Received: 5 December 2019 | Accepted: 23 October 2023

DOI: 10.1111/bjep.12641

REGISTERED REPORT STAGE 2

An experimental comparison of additional training in phoneme awareness, letter-sound knowledge and decoding for struggling beginner readers

Charlotte Webber[1,2] | Hetal Patel[1] | Anna Cunningham[3] | Amy Fox[1,4] | Janet Vousden[3] | Anne Castles[5] | Laura Shapiro[1]

[1]Institute of Health and Neurodevelopment, School of Psychology, Aston University, Birmingham, UK

[2]Moray House School of Eucation and Sport, University of Edinburgh, Edinburgh, UK

[3]Department of Psychology, Nottingham Trent University, Nottingham, UK

[4]Psychology and Human Development, Institute of Education, Faculty of Education and Society, University College London, London, UK

[5]Australian Centre for the Advancement of Literacy, Australian Catholic University, Sydney, New South Wales, Australia

Correspondence
Laura Shapiro, Institute of Health and Neurodevelopment and School of Psychology, Aston University, Birmingham, B4 7ET, UK.
Email: l.r.shapiro@aston.ac.uk

Funding information
British Academy/ Leverhulme, Grant/Award Number: SRG\170066; Aston University Research Impact Fund, Grant/Award Number: IFOC1-2021/2-004

Abstract

Background: Despite evidence that synthetic phonics teaching has increased reading attainments, a sizable minority of children struggle to acquire phonics skills and teachers lack clear principles for deciding what types of *additional* support are most beneficial. Synthetic phonics teaches children to read using a decoding strategy to translate letters into sounds and blend them (e.g., *c-a-t* = "k - æ – t" = "cat"). To use a decoding strategy, children require letter-sound knowledge (LSK) and the ability to blend sound units (phonological awareness; PA). Training on PA has been shown to benefit struggling beginning readers. However, teachers in English primary schools do not routinely check PA. Instead, struggling beginner readers usually receive additional LSK support.

Aims: Until now, there has been no systematic comparison of the effectiveness of training on each component of the decoding process. Should additional support for struggling readers focus on improving PA, or on supplementary LSK and/or decoding instruction? We aim to increase understanding of the roles of LSK and PA in children's acquisition of phonics skills and uncover which types of additional training are most likely to be effective for struggling beginner readers.

Sample and Method: We will compare training on each of these components, using a carefully controlled experimental design. We will identify reception-age children at risk of reading difficulties (target $n = 225$) and randomly allocate them to either PA, LSK or decoding (DEC) training. We

Charlotte Webber and Hetal Patel are joint lead authors.

will test whether training type influences post-test performance on word reading and whether any effects depend on participants' pre-test PA and/or LSK.

Results and Conclusions: Two hundred and twenty-two participants completed the training. Planned analyses showed no effects of condition on word reading. However, exploratory analyses indicated that the advantage of trained over untrained words was significantly greater for the PA and DEC conditions. There was also a significantly greater improvement in PA for the DEC condition. Overall, our findings suggest a potential advantage of training that includes blending skills, particularly when decoding words that had been included in training. Future research is needed to develop a programme of training on blending skills combined with direct vocabulary instruction for struggling beginner readers.

KEYWORDS

awareness, decoding, early reading, graphemephoneme-correspondences, letter-sound knowledge, phoneme awareness, phonics, phonological, reading difficulties, synthetic phonics

INTRODUCTION

Learning to read is fundamental to all subsequent learning, and therefore effective teaching of early reading is a key priority. The UK-government commissioned Rose Review (Rose, 2006) concluded that synthetic phonics is the most effective way to teach early reading of alphabetic languages. Synthetic phonics teaches children to read using a decoding strategy (translate written letters into sounds and blend them: e.g., *c-a-t* = "k - æ – t" = "cat"). In 2012, the "Phonics Screening Check" was introduced for all Year 1 pupils in England, which has effectively mandated the use of synthetic phonics teaching in English primary schools. There is evidence that school test scores significantly improved following the introduction of synthetic phonics teaching (Machin et al, 2018). Nevertheless, a sizable minority of children continue to struggle with basic reading skills despite receiving synthetic phonics instruction (Department for Education, 2016; NFER, 2018). The current study takes mainstream synthetic phonics teaching as the basis for effective early reading instruction and examines which types of *additional* support would most benefit struggling beginning readers. Synthetic phonics applies a 'part-to-whole approach' (Ehri et al., 2001) in which children learn to decode words by identifying the component phonemes and blending them together to pronounce the word. This approach teaches children to

(i) Develop knowledge of letter-sound correspondences (the letters or groups of letters that represent the individual speech sounds in language; e.g., the written letter *a* standing for the sound "æ").
(ii) Apply their letter-sound knowledge to written words by translating written letters into individual sounds, for example, *c-a-t* = "k - æ - t".
(iii) Blend these sounds to pronounce the word, for example, "k - æ – t" = "cat".

This third step requires "phonological awareness" – awareness of the individual sound units within a spoken word ("k - æ - t") and the ability to manipulate these individual sound-unit representations (i.e., blend them; Pfost et al., 2019). Phonological awareness is highly predictive of reading growth and is thought to be a key pre-requisite for learning to read (Melby-Lervåg et al., 2012). In support of this view, there is evidence that phonological awareness training (in addition to phonics) benefits beginning readers at risk of literacy difficulties (Hatcher et al., 2004; Shapiro & Solity, 2016). However, teachers in English primary schools do not routinely check phonological awareness or provide adapted support for children displaying phonological awareness difficulties. Instead, struggling beginner readers more commonly receive additional letter-sound knowledge support alongside mainstream phonics teaching. There are therefore strong theoretical and practical motivations for an experimental comparison of different types of training for struggling beginner readers.

WHAT IS THE DIFFERENCE BETWEEN PHONICS AND PHONOLOGICAL AWARENESS AND HOW IS THIS DIFFERENCE MANIFESTED IN MAINSTREAM TEACHING IN ENGLAND?

Phonological awareness (awareness of the individual sounds in spoken language) is taught in pre-school and may also be taught very early in Reception (first year of primary school). However, primary-level teaching often does not make a clear distinction between phonics (a strategy for translating *writing* into sounds) and phonological awareness (awareness of the individual sounds in *spoken* language). In most phonics programmes, children gradually build up their knowledge of letter-sound correspondences and learn to apply this knowledge to decoding written words. Great emphasis is placed on letter-sound knowledge and blending in the context of print, and these skills are taught very consistently across different phonics programmes. In contrast, there is less consistency in the way phonological awareness is taught. Although some teachers may check children's awareness of the individual sound units within a spoken word, this is not a requirement and phonological awareness is not normally explicitly taught once children are receiving phonics instruction (except in the context of decoding from print). A key question the proposed study aims to address is whether children can develop sufficient phonological awareness skills through the teaching of phonics, or whether some children require additional training on working with spoken sounds (e.g., segmenting and blending spoken words). Since decoding written words naturally involves working with speech sounds, it is conceivable that practicing decoding in the presence of print is an effective way to train phonological awareness (in fact there is evidence that phonological awareness develops reciprocally with reading, e.g., Perfetti et al., 1987). However, the demands of tasks that involve blending with print versus blending without print are different. When print is present, letter-sound knowledge is recruited, in addition to phonological awareness skills. The presence of print may help some children blend the sounds accurately (e.g., reducing the short-term memory load if letters act as a retrieval cue; National Reading Panel (US), 2000), whilst being unhelpful for others (e.g., causing them to shift attention away from the acoustic properties of the sounds). Therefore, it is plausible that training on blending sounds *in the absence of print* would benefit struggling beginner readers, helping them to acquire sufficient phonological awareness to learn a decoding strategy.

The synthetic phonics programmes commonly used in English Reception classes emphasize letter-sound knowledge over phonological awareness. For example, 'Letters and Sounds' (Department for Education and Skills, 2007) is structured in six phases to teach children 'phonic knowledge' (the alphabetic principle), blending and segmenting. Phase 1 begins in the Early Years Foundation Stage (EYFS) and has a strong focus on phonological awareness in the form of sound discrimination and oral segmenting and blending, however, in Phase 2 – introduced in Reception (first year of school) – the focus changes to letter-sound knowledge (LSK) and blending *written* words. In Phase 2 and beyond, the general procedure for introducing two-syllable words for reading incorporates a text-focussed approach to developing reading skills, for example, '(a) write a two-syllable word on the whiteboard making a slash between the two syllables (e.g., sun/set), (b) sound-talk the first syllable and blend it: s-u-n sun, (c) sound-talk the second syllable and blend it: s-e-t set, (d) say both syllables: sunset, (e) repeat and ask the children to join in, (f) repeat with another word' (Department for Education and Skills, 2007; pg. 65). Another commonly used programme ('Jolly Phonics'; Lloyd et al., 1998) focuses on learning 42 letter sounds and suggests 15 min a day of focussed phoneme tuition in which new letter sounds are introduced at a pace of 4–5 sounds per week. The programme uses a 'multi-sensory approach', with each letter sound being introduced alongside actions, stories and songs. Written letter formations are taught alongside each letter sound, with children typically being taught how to form and write the corresponding letters for each sound within the same session. Children are taught to recognize printed letters and their corresponding sounds and subsequently to blend and segment written words, without focussed tuition on developing these skills in the context of spoken words.

