, e.g., discussing how characters feel, or what is happening'.	Intervention	60	57.7	41	39.4	2	1.9	0	0	0	0	0	0	1	1	104	100
	Control	61	54.5	48	42.9	2	1.8	0	0	0	0	0	0	1	0.9	112	100
Talking about books they have read and say which ones they like.	Intervention	51	49.5	48	46.6	3	2.9	1	1	0	0	0	0	0	0	103	100
	Control	45	40.9	57	51.8	6	5.5	1	0.9	0	0	0	0	1	0.9	110	100

Source: Evaluation endline survey, 2022

Appendix table 9: Change in frequency of Reading Comprehension Activities by treatment group

	Treatment group	More frequently		Less frequently		About the same		Not sure		Never		Total	
		N	%	N	%	N	%	N	%	N	%	N	%
Applying phonic knowledge and skills to decode unfamiliar words accurately, before trying to understand them during reading.	Intervention	18	22.5	7	8.8	54	67.5	1	1.3	0	0	80	100
	Control	37	37.4	0	0	61	61.6	0	0	1	1	99	100
Using pictures and context to help understanding of unfamiliar vocabulary during reading, once the word has been decoded.	Intervention	23	28.8	2	2.5	55	68.8	0	0	0	0	80	100
	Control	33	33.7	2	2	62	63.3	0	0	1	1	98	100
Asking for help when they do not understand a word or text if something they are reading doesn't make sense.	Intervention	19	23.8	5	6.3	56	70	0	0	0	0	80	100
	Control	28	29.2	1	1	66	68.8	0	0	1	1	96	100
Discussing and answering questions about what they have read, to show they understand.	Intervention	26	32.1	0	0	54	66.7	0	0	1	1.2	81	100
	Control	31	32	5	5.2	59	60.8	1	1	1	1	97	100
Encouraging pupils to engage with the content of a text once they have mastered reading it, e.g., discussing how characters feel, or what is happening'.	Intervention	25	31.6	1	1.3	52	65.8	0	0	1	1.3	79	100
	Control	35	35.7	5	5.1	56	57.1	1	1	1	1	98	100
Talking about books they have read and say which ones they like.	Intervention	26	33.8	2	2.6	49	63.6	0	0	0	0	77	100
	Control	35	36.5	7	7.3	52	54.2	1	1	1	1	96	100

Source: Evaluation endline survey, 2022

Appendix table 10: Frequency of Spelling and Handwriting Activities by treatment group

	Treatment group	Daily		Weekly		Monthly		Half-termly		Termly		Yearly or less often		Never		Total	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Listening to sounds and identifying out loud the correct corresponding graphemes (letters or groups of letters), according to those correspondences that have been taught in the school's phonics programme.	Intervention	82	82	18	18	0	0	0	0	0	0	0	0	0	0	100	100
	Control	96	87.3	14	12.7	0	0	0	0	0	0	0	0	0	0	110	100
Writing a corresponding grapheme (letter or group of letters) after listening to a sound.	Intervention	81	81	18	18	1	1	0	0	0	0	0	0	0	0	100	100
	Control	88	81.5	20	18.5	0	0	0	0	0	0	0	0	0	0	108	100
Forming capital letters, as prompted by the phonics programme being followed or at least by the end of Reception.	Intervention	37	37.8	46	46.9	5	5.1	1	1	3	3.1	2	2	4	4.1	98	100
	Control	46	43	39	36.4	8	7.5	1	0.9	3	2.8	3	2.8	7	6.5	107	100
Spelling words by identifying the sounds in the order in which they occur and using their phonic knowledge to represent those sounds with graphemes.	Intervention	80	79.2	20	19.8	1	1	0	0	0	0	0	0	0	0	101	100
	Control	81	76.4	25	23.6	0	0	0	0	0	0	0	0	0	0	106	100
Writing, from dictation, simple English words made up of the grapheme-phoneme correspondences learned.	Intervention	57	57	36	36	4	4	0	0	0	0	0	0	3	3	100	100
	Control	60	55.6	36	33.3	4	3.7	1	0.9	0	0	3	2.8	4	3.7	108	100
Writing some common exception words that have been learned for reading.	Intervention	53	55.2	42	43.8	1	1	0	0	0	0	0	0	0	0	96	100
	Control	50	47.2	49	46.2	6	5.7	0	0	0	0	0	0	1	0.9	106	100
Writing stories which include letter- sound combinations they have learnt.	Intervention	32	32.7	41	41.8	15	15.3	4	4.1	1	1	0	0	5	5.1	98	100
	Control	30	27.8	49	45.4	21	19.4	2	1.9	3	2.8	1	0.9	2	1.9	108	100

Source: Evaluation endline survey, 2022

Appendix table 11: Change in frequency of Spelling and Handwriting Activities by treatment group

	Treatment group	More frequently		Less frequently		About the same		Not sure		Never		Total	
		N	%	N	%	N	%	N	%	N	%	N	%
Listening to sounds and identifying out loud the correct corresponding graphemes (letters or groups of letters), according to those correspondences that have been taught in the school's phonics programme.	Intervention	21	27.3	3	3.9	52	67.5	0	0	1	1.3	77	100
	Control	32	34.8	2	2.2	58	63	0	0	0	0	92	100
Writing a corresponding grapheme (letter or group of letters) after listening to a sound.	Intervention	22	28.6	2	2.6	52	67.5	0	0	1	1.3	77	100
	Control	32	34.8	2	2.2	58	63	0	0	0	0	92	100
Forming capital letters, as prompted by the phonics programme being followed or at least by the end of Reception.	Intervention	24	30.8	4	5.1	48	61.5	1	1.3	1	1.3	78	100
	Control	30	32.6	5	5.4	51	55.4	2	2.2	4	4.3	92	100
Spelling words by identifying the sounds in the order in which they occur and using their phonic knowledge to represent those sounds with graphemes.	Intervention	26	33.3	2	2.6	49	62.8	0	0	1	1.3	78	100
	Control	34	37.4	0	0	57	62.6	0	0	0	0	91	100
Writing, from dictation, simple English words made up of the grapheme-phoneme correspondences learned.	Intervention	30	38.5	1	1.3	43	55.1	3	3.8	1	1.3	78	100
	Control	33	36.3	2	2.2	53	58.2	0	0	3	3.3	91	100
Writing some common exception words that have been learned for reading.	Intervention	29	37.2	1	1.3	47	60.3	0	0	1	1.3	78	100
	Control	31	36	1	1.2	51	59.3	2	2.3	1	1.2	86	100
Writing stories which include letter- sound combinations they have learnt.	Intervention	26	32.9	5	6.3	44	55.7	2	2.5	2	2.5	79	100
	Control	29	32.2	6	6.7	52	57.8	2	2.2	1	1.1	90	100

Source: Evaluation endline survey, 2022

Appendix P: Detailed costs and alternate cost model

As described in the Costs chapter, the intervention group indicated that, as it was currently delivered, Flexible Phonics could be run at little to no additional cost. However, it is important to note that there were some resources that were provided to the schools free of charge by the delivery team. These were the Flexible Phonics manuals at a total cost of £2965.20 for the programme as a whole and sets of high-quality children's books at a total cost of £19,212.27 for the programme as whole. In future these costs (and others noted below) may need to be covered by a fee which we have estimated in an alternate cost model. This appendix explores costs for both the main model and the alternate model in detail to create a deeper understanding of current and possible future costs for delivering the Flexible Phonics programme.

Control schools and prerequisite costs

Staff at schools in the control group were asked about the costs for delivering their usual phonics teaching approach to allow comparison with the costs of participating in Flexible Phonics. The control group reported a number of costs that were part of their business-as-usual phonics programme. We are treating this data as a more qualitative indicator of the sorts of costs schools may have as part of their usual phonics programme and are not treating these as itemised pre-requisites included in the costing models for a number of reasons. These reasons are as follows: that the number of respondents for the costs questions were low and the data was quite variable across different phonics programmes so we do not feel it is robust enough to inform cost estimates in the model. In addition to this, all schools in England will already have materials to deliver phonics and so these should not have to be additionally sourced for any school. The types of items described by the control group as often used for phonics delivery were: phonics books, phonics materials (games, flashcards, magnetic letters etc), phonics staff handbooks, and whiteboards. The control group also mentioned that for some phonics programmes there is a subscription fee. It is not clear from our data if this is a one-off fee or a recurring fee and the number of respondents mentioning this is low. Recurring costs of general resources such as printing were mentioned but again only by a few respondents.