Although additional support programmes can be selected at the discretion of the teacher, there is no requirement that phonological awareness be specifically checked or supported for children displaying reading difficulties. Some additional support programmes include an element of phonological awareness training, for example, Easyread scheme, Reading Recovery, Sound Check and the Reading Intervention Programme (Brooks, 2016). However, these are not purely phonological awareness programmes, and all contain a text-based element. Most other programmes available take the form of additional phonics, rather than focussing on phonological awareness (e.g., Catch-Up Literacy Scheme, Lexia Scheme and Project X Code Scheme). These programmes are usually implemented as a 'catch-up' for children who already display reading difficulties, or who did not pass their Year 1 Phonics Screening Check, rather than serving as preventative schemes during the Reception year. One programme that was designed as a preventative programme from the beginning of primary school is the Early Reading Research (ERR; e.g., Shapiro & Solity, 2016). ERR is delivered to a whole class of children with differing levels of phonological awareness and reading abilities. In the phoneme awareness aspect of the intervention, children are taught to synthesize up to five individual phonemes to pronounce a word and segment words into a maximum of five individual phonemes. The ERR programme is unusual in continuing to practice phoneme awareness orally in the absence of print for the first 2 years of school. Children attending ERR intervention schools outperformed those in comparison schools, and a significantly lower proportion of children were classified as having reading difficulties by the end of Year 2. These benefits may have been partly due to the extra practice children received on phonological awareness; however, this hypothesis would need to be tested by comparing different types of training more systematically.

It is likely that having good phonological awareness is especially important when children are learning to read via synthetic phonics because of the focus on letter sounds at the phoneme level. Shapiro et al. (2013) found that children rely most heavily on their phonological skills when reading words that can only be read using a phonics strategy (i.e., nonwords which cannot be recognized as whole words and must be broken down into letter sounds; e.g., *s-a-n*) versus when reading highly familiar words that may be recognized by sight as whole words (e.g., *the*). It follows that for some children, poor phonological awareness may be a barrier to learning key phonics skills (e.g., blending phonemes to make a word). In support of this hypothesis, Shapiro and Solity (2016) found that children who began school with poor phonological awareness did not respond well to intensive phonics teaching (in fact they responded better to teaching that included regular practice on "sight words"). Perhaps these children require more opportunities to cement their phonological awareness skills (Bowyer-Crane et al., 2008).

Until now, there has been no systematic comparison of the effectiveness of training on each component of the decoding process. Therefore, it is not currently clear how best to support children who are struggling to develop phonics knowledge in their early school years. Should additional support for struggling readers focus on improving phonological awareness skills in the absence of print? Or should additional support focus on enhancing print-based skills in the form of supplementary letter-sound knowledge and/or decoding instruction? Much of the available research into the effectiveness of phonological awareness training focuses on primary school-age children (6 years and above). We review this research below, before making predictions about how phonological awareness training could benefit beginning readers in their first year of phonics instruction.

PHONOLOGICAL AWARENESS TRAINING FOR POOR READERS

There is a well-established link between poor phonemic awareness and deficits in word decoding, which subsequently impairs reading comprehension and fluency (e.g., Elbro & Jensen, 2005; Hulme et al., 2012; Mann & Foy, 2003). Indeed, interventions which focus on strengthening representations of phoneme segments have been shown to aid phonological awareness (e.g., Rvachew et al., 2004). Phonological awareness training has also been used as a therapeutic aid for older children with developmental dyslexia (Snowling, 2001) as well as with 6–7-year-old children with dyslexic parents (defined as "at risk of reading difficulties"; Elbro & Petersen, 2004). Recently, Pfost et al. (2019) tested a 20-week-long phonological awareness intervention for German pre-school children (age 4 years 9 months to 6 years 6 months; early primary in the UK). They found strong effects on phonological awareness, which translated to a benefit in reading only for the lowest performing children. They suggest that screening procedures within schools could help identify children with specific phonological awareness deficits for which adapted training could be beneficial. Similarly, Hatcher et al. (2004) found benefits of additional phonological awareness training, but only for children at risk of reading difficulties. They compared four treatment groups: (i) reading alone; (ii) reading with rhyme training; (iii) reading with phoneme training; and (iv) reading with rhyme and phoneme training. Children (*M* age 4.65 years at the start of training) received 10-minute intervention sessions 3 times a week over 14.5 months. The results showed that only phoneme training was "truly selective in facilitating phoneme skills" (p.350) and that reading with rhyme training did provide a benefit to phoneme skill when combined with phoneme training but not over and above that of phoneme training alone. Importantly, specific phoneme training generalized to reading skill in 'at risk' children but not in 'normally developing' children. This study provides good evidence that an extensive period of phoneme awareness training (14.5 months) at an early age can be of specific benefit to those 'at risk' of phonological awareness deficit. However, this study was conducted prior to the introduction of phonics in English primary schools, and since there was considerable overlap in terms of the training provided in each condition, it does not indicate which component skill is most critical to support in beginning readers.

Overall, although there is good reason to believe phonological awareness training would be beneficial for some beginning readers, there has not yet been a systematic comparison with other types of training. There remains an open question as to whether phonological awareness training in the absence of print (spoken words only) is necessary for the development of a robust decoding strategy or whether print-based training is sufficient.

THE PRESENT STUDY

In the proposed study, we aim to compare the benefits of additional training on each of the components of the decoding process with a view to establishing key principles that help teachers decide how best to support struggling beginner readers. Additional phoneme awareness training may help some children to overcome an initial short-term barrier to their learning, enabling them to access mainstream classroom phonics teaching, which in turn will further develop their phonological awareness (the ability to manipulate the sounds that make up words) and reading skills. Alternatively, children who are struggling to access class-based phonics may benefit from additional support in securing letter-sound knowledge, or additional practice on the full process (i.e., decoding printed words). We will compare the effectiveness of three types of additional training (phoneme awareness, letter-sound knowledge and decoding) on children's phonological awareness skills, letter-sound knowledge and word reading. Since our target group of participants is struggling beginner readers, we anticipate very low performance on standardized tests of reading. We therefore use a dynamic testing schedule, in order to explore in more detail the components of the reading process that are being supported by each training technique. A dynamic test was chosen rather than static measures to allow pupils to respond after some prompting. Using a static measure where pupils can either correctly or incorrectly respond on their first attempt is likely to result in floor performance for lower attaining readers in their first year of school. This dynamic technique will also enable us to explore the rate at which children pick up the phonics strategy, provide insight into which particular skill (PA, LSK or decoding) children are struggling with, and whether PA, LSK and/or DEC training is most effective at improving the rate at which children are able to apply phonics strategies to word reading. Since reading outcomes may not have the same "proximal" and "distal" causes in all poor readers (McArthur & Castles, 2017), the present study will also examine whether children's phonological awareness and letter-sound knowledge at baseline influence the type of training that is most beneficial. Ultimately, we aim to inform teachers and education providers about the most effective strategies to use to support struggling beginner readers.

The research questions to be explored in the proposed research are as follows:

1. Does type of training (PA vs. LSK vs. Decoding) affect post-test performance on a dynamic decoding task?
   Hypothesis: The PA-training group will outperform the LSK-training group and decoding training group on reading of trained words at post-test (we anticipate that these children will already be receiving LSK and decoding practice as part of mainstream teaching, but they will not be receiving PA support).

2. Is any effect of training type on post-test performance on a dynamic decoding task moderated by pre-test PA or LSK scores?
   Hypotheses: The advantage of the PA-training group will be moderated by pre-test PA and LSK. We anticipate that lower initial PA scores will be associated with greater effects of PA training, and lower initial LSK scores will be associated with greater effects of LSK and decoding training.

3. Does training type affect performance on standardized tests of phonological awareness and letter-sound knowledge?
   Hypotheses: Type of training will affect post-test performance on YARC Sound Isolation, Sound Deletion and Letter-Sound Knowledge measures. Specifically, PA training will deliver the greatest benefit for Sound Isolation and Deletion, whereas LSK training will deliver the greatest benefit for LSK.

METHODS

Ethics

Our University Ethics Committee has approved this study. All parents gave written informed consent prior to participation.

Data and material release

This project was pre-registered on the Open Science Framework. All anonymized data collected as part of this study has been made openly available on the Open Science Framework (OSF). We have also uploaded materials, such as the word lists, so that interested readers can reproduce the training and assessment programmes (10.17605/OSF.IO/AF56J).

Pilot study

We have conducted a pilot study in order to validate our methods and guide selection of words for the present study. In our pilot study, children were randomly assigned to a training group ("decoding" or "whole word") or a control group. Children in both training groups received individual reading practice sessions using the same set of word cards but with condition-specific instructions. This pilot work informed the script used for the current study, using transparent and simple instructions and maintaining the child's interest by including a picture and standardized sentence after children attempt each word (see Table 1 for full script). Children in the pilot study appeared to enjoy the sessions and were engaged with the task. Importantly, this pilot work demonstrated that a short period of training (10 sessions of 20 min) resulted in a large advantage for trained over untrained groups ($d = .74$). See Appendix S1: Section 1 for details of this pilot work. In the main study, we adapt this design to deliver three different types of training on the same set of words (LSK: say each letter; PA: isolate, delete and blend the phonemes; DEC: say the letters and blend the phonemes; see Table 1).

Power analysis

Our pilot study demonstrated that 225 participants were sufficient to detect effects of trained versus untrained groups on post-test word reading, and to detect an interaction between word type (trained vs. untrained words) and time (see Appendix S1). We additionally conducted power simulations to gain sample size estimates for each of our proposed analyses (see analysis section below for R script). All power simulations gave estimates of fewer than 225 participants to have 90% power for each of our main effects (alpha $<.05$). To allow for the same level of attrition observed in our pilot study, we will recruit 232 children in the main study. See Appendix S2 for details of our power analyses.

Participants

We asked participating teachers to identify children who they consider to be "below average" in their reading performance, for their class (lowest performing 15 per class; in England, this is approximately half of the class) and invited parents to give consent for these children to participate (as in our pilot study; Appendix S1). All children with parental consent who met the study criteria were be invited to participate. However, we explicitly asked the child whether they would like to participate (using a child-appropriate form in which they point to one of two pictures to indicate their choice: taking part in our activities or returning to their classroom). If the child did not actively choose to take part, or if their response was ambiguous, they were thanked and given a sticker (using the same script and reward as for participants who completed the session). Children with parental consent who indicated they would like to take part then completed our screener, and those gaining a standard score of 85 or below were eligible for the training study (equivalent to the bottom 16% of readers, average ~ 5 per class). Based on our previous research, we anticipated gaining consent for 4/5 of eligible children, requiring c. 46–47 classes of children from c. 20 schools (many of our partner schools expressed an interest in participating and so this level of recruitment was realistic). Since the study was completed within a school term, attrition was expected to be low (<3% observed in pilot study). To compensate for missing data due to child absence or withdrawing of consent, we aimed to recruit 232 eligible participants and expected to retain 225 participants at post-test. This would give 75 participants in each condition: (i) letter-sound knowledge (LSK) training, (ii) phoneme awareness (PA) training and (iii) decoding (DEC) training. Eligible children were invited to complete our pre-tests and then randomly assigned to a training condition. We additionally recorded information on English as first language (EAL) and special education needs (SEN) so that we can note whether these characteristics are distributed equally among our conditions. However, at age 4–5 years, even participants without an assessment of SEN will vary in their speech

TABLE 1 Condition-specific script for PA, LSK and decoding conditions.

Phoneme awareness training	Letter-sound knowledge training	Decoding training
Introduction to Part 1: Phonemes		
We are going to think about the sounds that make up words. I'm going to say a word and you tell me what the sounds are.	We are going to think about the letters that make up words. I'm going to show you a word and you tell me what the letters sound like.	We are going to think about the sounds that make up words. I'm going to show you a word and you tell me what the sounds are.
Part 1: Phonemes		
[say spoken word, e.g., "much"] What sound does ["much"] begin with? Yes ["much"] begins with ["m"]/good try, ["much"] begins with ["m"], can you say ["m"]? Now can you say ["much"] without the ["m"]? Yes ["much"] without the ["m"] is ["uch"]/good try, ["much"] without the ["m"] is ["uch"], can you say ["uch"]?	[present written word, e.g., *much*] What sound do(es) this (these) letter(s) make [Point to first grapheme, *m*]? Yes, this letter (these letters) make(s) ["m"] / good try, this letter (these letters) make(s) ["m"], can you say ["m"]?	[present written word, e.g., *much*] What sound does this letter make [Point to first grapheme, *m*]? Yes, this letter (these letters) make(s) ["m"] /good try, this letter (these letters) make(s) ["m"], can you say ["m"]?
What sound does ["much"] end with? Yes ["much"] ends with ["ch"]/good try, ["much"] ends with ["ch"], can you say ["ch"]?. Now can you say ["much"] without the ["ch"]? Yes ["much"] without the ["ch"] is ["mu"]/good try, ["much"] without the ["ch"] is ["mu"], can you say ["mu"]? *Go to Part 3: Context. Repeat for all words.*	What sound do(es) the next letter(s) make? [Point to next grapheme] Yes, this letter (these letters) make(s) [sound] / good try, this letter (these letters) make(s) [sound], can you say [sound]? *Repeat for all graphemes in a word, then go to Part 3: Context. Repeat for all words.*	What sound do(es) the next letter(s) make? [Point to next grapheme] Yes, this letter (these letters) make(s) [sound] /good try, this letter (these letters) make(s) [sound], can you say [sound]? *Repeat for all graphemes in a word, then go to Part 3: Context. Repeat for all words.*
Introduction to Part 2: Blending		
This time we are going to blend some sounds together to make a word. [shuffle word cards]	We are going to do this set again one more time. [shuffle word cards]	This time we are going to blend some sounds together to make a word. [shuffle word cards]
Part 2: Blending		
What do we get if we blend ["m-u-ch"]? Yes, they make ["much"] /good try, ["m-u-ch"] makes ["much"]. Can you say ["much"]? *Repeat 3 times per word, blending sounds to make the (i) word (as above): ["m-u-ch"], ["much"] (ii) rime: ["u-ch"], ["uch"] (iii) onset and rime: ["m-uch"], ["much"].* *Then go to Part 3: Context. Repeat for all words.*	*No blending. Repeat Part 1 and Part 3 for all words a second time.*	What do we get if we blend [Point to each grapheme and say corresponding phonemes]? Yes, [point and say each phoneme] makes [sound] /good try, [point and say each phoneme] makes [sound]. Can you say [sound]? *Repeat for all graphemes in a word then go to Part 3: Context.* *Repeat for all words.*
Part 3: Context		
[Present picture] What can you see in this picture [Read standardized sentence]	[Present picture] What can you see in this picture? [Read standardized sentence]	[Present picture] What can you see in this picture? [Read standardized sentence]

Note: Wording based on the Phoneme Discrimination and Alliteration and Phoneme blending tasks from The Highland Council Phonological Awareness Informal Assessment for school-aged children (The Highland Council, 2016).

production abilities, which poses a challenge for coding verbal responses. To facilitate coding, we have avoided words that are most commonly mispronounced by children with speech production difficulties (e.g., avoiding fricative + liquid clusters such as /sl/; McLeod et al., 2001).

Screening and background measures

York Assessment of Reading for Comprehension (YARC) Single-Word Reading, Sound Isolation, Sound Deletion and Letter-Sound Knowledge subtests. Participants completed the single-word reading (SWRT) subtest of the YARC (Hulme et al., 2009). Children who obtained a standard score of 85 or below on this test then also completed a battery of pre-tests comprised of the sound isolation, sound deletion and letter-sound knowledge subtests from the YARC to assess PA and LSK. In the isolation subtest, children are asked to repeat a word and then isolate a sound from the beginning, or end of the word. Children are awarded one point for each correct response with a maximum score of 12. In the deletion subtest, the participant hears a word, and is asked to repeat it, dropping a specified phoneme (e.g., 'without the /b/'). In the letter-sound knowledge subtest, participants are asked to identify the sound that individual letters make, providing a measure of LSK. Tests were administered by a trained research assistant in the participants' school, in a quiet area free from distractions.

Socioeconomic Status and Home Language. Students from lower socioeconomic backgrounds tend to perform more poorly on school assessments (Education Policy Institute, 2018), and disadvantages in measures of language have been observed even in pre-school (Roy et al., 2014). In addition, children speaking languages other than English at home typically have lower levels of English Vocabulary and Grammar (e.g., Babayiğit & Shapiro, 2020) potentially causing slower growth in reading. We recorded home background information using a short questionnaire for parents tapping into parents' level of education (an indicator of SES most closely related to academic attainments; Ilie et al., 2017) and the languages spoken at home (Appendix S4). These measures were used to check the distribution of participants across conditions and whether participants who reach criterion were broadly representative of the full sample of screened children. Depending on the range of responses gained, these measures may be used in an exploratory analysis of the effectiveness of our training for children from different home backgrounds.