Training

Training for Flexible Phonics consisted of three half-day sessions lasting 3 hours each. In the original programme delivery plan, the training would have taken place in person over 2 full days and would have required teaching staff to be absent for the full school day. However, due to the Covid-19 pandemic, training delivery took place online and the requirement changed to a half-day away from teaching for each of three sessions. Schools reported that as most children were learning remotely, in-person class sizes were small, and so it was possible for one Teacher or TA to cover classroom teaching while other Teachers and TAs attended the online Flexible Phonics training. However, we are aware that in a more typical year with in-person training, there would have been a cost for Teacher/TA cover. We have explored this potential cost by asking schools about typical supply teacher costs and we have estimated teacher cover costs for the in-person model of training delivery that we might expect to see in a more typical year in an alternate cost model. At the outset of the trial, the delivery team had a budget for training costs (such as venue hire) the majority of which was not needed as the training took place online. However, in interviews the project director and project manager both expressed a preference for some or all of the initial training delivery to take place in-person in future where this was possible.

Time

Time needed for planning and delivery of Flexible Phonics is described in the Costs chapter. The alternate model assumes that the time required for preparation and delivery of Flexible Phonics would remain the same but that Teachers and TAs would need additional time to attend two days of in-person training as specified in the original delivery plan prior to the Covid-19 pandemic.

Cost models

In light of changes to programme delivery due to the Covid-19 pandemic, we present two costing scenarios. The first is the main model which reflects the actual costs to the schools during this EEF funded trial and is reported in the Costs

chapter as the basis for calculating cost-effectiveness. However, the design of the programme changed significantly in response to the Covid-19 pandemic, and in-person training and support changed to an online format which impacted on the costs experienced by schools. Our second 'alternate' model here includes costs for school staff to attend training face-to-face as originally planned, including Teacher and TA cover costs. In addition to this, the delivery team provided key items needed for Flexible Phonics delivery for free, as well as spending time beyond their contracted hours. We have added costs for the schools to purchase the Flexible Phonics manuals and sets of children's books to our alternate model, as well as calculating an estimated programme fee to cover the further costs that UCL covered. These include the training venue costs, and their staff time as contracted. We have not attempted to estimate any extra delivery team hours needed as the year of delivery (2021) would not be representative of a normal year because of the Covid-19 pandemic.

The estimated programme fee was calculated using the following assumptions:

- Training venues, food and refreshments (£38,550)
- Delivery team travel to training and support sessions (£16,000)
- 1 full time project manager salary for 1 year (£33,500),
- 1/8th of a professor's salary for 1 year (£10,562.50)
- 6 x Support Assistants working part time (0.5 FTE) for 1 year (£96,651).

Information regarding budgeted costs for venue hire, food and travel were provided by the delivery team in the qualitative interviews and costs pro- forma, All salaries were estimates taken from the UCL website/or online job adverts for similar roles and with the assumption that after the 1st year schools will run Flexible Phonics without any support so no staff time from UCL would be needed to run the programme in subsequent years. We totalled up these costs and divided by the number of schools to which the intervention was delivered – 120 – to create an estimate of a likely per school programme fee. Thus, this alternate model gives an estimate of the maximum costs that might be incurred by schools in a future non grant funded scenario. Even under these circumstances it is still a very low-cost project.

Costs are calculated for a two-form class school with an average of 27 pupils per class, and one teacher and one TA per class.

The tables below present a detailed breakdown of the costs of implementing Flexible Phonics by ingredient for both the main model and the alternate model (Appendix table 12) and recurring costs of implementing Flexible Phonics by ingredient for both the main costs model (Appendix table 13) and the alternate model (Appendix table 14).

Appendix table 12: Costs of the implementation of Flexible Phonics, per ingredient

			Year 1			Year 2			Year 3		
		Start-up or Recurring?	Mean quantity required (other measures of dispersion)	Price per unit required (other measures of dispersion)	Mean cost (min-max)	Mean quantity required (other measures of dispersion)	Price per unit required (other measures of dispersion)	Mean cost (other measures of dispersion)	Mean quantity required (other measures of dispersion)	Price per unit required (other measures of dispersion)	Mean cost (other measures of dispersion)
Personnel	N/A – no extra personnel needed.	N/A									
Personnel for training	Teacher cover	Start up	0 this year (2 full days for 2 teachers in future years)	£0 this year (£200 per day in usual years)	£0 this year (£800 in future years)						
	TA cover	Start up	0 this year (2 full days for 2 Tas in future years)	£0 this year (£100 per day in usual years)	£0 this year (£400 in future years)						
Training and programme costs	Training fee	Start up	Currently covered by UCL		Free provided by UCL						
	Travel and subsistence	Start up	Online this year		Online so no cost (but estimate travel within London £20 zone 1-6 daily travelcard per person per day – 4 people for 2 days = £160)						