Training protocol and selection of words

Following the completion of screener and pre-tests, participants were randomly allocated to one of three training conditions. They received 10 training sessions (2 sessions a week for 5 weeks) in addition to their standard classroom teaching, estimated time of 20 min per session. The sessions were carried out by a trained research assistant who delivered training using the relevant condition-specific script (see Table 1) and the same word list for each condition (Appendix S3). A matched set of generalization words was included in our post-tests. The training and generalization words were selected to ensure that children who meet our screening criteria could not read them independently from the outset (i.e., we can assume floor performance at pre-test, without having to pre-expose participants to the words). The words are also age-appropriate in terms of Reception children's vocabulary (Masterson et al., 2010) and contained grapheme–phoneme correspondences (GPCs) that are taught in Reception. Participants were trained on 10 words from the training set per session. In each session, two sets of five training words will be used with the sets organized in increasing difficulty, as determined by number of GPCs and word frequency (training session 1: set A&B, session 2: AB, session 3: BC, 4: BC, 5: AB, 6: AB, 7: BC, 8: BC, 9: AB, 10: BC; Appendix S3). As shown in the script in Table 1, all conditions presented the graphemes/phonemes within words, including the LSK condition. In the LSK and DEC conditions, the child was shown the printed word, whereas in the PA condition, they were only presented with the spoken word (no print). Both the PA and the DEC conditions trained children on blending phonemes, and in the PA condition, children were additionally trained to segment sounds within the words (i.e., isolating or deleting phonemes). Both the LSK and DEC conditions presented the written words, but in the LSK condition children received no training on blending the sounds. In all conditions, after a word had been attempted, a picture card was used to create a context, and a standardized sentence read out. Our pilot participants enjoyed this addition, and it also has the advantage of providing an oral presentation and a meaningful context for each word (even in the LSK condition, where the child would otherwise not hear or say the words).

Data collection timeline

Stage 1: Screening. The screening test (SWRT) was administered by a Research Assistant on a one-to-one basis in a quiet location within the child's school. Arrangements were made with teachers to accommodate participants who require extra support (e.g., a teaching assistant present). Children with a standard score of 85 or below proceeded to stage 2.

Stage 2: Pre-tests. Eligible participants completed our pre-tests with a trained Research Assistant for approximately 20 min to measure their initial PA and LSK skills.

Stage 3: Training. Participants were randomly allocated to one of the three conditions (PA, LSK and DEC). Participants in all conditions completed 10 sessions of training (2 sessions a week for 5 weeks; total participation time was 6 weeks, accounting for screening, pre-tests and post-tests). A trained Research Assistant worked with children individually in a quiet location at school, using the Training Words list and the relevant training script for each condition (see Table 1). Each Research Assistant was observed twice over the 10 sessions, to monitor fidelity to the training protocol, following a structured observation record (adherence to each section of the training script, number of words trained and length of session time). When in-person observation was not possible (e.g., due to limits on the number of visitors to a school), this observation was conducted using an audio recording of the session.

Stage 4: Post-test. Following the 5-week training period, participants in all three conditions completed the post-tests: the YARC sound isolation, sound deletion and LSK subtests (see pre-tests above), plus an additional dynamic decoding test using the training words plus a matched set of untrained words (generalization words; see Appendix S3 for details of matching process). Dynamic testing was used as it was expected that a large number of participants may still not be able to read a majority of the words independently, resulting in floor effects from a static task. Participants were presented with sample words from each word list and asked to read them aloud. Responses were scored dynamically, based on the method outlined by Spector (1992) and implemented by Cunningham and Carroll (2011), among others. If children could read a word, they were given increasingly explicit prompts to aid word reading and scored appropriately. Each participant received a scaled score for each word rather than a categorical 'learned'/'not learned' score in order to give a clearer indication of proficiency in each of the component skills of reading (LSK, blending). The prompts implemented for each word are outlined below; the skill that each prompt assists with is given in parentheses – note that performance on the dynamic task will be affected by which skill the child struggles with. If they require prompts for LSK, but have good blending skills, they will score relatively higher (requiring fewer prompts) than for a child with the opposite balance of skills. Participants were given an individual score for each word as well as a mean Word Reading score.

- Prompt 1: [point to first grapheme] "What sound does the first letter in this word make? …. is the first sound". *(Skill: LSK)*
- Prompt 2: [point to second grapheme] "What sound does the next letter make? …. is the second sound". *(Skill: LSK)*
- Prompt 3: [point to next/last grapheme] "What sound does the next/last letter make? …. is the next/last sound". Continue to prompt for each grapheme up to the final grapheme. *(Skill: LSK)*

- Prompt 4: "Now we're going to put all those sounds together. What do we get if we blend together [point to each grapheme and say corresponding sounds]?" [Do not give correct answer.] (*Skill: blending)*
- Prompt 5: "What do we get if we blend these sounds together [point to graphemes that form the rime and say corresponding sounds]? ….is what we get". (*Skill: blending)*
- Prompt 6: "What do we get if we blend the first sound with the final sounds together?" [Point to letters and say corresponding sound for the onset and rime. Do not give correct response.] (*Skill: blending)*
- Prompt 7: [Present picture] "What can you see in this picture?" [Read standardized sentence.] (*Skill: picture recognition)*

Dynamic scores for each word indicate the degree of independence that the child achieved in performing the decoding task (with higher scores reflecting better performance). The task was scored as follows:

- 7 = correct response with no prompts required;
- 6 = correct response after Prompt 1;
- 5 = correct response after Prompt 2;
- 4 = correct response after Prompt 3;
- 3 = correct response after Prompt 4;
- 2 = correct response after Prompt 5;
- 1 = correct response after Prompt 6;
- 0 = correct response or no correct response after prompt 7.

Data analysis plan

Data comprised raw scores obtained from standardized and bespoke tasks. We report participant means and standard deviations for all measures, alongside correlations between all measures.

Manipulation and quality check

We checked (i) the distribution of participants over conditions and (ii) for floor and ceiling effects on pre- and post-test measures.

Testing hypotheses

Data was analysed using linear mixed effects modelling for each of the conditions to explore the effects of item-level and participant-level variables on post-training outcomes. Statistical analysis were performed in R (R Core Team, 2013).

Question 1: Does type of training (PA vs. LSK vs. DEC) affect post-test performance on a dynamic decoding task?

Data was analysed using the *lmer* function of the *lme4* package (Bates et al., 2014) in R (R Core Team, 2013), with dynamic decoding performance (as measured at post-test) as the dependent variable, training type (PA vs. LSK vs. decoding) as the fixed effect and item and participant as random effects.

$$\text{Dynamic decoding} \sim 1 + \text{training type} + (1|\,\text{item}) + \left(1|\,\text{participant}\right)$$

Question 2: Is any effect of training (PA vs. LSK vs. DEC) on post-test performance on a dynamic decoding task moderated by pre-test PA or LSK scores?

In order to explore whether effects of training type differ based on participants' pre-test PA and LSK scores, we planned a lme analysis with post-test dynamic decoding score as the dependent variable, pre-test scores (PA or LSK) as fixed effects and random effects on both participant and item.

$$\text{Dynamic decoding} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test PA} + (1|\ \text{item}) + \left(1|\ \text{participant}\right)$$

$$\text{Dynamic decoding} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test LSK} + (1|\ \text{item}) + \left(1|\ \text{participant}\right)$$

Question 3: Does type of training affect performance on standardized tests of phonological awareness and letter-sound knowledge?

In order to establish whether there was an effect of type of training on the specific skills that are trained, we planned two glmer analyses:

$$\text{Post} - \text{test PA} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test PA} + (1|\ \text{item}) + \left(1|\ \text{participant}\right)$$

ondition (50%), 36 out of 54 respondents i

$$\text{Post} - \text{test LSK} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test LSK} + (1|\ \text{item}) + \left(1|\ \text{participant}\right)$$

Study timeline

Stage 1 report in principle acceptance on 30 December 2020; data collection commenced in April 2021 and was completed in June 2023.

RESULTS

Participant-level descriptive analyses

Teachers from 35 classes across 12 schools sent letters to the parents of 525 children inviting them to participate in the study (lowest performing 50% of children, per class). We received parent consent to invite 271 children to participate, and of these, 251 children agreed to take part. Of the 251 children who completed the screener, 18 children were excluded with screener standard scores above 85. The remaining 233 participants were eligible for the study and were randomly assigned to conditions. Four of these children were excluded due to persistent absence and seven children opted out at stage 2 (pre-tests). The remaining 222 participants completed the phonological awareness (PA) ($n = 69$), the letter-sound knowledge (LSK) ($n = 76$), or the decoding (DEC) training ($n = 77$).

One-hundred and seventy-four completed parent questionnaires were returned (174/271), with 5 parents identifying their child as SEN (2 children in PA, 1 in LSK, and 2 in DEC) and 103 parents indicating English as an additional language (EAL); 24 out of the 48 respondents in the DEC condition (50%), 36 out of 54 respondents in the LSK condition (66.7%) and 27 out of the 46 respondents in the PA condition (58.7%) spoke languages other than English at home. There were also 16 out of the 26 (61.5%) respondents from the excluded participants who identified English as an additional language. The most common languages spoken at home were Urdu, Punjabi, Arabic and Hindko.

Although 174 questionnaires were completed, many parents omitted the parent education question, leaving 118 responses for female caregiver's qualifications and 104 responses for male caregiver's qualifications. For female caregivers, 16 (13.6%) indicated GCSEs or equivalent, 25 (21.2%) indicated A-levels or equivalent and 62 (52.5%) indicated a qualification higher than A-level (e.g., undergraduate, higher degree, etc.). For male caregivers, 25 (24%) chose GCSEs as the father's highest qualification, 20 (19.2%) chose A-levels, and 36 (34.6%) chose higher than A-levels. For a full breakdown of the qualifications of the participants' parents by condition, see Tables S3 and S4 in Appendix S5.

Screener and pre-test scores

From this point onwards, we report data for children who completed the training ($n = 222$). For these children, standardized scores ranged between 69 and 85 (equivalent to a raw score of between 0 and 4 words read correctly), with an average score of 70 (raw score of .38). We must note that SWRT norms are not available for children under 5 and so the standard score will be an underestimate for many children (the average age of our sample was 4 years and 10 months). Therefore, our sample may include children reading at, or close to, age-appropriate levels.

The average score on our pre-tests for these 222 children was 3.59/12 (SD = 2.97) for sound isolation; 1.83/12 (SD = 1.83) for sound deletion and 20/30 (SD = 9.51) for letter-sound knowledge. A full breakdown of average scores and standard deviations on pre-test measures by condition can be seen in Table 2. Floor and ceiling effects were checked for each pre-test measure. In the sound isolation test, 46 (20.7%) participants scored 0 whilst 4 (1.8%) scored the maximum of 12. Sixty-five (29.3%) of the participants scored 0 on the sound deletion test, and no participants scored the maximum score. Finally, 9 (4.05%) participants scored 0 on the letter-sound knowledge test, whilst 4 (1.80%) participants scored the maximum of 32.

Distribution of post-test scores

Floor and ceiling effects were checked for each post-test measure. In the sound isolation test, 19 (9%) participants scored 0 whilst 23 (10%) scored the maximum of 12. Twenty-eight (13%) of the participants scored 0 on the sound deletion test and no participants scored the maximum score. Finally, one participant scored 0 on the letter-sound knowledge test, whilst 12 (5%) participants scored the maximum of 32 (a full breakdown of the floor and ceiling effects of pre- and post-test scores by condition is shown in Tables S5 and S6 in Appendix S5). As shown in Table 2, there was a trend for greater improvement in sound isolation at post-test for the decoding condition and a trend for greater improvement in letter-sound knowledge at post-test for the LSK condition. There are no clear trends for sound deletion, and scores remain low for this measure even at post-test (average of 4/12). Dynamic decoding was only completed at post-test. Participants could score between 0 and 7 on each word, and the average score per word was 2.57 (SD = 2.51). Table 2 shows the total scores for dynamic decoding across trained and untrained words (30 words).

Planned analyses

Two hundred and twenty participants completed the dynamic decoding task and were included in the following models. The models compare the differential effect of each condition on outcomes and use sum contrasts to code the conditions (as recommended by Schad et al., 2020). Sum contrasts effectively compare each condition to the average across all conditions. If the effect of a condition is significant, then we can interpret this as a differential effect of that condition. If there are no effects of condition,

TABLE 2 Mean scores (and standard deviations) for pre- and post-test measures by condition.

	Sound isolation		Sound deletion		Letter-sound knowledge		Dynamic decoding
Condition	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Post-test
PA	3.73 (2.94)	6.29 (3.54)	1.87 (1.57)	4.04 (2.73)	20.55 (9.83)	26.41 (6.52)	81.10 (45.60)
LSK	3.63 (3.04)	6.08 (3.76)	1.79 (1.85)	4.12 (2.77)	19.88 (9.17)	27.20 (5.37)	72.40 (53.10)
DEC	3.44 (2.94)	6.81 (3.40)	1.83 (2.04)	4.05 (2.51)	19.56 (9.62)	26.36 (5.08)	75.50 (46.60)

Note: Maximum possible score on sound isolation and sound deletion tasks was 12. Maximum score on letter-sound knowledge test was 32. Maximum score on dynamic decoding was 210 (30 words, each scored 0–7).

then this would indicate that scores were similar across the three types of training. Fixed effect variables and interactions were added in a stepwise manner and compared to the last significant model. Each subsequent model was accepted if it led to a significant improvement in fit (as determined by a chi-square test). Our R script and all output are included in Appendix S6.

Question 1: Does training type (condition) affect post-test performance on a dynamic decoding task?

A model was fitted with dynamic reading as the outcome and condition as a fixed effect. The lmerTest package was used to obtain *p*-values, and the ANOVA function was used to compare the model to the below null model (see Appendix S6 for the full script and output for all analyses):

$$\text{lmer}(\text{formula} = \text{Score} \sim 1 + (1|\,\text{ItemNumber}) + \left(1|\,\text{ParticipantNumber}\right)$$

As shown in Table 3 there were no effects of condition on dynamic reading score (and the model was not significantly better than the null model; non-significant chi-square). For this planned analysis, the total dynamic reading score for trained and untrained words was used (as reported in Table 2). In our exploratory analysis section, we compare the effects of trained versus untrained words.

Question 2: Is any effect of training (PA vs. LSK vs. decoding) on post-test performance on a dynamic decoding task moderated by pre-test PA or LSK scores?

To explore whether pre-test scores on the PA and LSK tasks had any effect on the condition effects on dynamic reading scores, a series of models was fitted with dynamic decoding score as the outcome. Our analysis plan did not specify whether sound isolation and sound deletion scores would be combined or kept as separate measures of PA. We decided the safer option was to maintain these subtests as separate measures to avoid effects being missed due to floor performance on sound deletion or ceiling performance on sound isolation. We therefore built three models to investigate this question:

$$\text{Dynamic decoding} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test SoundIsolation} + (1|\,\text{item}) + \left(1|\,\text{participant}\right)$$

$$\text{Dynamic decoding} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test SoundDeletion} + (1|\,\text{item}) + \left(1|\,\text{participant}\right)$$

$$\text{Dynamic decoding} \sim 1 + \text{training type}^{*}\ \text{pre} - \text{test LSK} + (1|\,\text{item}) + \left(1|\,\text{participant}\right)$$

TABLE 3 Model testing the effect of condition on dynamic reading scores.

Fixed effects	Estimate	St. Error	*t*-Value	*p*-Value	χ^2
PA	.191	.157	1.22	.225	1.55
LSK	−.130	.153	−.846	.398	1.55
DEC	−.061	.152	−.402	.688	1.55
Random effects	**Name**	**Variance**	**SD**		
Participant number	(Intercept)	2.51	1.58		
Item number	(Intercept)	.85	.92		
Residual		2.96	1.72		

Table 4 presents the results from all three planned models. We found that pre-test scores of sound isolation and letter-sound knowledge significantly predicted dynamic reading scores. This is in line with expectations: children who scored more highly on these pre-tests went on to score more highly on dynamic reading. It is interesting to note that this effect was not significant for sound deletion, suggesting that this measure may not have been sensitive for low attaining beginner readers (supported by our earlier observation of floor effects). However, none of the pre-test measures significantly interacted with condition to predict dynamic reading.

Question 3: Does type of training affect performance on standardized tests of phonological awareness and letter-sound knowledge?

For our final question, the model we used deviated slightly from the original model outlined in the analysis plan. We originally proposed a model with post-test score as the outcome and pre-test scores as the fixed effects. However, before conducting our analyses we realized that this was not the most appropriate model as we had the same measures at pre- and post-test. We therefore had the potential to test whether the increase was greater for some conditions over others. Therefore, instead of the planned models, three models were fitted with each standardized test as the outcome and condition and time as fixed effects.

$$\text{SoundIsolationScore} \sim 1 + \text{Condition}^{*}\text{Time} + (1|\text{item}) + (1|\text{participant})$$

$$\text{SoundDeletionScore} \sim 1 + \text{Condition}^{*}\text{Time} + (1|\text{item}) + (1|\text{participant})$$

$$\text{LSKScore} \sim 1 + \text{Condition}^{*}\text{Time} + (1|\text{item}) + (1|\text{participant})$$

The results for each of these models are summarized separately for each outcome (Tables 5–7). First, note that there was a significant effect of time for all three models. When sound isolation was the outcome, no significant effects of condition were found although the effect of DEC on sound isolation was approaching significance. There was also a significant interaction between the decoding condition and time. This is in line with the pattern highlighted earlier for greater improvement in sound isolation scores for the DEC condition. This pattern is depicted in Figure 1. For the sound deletion model, no significant effects of condition or interactions between condition and time were found (Table 6). For the model with letter-sound knowledge score as the dependent variable (Table 7), there was a significant interaction between the LSK condition and time at $p < .05$. However, note that the letter-sound knowledge model was not significant overall. Results for all planned analyses are included in the tables below, including non-significant models.

TABLE 4 Summary of models testing effects of pre-test scores and condition on dynamic reading scores.

Fixed effects	Estimate	St. Error	*t*-value	*p*-value	χ^2
PA	.142	.132	1.07	.285	1.77
LSK	−.150	.130	−1.16	.249	1.77
DEC	.008	.128	.063	.950	1.77
PreIsolation	.149	.033	4.50	<.001	40.0***
PreDeletion	.070	.065	1.07	.286	1.17
PreLSK	.072	.010	6.96	<.001	44.4***
PA: PreIsolation	−.021	.045	−.468	.640	3.14
LSK: PreIsolation	.074	.044	1.69	.093	3.14
DEC: PreIsolation	−.053	.044	−1.21	.229	3.14
PA: PreDeletion	−.105	.083	−1.27	.207	2.75
LSK: PreDeletion	.112	.075	1.49	.137	2.75
DEC: PreDeletion	−.007	.080	−.084	.933	2.75
PA: PreLSK	−.010	.014	−.742	.459	.806
LSK: PreLSK	−.000	.014	−.015	.988	.806
DEC: PreLSK	.010	.014	.771	.441	.806
Random effects	**Name**	**Variance**	**SD**		
Participant number	(Intercept)	1.69	1.30		
Item Number	(Intercept)	.81	.899		
Residual		2.94	1.71		

Note: $*p \le .05$; $**p \le .01$; $***p \le .001$.