			Year 1			Year 2			Year 3		
		Start-up or Recurring?	Mean quantity required (other measures of dispersion)	Price per unit required (other measures of dispersion)	Mean cost (min-max)	Mean quantity required (other measures of dispersion)	Price per unit required (other measures of dispersion)	Mean cost (other measures of dispersion)	Mean quantity required (other measures of dispersion)	Price per unit required (other measures of dispersion)	Mean cost (other measures of dispersion)
Facilities, equipment and materials	Set of children's books	Start up	Unknown	Free this year (but UCL spent £19,212.27 / 120 schools = £160.10 per school)	Free provided by UCL						
	Flexible Phonics training manual	Start up	1 per class	Free this year (but UCL spent £2965.20 / 120 schools = £24.71 per school)	Free provided by UCL						
Other programme inputs	N/A										

Note: Maximum values presented in brackets are estimates based on the original delivery plan as used for the alternate cost model.

Appendix table 13: Recurring costs of the implementation of Flexible Phonics, per ingredient – Main model

		Start up or Recurring?	Nominal values			Real values (deflate using Y1 as Base)			Present value	Cost in analysis year*
			£ Year 2021	£ Year 2022	£ Year 2023	£ Year 1 (in 2021 prices)	£ Year 2 (in 2022 prices)	£ Year 3 (in 2023 prices)	£PV (in 2021 prices)	£PV (in 2021 prices)
Personnel	N/A					0	0	0	0	0
Personnel for training	Teacher cover	Start up (but not needed in trial year due to pandemic)	0	N/A	N/A	0	0	0	0	0
	TA cover	Start up (but not needed in trial year due to pandemic)	0	N/A	N/A	0	0	0	0	0
Training and programme costs	Programme fee (should cover cost of training and delivery team costs (i.e., staff time))	Start up (covered by UCL)	0	N/A	N/A	0	0	0	0	0
	Travel and subsistence for training	Start up (but not needed in trial year due to pandemic)	0	N/A	N/A	0	0	0	0	0
Facilities, equipment and materials	Set of children's books	Start up Provided by UCL	0	N/A	N/A	0	0	0	0	0
	Flexible Phonics training manual	Start up Provided by UCL	0	N/A	N/A	0	0	0	0	0
Other programme inputs	N/A									
	Total cost per school		0	0	0					0
			Number of pupils-per-school-year							162
			Cost per pupil-school-year							0

**Analysis year was the same as the year the programme was delivered.*

Appendix table 14: Recurring costs of the implementation of Flexible Phonics, per ingredient – Alternate model

		Start up or Recurring?	Nominal values			Real values (deflate using Y1 as Base)			Present value	Cost in analysis year*
			£ Year 2021	£ Year 2022	£ Year 2023	£ Year 1 (in 2021 prices)	£ Year 2 (in 2021 prices)	£ Year 3 (in 2021 prices)	£PV (in 2021 prices)	£PV (in 2021 prices)
Personnel										
Personnel for training	Teacher cover	Start up	800	0	0	800	0	0	800	800
	TA cover	Start up	400	0	0	400	0	0	400	400
Training and programme costs	Programme fee (should cover cost of training venue and delivery team costs (i.e., staff time))	Start up	1,627	0	0	1,627	0	0	1,627	1627
	Travel for training	Start up	160	0	0	160	0	0	160	160
Facilities, equipment and materials	Set of children's books	Start up	160.10	0	0	160.10	0	0	160.10	160.10
	Flexible Phonics manuals	Start up	24.71	0	0	24.71	0	0	24.71	24.71
Other programme inputs										
	Total cost per school		3,171.81	0	0					3,171.81
			Number of pupils-per-school-year							162
			Cost per pupil-school-year							19.58

**Analysis year was the same as the year the programme was delivered.*

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