Exploratory analyses

Our original analysis plan did not specify models to test the effect of trained versus untrained words on the dynamic reading task. The inclusion of untrained words was an important design feature which allows us to test the potential of each type of training to generalize to new words. We therefore created models to investigate whether word type (trained vs. new) influenced dynamic reading score, and whether the effect of condition varied for each word type.

$$\text{Dynamic decoding} \sim 1 + \text{Condition} + \text{WordType} + (1|\,\text{ItemNumber}) + \left(1|\,\text{ParticipantNumber}\right)$$

$$\text{Dynamic decoding} \sim 1 + \text{Condition}^{*}\text{WordType} + (1|\,\text{ItemNumber}) + \left(1|\,\text{ParticipantNumber}\right)$$

As shown in Table 8, a significant effect of word type was found. As shown in Figure 2, participants performed more highly for trained than untrained words, as expected. Importantly, there was also an interaction between word type and condition on dynamic reading score. This interaction was significant for all three conditions, suggesting that word type interacted differently with each condition in driving decoding outcomes. As shown in Figure 2 and Table 8, the advantage for trained words was greater for the PA and DEC conditions. In contrast, there was a smaller difference between trained and untrained words for the LSK condition. It is important to note that Figure 2 does not show an advantage for the LSK condition and that no significant condition effects were found on decoding performance in our first planned analysis. Instead, there is a trend for participants in the LSK condition to perform more poorly than the other two conditions on trained words, with the highest scores observed for the PA condition. All three conditions show similar scores for the untrained words, with a slight trend for lower performance for the decoding condition.

TABLE 5 Summary of models testing condition effects on sound isolation scores.

Fixed effects	Estimate	St. Error	*z*-Value	*p*-Value	χ^2
PA	.281	.226	1.24	.214	.76
LSK	.149	.222	.670	.503	.76
DEC	−.429	.223	−1.92	.055	.76
Time	1.56	.079	19.8	<.001	435***
PA: Time	−.161	.108	−1.49	.136	9.71**
LSK: Time	−.174	.106	−1.64	.101	9.71**
DEC: Time	.336	.108	3.12	<.01	9.71**
Random effects	**Name**	**Variance**	**SD**		
Participant number	(Intercept)	2.10	1.45		
Item number	(Intercept)	1.19	1.09		

Note: *$p \leq .05$; **$p \leq .01$; ***$p \leq .001$.

TABLE 6 Summary of models testing the effects of condition on sound deletion scores.

Fixed effects	Estimate	St. Error	*z*-value	*p*-Value	χ^2
PA	.068	.139	.488	.626	.224
LSK	−.025	.136	−.186	.853	.224
DEC	−.042	.134	−.315	.753	.224
Time	1.44	.091	15.8	<.001	269***
PA: Time	−.113	.126	.902	.367	.923
LSK: Time	.020	.124	.158	.874	.923
DEC: Time	.094	.122	.771	.441	.923
Random effects	**Name**	**Variance**	**SD**		
Participant number	(Intercept)	1.62	1.27		
Item number	(Intercept)	2.88	1.70		

Note: *$p \leq .05$; **$p \leq .01$; ***$p \leq .001$.

TABLE 7 Summary of models testing the effects of condition on letter-sound knowledge scores.

Fixed effects	Estimate	St. Error	*z*-Value	*p*-Value	χ^2
PA	.108	.182	.595	.552	.43
LSK	.015	.177	.087	.931	.43
DEC	−.124	.176	−.702	.483	.43
Time	1.87	.065	29.0	<.001	983***
PA: Time	−.114	.088	−1.30	.194	5.09
LSK: Time	.196	.086	2.27	<.05	5.09
DEC: Time	−.082	.0831	−.984	.325	5.09
Random effects	**Name**	**Variance**	**SD**		
Participant number	(Intercept)	3.27	1.81		
Item number	(Intercept)	2.94	1.72		

Note: *$p \leq .05$; **$p \leq .01$; ***$p \leq .001$.

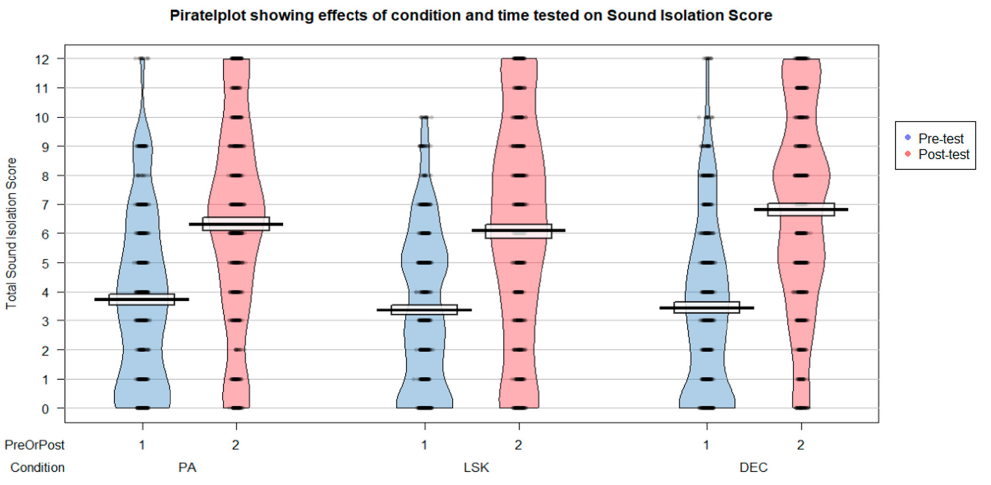

FIGURE 1 Graph illustrating the interaction between condition and time on sound isolation scores.

TABLE 8 Summary of models testing the effect of word type and condition on dynamic reading score.

Fixed effects	Estimate	St. Error	*t*-value	*p*-Value	χ^2
PA	.191	.157	1.22	.225	1.55
LSK	−.130	.157	−.846	.398	1.55
DEC	−.0612	.152	−.402	.689	1.55
WordType	−1.05	.283	−3.70	<.001	11.8***
PA: WordType	−.168	.061	−2.77	<.01	43.2***
LSK: WordType	.388	.059	6.54	<.001	43.2***
DEC: WordType	−.221	.059	−3.74	<.01	43.2***
Random effects	**Name**	**Variance**	**SD**		
Participant number	(Intercept)	2.51	1.58		
Item number	(Intercept)	.59	.77		
Residual		2.94	1.71		

Note: $*p \leq .05$; $**p \leq .01$; $***p \leq .001$.

DISCUSSION

The current study aimed to identify the relative benefits of training each specific element of the decoding process through three carefully matched conditions using the same set of stimuli. In the PA training, participants were asked to segment and then blend the phonemes of the stimuli without ever seeing the words. In the LSK condition, participants saw each word but did not produce the blended form. Instead, they were asked to say the phoneme corresponding to each grapheme within the word. Finally, the DEC condition included both blending (as in the PA condition) and the written form (as in the LSK condition). Our study focused on the differential impact of each type of training, comparing these three conditions to each other, rather than to a control condition. Our pilot data (Appendix S1) already demonstrated a clear advantage of similar types of training over an untreated control which justified our current focus on whether there were any differences between the experimental conditions. Our study is the first to experimentally break down the components of the decoding process and compare the effect of training blending, letter sounds or both on decoding outcomes (see Castles et al., 2009, for a similar approach to training letter-sound knowledge).

We found no differences between the conditions on our reading outcome task (dynamic reading). However, we did find that participants from the DEC condition showed greater improvement on the

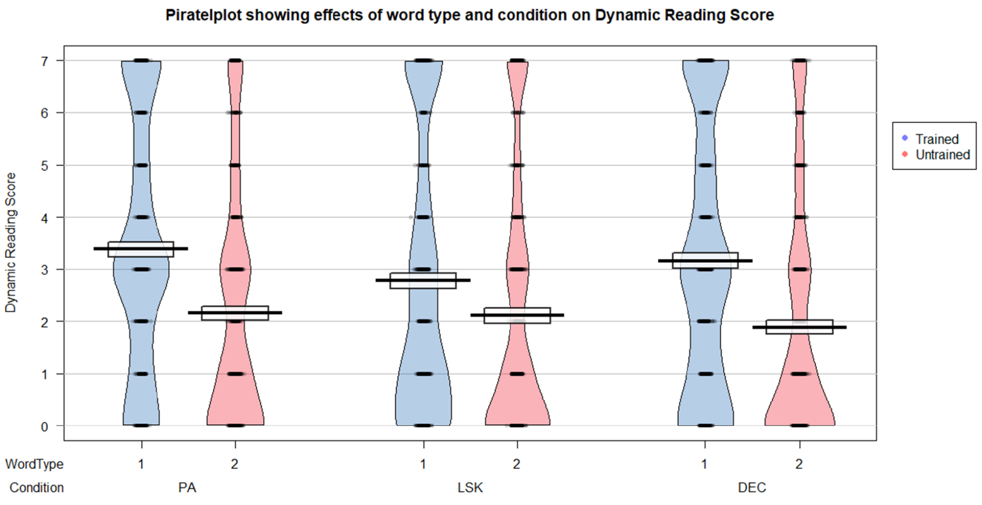

FIGURE 2 Graph illustrating the effect of word type on dynamic reading by condition.

sound isolation subtest when compared with participants from the other training conditions. This is in line with expectations that including blending as part of the training would be beneficial for phonemic awareness. However, it is surprising that this advantage was not also observed for the PA condition and suggests that the presence of the written form, alongside blending training, was beneficial for children's development of phonological awareness. Although print has been assumed to be beneficial when learning phonological awareness (e.g., National Reading Panel (US), 2000), this has not previously been tested using carefully matched conditions. There was also a trend for greater improvement for participants in the LSK condition on the letter-sound knowledge test. Again, this is in line with expectations that the condition targeting letter-sound knowledge should lead to the greatest improvement in this skill, although it is important to emphasize that this was not a significant pattern.

As expected, children's pre-test scores (on sound isolation and letter-sound knowledge) were positively predictive of performance on our outcome reading test (dynamic reading) but did not significantly interact with any of the conditions suggesting there was no differential effect of condition based on children's pre-test scores. Additionally, our exploratory analyses showed a significant effect of word type such that participants performed significantly better with trained words on the dynamic reading test versus untrained words. This is in line with expectations that training would have the greatest impact on the words included within the training sessions. Condition also significantly interacted with word type such that the advantage for trained words was greatest for the two conditions that included blending (PA and DEC) and smallest for the LSK condition. One interpretation is that training on blending is less likely to generalize to new words than training on letter sounds. Importantly, there was no overall effect of condition on reading outcome, so this interaction does not suggest that the LSK condition was most beneficial overall. It also does not mean that there was no generalization for the PA and DEC conditions, only that there was a greater advantage for trained words, for these conditions. Given that reading scores were close to floor at the beginning of the study, any success in reading untrained words is positive. Especially considering the difficulty of the words used in the dynamic reading task (selected specifically to avoid words taught in class; see Appendix S3). Below, we discuss the implications of our findings from each question in turn.

Question 1: Does type of training (PA vs. LSK vs. DEC) affect post-test performance on a dynamic decoding task?

Our key finding is that there was no significant effect of condition on performance on the dynamic decoding post-test, meaning that all three training types performed similarly and that training one specific element of the decoding process did not prove to be more advantageous than another for decoding outcomes. On one level, this finding appears in line with Pfost et al. (2019) who found that an intervention which trained both letter knowledge and phonological awareness improved those respective skills but not overall word reading. However, our interpretation differs for two reasons. First, we worked with beginner readers who mostly scored at floor on a standardized test of reading at the beginning of the study, so we did not attempt to measure improvement in word reading. Instead, we assume that word reading must have improved since participants scored above floor on the dynamic reading task at the end of the study. Second, we focus on the difference between the three conditions, each of which target one or more elements of the decoding process. We conclude that there was not a differential effect of condition on decoding outcomes and instead training on any part of the decoding process is likely to be beneficial. Nevertheless, it is important to consider that our outcome measure may not have been sensitive enough to reveal effects of condition. A dynamic task certainly had advantages over static measures (Spector, 1992) by avoiding floor effects by providing successive prompts to success. However, it inevitably remained a difficult task and our participants required prompting for many items (approximately half of all responses required a prompt), and even with prompts, scores of zero were achieved (31% of all responses resulted in a score of zero). Teachers also commented that the words we used did not follow the consonant–vowel–consonant (CVC) structure that is often taught in reception. The words used in the current study were chosen so that they could be decoded using grapheme–phoneme correspondences taught during the first year of school but avoiding words that had been explicitly taught in class. The sensitivity of this task could have been improved using a larger range of words, perhaps including nonwords with a CVC structure, or by using simpler real words alongside a way to factor in their familiarity level in our analysis.

Another possible reason for the lack of a condition effect is that our training was relatively short. Most of the studies discussed previously used interventions lasting many months compared to our training which consisted of 10 sessions over 5 weeks. The current training study may not have been extensive enough to judge differential effects of the conditions on the participants' reading.

Question 2: Is any effect of training type on post-test performance on a dynamic decoding task moderated by pre-test PA or LSK scores?

Pre-test performance on the sound isolation and letter-sound knowledge test significantly predicted post-test dynamic decoding performance, in line with expectations that these are crucial underlying skills driving reading development (Hatcher et al., 2004; Hulme et al., 2012; Shapiro et al., 2013). We found no significant effect of pre-test sound deletion score, perhaps due to this task being relatively difficult for the children, compared with our other pre-tests. We found no interactions between these pre-test scores and condition on decoding outcomes, suggesting that pre-test PA and LSK scores did not moderate the effect of condition on decoding outcomes. This is surprising since we would expect that targeted training would be most beneficial, for example, a child struggling with LSK would benefit the most from LSK training. However, as highlighted by Carroll et al. (2016) pre-reading skills are heavily intercorrelated and so identifying one specific weakness may not be feasible for beginner readers. Instead, it is likely to be more important that training directly targets the crucial components of the skill the children are aiming to achieve. In our study, each training condition addressed one or more crucial component of the decoding process, and these were found to be equally beneficial for all children, regardless of their initial pre-test performance.

Question 3: Does training type affect performance on standardized tests of phonological awareness and letter-sound knowledge?

Contrary to our predictions, it was the DEC condition which delivered the greatest benefit on the sound isolation test, rather than the PA condition. This is consistent with Hatcher et al. (2004)'s finding that a phonological awareness intervention paired with reading was beneficial for overall phonological awareness. However, our study was the first to make a direct comparison across matched training conditions. Our findings can be interpreted in line with two key theories. First, that there is a reciprocal relationship between phonological awareness and letter knowledge (e.g., Cunningham et al., 2021), whereby as letter knowledge develops, this leads to improved phonological awareness and vice versa. Although all training conditions included at least one of these components, only DEC included both. This finding contributes to the evidence base supporting interventions that include both blending and letter knowledge (e.g., Bowyer-Crane et al., 2008; Hulme et al., 2012).

Our findings are also consistent with theories that orthographic cues are helpful to solve phonological awareness tasks. In the DEC condition, participants had access to the printed words while practicing their blending skills, unlike participants in the PA condition. Although the words used in the sound isolation test were different from the words they were trained on, their practice in blending with printed cues may have helped them to use a similar strategy of visualizing the letters when completing the sound isolation task (Deacon et al., 2018). Magnan and Ecalle (2006) also found that a letter-sound mapping intervention improved phonological awareness in dyslexic children, consistent with this interpretation. It is important to note that we only found a condition effect on sound isolation. There were no condition effects on the performance of sound deletion, potentially because of the higher levels of floor performance. Muter et al. (2004) also found poor performance on phoneme deletion for children of the same age range, and Carroll et al. (2003) suggest that the phoneme deletion task requires a level of explicit phoneme awareness which reception-aged children may not have developed yet.

We also found a trend for a condition effect on letter-sound knowledge. Although our model was not significant, this pattern is in line with expectations that the training that most closely matches the outcome is most effective.

Exploratory analysis: Does word type affect performance on the post-test dynamic decoding task? Does word type affect the condition effects on the post-test dynamic decoding task?

As expected, our participants performed significantly better on trained words than untrained words. This is reassuring in that it demonstrates that our training had the expected impact. It was interesting to see that word type had a different effect for each of the training conditions. In particular, the two conditions which included blending (PA and DEC) showed a greater advantage for trained over untrained words. The LSK condition showed a significantly smaller advantage for trained over untrained words, perhaps because the participants did not benefit from exposure to the blended form of the word during training. It is interesting that dynamic reading performance was at least as great for the PA condition as the DEC condition, even though children in the PA condition were seeing the printed word for the first time when completing the dynamic decoding outcome task. Overall, this pattern suggests that familiarity with the blended form of the spoken word is helpful when decoding that word from print, consistent with theories of semantic involvement in word reading (Ricketts et al., 2016). This finding also highlights the importance of carefully choosing items to use in training sessions, because the benefit of the training will be strongest for these particular words.

Limitations

A key limitation of the current study was the lack of a baseline dynamic decoding score. Even though children would be expected to score at floor for a static task, we may have gained above-floor dynamic reading scores, and this would have increased the power of our analyses. However, the disadvantage of using a dynamic task at the beginning of the study would be to make the training conditions less clearly distinct (i.e., all participants would then have seen the word in print and heard the blended form). This may have been less of a problem with a more extended training period, and an increased gap between pre- and post-tests. Another limiting factor of our study was the lack of a control group. Including a control group would have provided a baseline with which we could compare the results of each training type, potentially revealing clearer effects of each condition.

Implications for education

We have highlighted the benefit of training that targets key components of the decoding process. Although targeting any of these components is likely to be beneficial, the clearest benefit was for the conditions that included training on blending, and in particular for the condition that included both blending and access to the written form of the word. This type of training has the advantage of familiarizing children with both the written and spoken forms of words and can also be readily combined with direct teaching of vocabulary teaching. Although our training was not designed to teach vocabulary, we included a sentence and picture context to make the training more interesting for the children. They appeared to enjoy this aspect and readily recalled the context of the words during the outcome tasks, even though this information was not asked for. Selecting words that children are likely to encounter, whilst also targeting those they may struggle with is key to the success of this approach. A valuable tool in selecting words is the "three tiers" framework for vocabulary teaching (Beck et al., 2013). Children would most likely benefit from targeted teaching of "tier 2" words (such as *prefer, explore, special*). Delivering individualized decoding training using words that are most useful for children's oral and written language comprehension would be expected to have dual benefits on reading and vocabulary outcomes.

CONCLUSION

We compared the effects of three training conditions developed to target one or more components of the decoding process. Low attaining readers were selected for the study and children in all conditions improved in their phonological awareness and letter knowledge, and using a dynamic reading task, were able to decode more complex words than expected for their age. Although the differences between conditions were small, there was a pattern for the conditions that included blending skills to be more beneficial, particularly when decoding words that had been included in training. Future research is needed to develop this approach further and investigate the effectiveness of a programme of training on blending skills combined with direct vocabulary instruction for struggling beginner readers.

AUTHOR CONTRIBUTIONS

Charlotte Webber: Formal analysis; methodology; writing – original draft; writing – review and editing. **Hetal Patel:** Data curation; formal analysis; project administration; validation; writing – original draft; writing – review and editing. **Anna Cunningham:** Conceptualization; formal analysis; methodology; supervision; writing – review and editing. **Amy Fox:** Conceptualization; data curation; funding acquisition; methodology; project administration; supervision; writing – review and editing. **Janet Vousden:** Conceptualization; formal analysis; funding acquisition; methodology; software; writing – review and editing. **Anne Castles:** Conceptualization; funding acquisition; methodology; writing – review and editing. **Laura Shapiro:** Conceptualization; formal analysis; funding acquisition; methodology; project administration; resources; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

This study is supported by the British Academy and Leverhulme Trust (SRG\170066; PI Shapiro, Co-I Castles, Fox and Vousden).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on the Open Science Framework at https://osf.io/af56j/; DOI: 10.17605/OSF.IO/AF56J.

ORCID

Laura Shapiro https://orcid.org/0000-0002-3276-457X

REFERENCES

Babayiğit, S., & Shapiro, L. R. (2020). Component skills that underpin listening comprehension and reading comprehension in learners with English as first and additional language. *Journal of Research in Reading*, *43*, 78–97. https://doi.org/10.1111/1467-9817.12291

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823*.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry*, *49*(4), 422–432.

Brooks, G. (2016). *What works for children and young people with literacy difficulties. The effectiveness of intervention schemes* (5th ed.). Frensham, Dyslexi-SpLD Trust.

Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*(5), 913–923.

Carroll, J. M., Solity, J., & Shapiro, L. R. (2016). Predicting dyslexia using prereading skills: The role of sensorimotor and cognitive abilities. *Journal of Child Psychology and Psychiatry*, *57*(6), 750–758.

Castles, A., Coltheart, M., Wilson, K., Valpied, J., & Wedgwood, J. (2009). The genesis of reading ability: What helps children learn letter–sound correspondences? *Journal of Experimental Child Psychology*, *104*(1), 68–88.

Cunningham, A., & Carroll, J. (2011). Age and schooling effects on early literacy and phoneme awareness. *Journal of Experimental Child Psychology*, *109*(2), 248–255.

Cunningham, A. J., Burgess, A. P., Witton, C., Talcott, J. B., & Shapiro, L. R. (2021). Dynamic relationships between phonological memory and reading: A five year longitudinal study from age 4 to 9. *Developmental Science*, *24*(1), e12986.

Deacon, S. H., Pasquarella, A., Marinus, E., Tims, T., & Castles, A. (2018). Orthographic processing and children's word reading. *Applied PsychoLinguistics*, *40*(2), 509–534. https://doi.org/10.1017/s0142716418000681

Department for Education. (2016). *Achievement of 15-YearOlds in England: PISA 2015 National Report*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/574925/PISA-2015_England_Report.pdf

Department for Education and Skills. (2007). *Letters and Sounds: Principles and Practice of High Quality Phonics*.

Education Policy Institute. (2018). *Education in England: Annual Report 2018*. Retrieved from https://epi.org.uk/publications-and-research/annual-report-2018

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393–447.

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*(4), 375–384.

Elbro, C., & Petersen, D. M. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology*, *96*(4), 660–670.

Hatcher, P. J., Hulme, C., & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry*, *45*(2), 338–358.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science*, *23*(6), 572–577.

Hulme, C., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Snowling, M. J. (2009). *YARC York assessment of reading for comprehension. Early reading*. GL Publishers.

Ilie, S., Sutherland, A., & Vignoles, A. (2017). Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation. *British Educational Research Journal*, *43*(2), 253–274.

Lloyd, S., Wernham, S., Jolly, C., & Stephen, L. (1998). *The phonics handbook*. Jolly.

Machin, S., McNally, S., & Viarengo, M. (2018). Changing how literacy is taught: Evidence on synthetic phonics. *American Economic Journal: Economic Policy*, *10*(2), 217–241.

Magnan, A., & Ecalle, J. (2006). Audio-visual training in children with reading disabilities. *Computers & Education*, *46*(4), 407–425.

Mann, V. A., & Foy, J. G. (2003). Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia*, *53*(1), 149–173.

Masterson, J., Stuart, M., Dixon, M., & Lovejoy, S. (2010). Children's printed word database: Continuities and changes over time in children's early reading vocabulary. *British Journal of Psychology*, *101*(2), 221–242.

McArthur, G., & Castles, A. (2017). Helping children with reading difficulties: Some things we have learned so far. *Npj Science of Learning*, *2*(1), 7.

McLeod, S., Van Doorn, J., & Reed, V. A. (2001). Normal acquisition of consonant clusters. *American Journal of Speech-Language Pathology.*, *10*, 99–110.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665–681.

National Reading Panel (US). (2000). *National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on Reading and its implications for Reading instruction*. National Institute of Child Health and Human Development, National Institutes of Health.

NFER. (2018). *NFER Education Briefings: Key insights for England from PIRLS, THRASS and PISA*. Retrieved from http://www.easybib.com/reference/guide/apa/website

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal – a longitudinal-study of 1st-grade children. *Merrill-Palmer Quarterly-Journal of Developmental Psychology*, *33*(3), 283–319.

Pfost, M., Blatter, K., Artelt, C., Stanat, P., & Schneider, W. (2019). Effects of training phonological awareness on children's reading skills. *Journal of Applied Developmental Psychology*, *65*, 101067.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. http://www.r-project.org

Ricketts, J., Davies, R., Masterson, J., Stuart, M., & Duff, F. J. (2016). Evidence for semantic involvement in regular and exception word reading in emergent readers of English. *Journal of Experimental Child Psychology*, *150*, 330–345.

Rose, J. (2006). *Independent review of the teaching of early reading: Final report*. Department for Education and Skills.

Roy, P., Chiat, S., & Dodd, B. (2014). *Language and socioeconomic disadvantage: From research to practice*. City University London. http://openaccess.city.ac.uk/4989/

Rvachew, S., Nowak, M., & Cloutier, G. (2004). Effect of phonemic perception training on the speech production and phonological awareness skills of children with expressive phonological delay. *American Journal of Speech-Language Pathology.*, *13*(3), 250–263.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038.

Shapiro, L. R., Carroll, J. R., & Solity, J. (2013). Separating the influences of pre-reading skills on early word and nonword reading. *Journal of Experimental Child Psychology*, *116*(2), 278–295.

Shapiro, L. R., & Solity, J. (2016). Differing effects of two synthetic phonics programmes on early reading development. *British Journal of Educational Psychology*, *86*(2), 182–203.

Snowling, M. J. (2001). From language to reading and dyslexia 1. *Dyslexia*, *7*(1), 37–46.

Spector, J. E. (1992). Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of Educational Psychology*, *84*(3), 353–363.

The Highland Council. (2016). *Phonological awareness informal assessment for school-age children*. Designed by: Care & Learning Service – Speech and Language Therapy.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Webber, C., Patel, H., Cunningham, A., Fox, A., Vousden, J., Castles, A., & Shapiro, L. (2024). An experimental comparison of additional training in phoneme awareness, letter-sound knowledge and decoding for struggling beginner readers. *British Journal of Educational Psychology*, *94*, 282–305. https://doi.org/10.1111/bjep.12641