



## OPEN ACCESS

## EDITED BY

Seth King,  
The University of Iowa, United States

## REVIEWED BY

Kimberly Murphy,  
Old Dominion University, United States  
Fahisham Taib,  
Universiti Sains Malaysia, Malaysia

## \*CORRESPONDENCE

Lisa M. D. Archibald  
✉ larchiba@uwwo.ca

RECEIVED 22 November 2022

ACCEPTED 20 July 2023

PUBLISHED 16 August 2023

## CITATION

Archibald LMD, Davison C, Kuiack A, Doytchinova S, King C, Shore-Reid D, Cook P and Vollebregt M (2023) Comparing community-based reading interventions for middle school children with learning disabilities: possible order effects when emphasizing skills or reasoning. *Front. Educ.* 8:1104945. doi: 10.3389/feduc.2023.1104945

## COPYRIGHT

© 2023 Archibald, Davison, Kuiack, Doytchinova, King, Shore-Reid, Cook and Vollebregt. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Comparing community-based reading interventions for middle school children with learning disabilities: possible order effects when emphasizing skills or reasoning

Lisa M. D. Archibald \*, Christine Davison<sup>1</sup>, Alyssa Kuiack<sup>1</sup>, Stella Doytchinova<sup>1</sup>, Colin King<sup>2</sup>, Deborah Shore-Reid<sup>3</sup>, Paul Cook<sup>2,3</sup> and Meghan Vollebregt<sup>1</sup>

<sup>1</sup>School of Communication Sciences and Disorders, The University of Western Ontario, London, ON, Canada, <sup>2</sup>Faculty of Education, The University of Western Ontario, London, ON, Canada, <sup>3</sup>Learning Disabilities Association, London, ON, Canada

**Introduction:** There is an abundance of community-based reading programs for school-age children who are struggling learners. The aim of this study was to compare two community-based programs (i.e., skill and reason-based programs) and to analyze any complementary benefits.

**Methods:** In this randomized cross-over study, 20 children completed two 8-week literacy intervention programs. The skills-based program, *Leap to Literacy*, focused on explicit teaching and repeated practice of the five key components of literacy instruction (phonemic awareness, phonics, fluency, vocabulary, comprehension). The reason-based program, *Wise Words*, focused on morphological knowledge, hypothesis testing, and critical thinking.

**Results:** Results revealed study-wide improvements in phonemic awareness, nonword reading, passage reading accuracy, spelling words and features, and affix identification. There were consistent program by program order effects with robust effects of completing the skills-based program first for phonemic awareness, the reason-based program first for passage reading accuracy, and both programs for affix identification. A significant increase in an oral language measure, recalling sentences, was observed for the group who completed the reason-based program first, although they also started off with a lower initial score.

**Discussion:** Findings indicated improvements from participating in either program. The observed order effects suggest potential additive effects of combining reason- and skills-based approaches to intervention.

## KEYWORDS

community-based, skill-based, reason-based, reading programs, randomized cross-over design

## 1. Introduction

For a variety of reasons, there is an abundance of community-based reading programs for school-age children who are struggling learners. Extracurricular tutoring, however, is associated with an economic cost to families, and social and cognitive costs for the children who participate

in them after a full school day. Given these costs, it is important that short-term, community-based reading interventions be evidence-based and effective. [National Institute of Child Health and Human Development et al. \(2000\)](#) identified five key components of effective reading instruction: phonological awareness, phonics, reading fluency, vocabulary, and text comprehension. An intervention focused on explicit instruction of these skills and offering ample opportunities for practice should have a positive impact on a child's reading. Such a skills-based approach, however, may not address all of the needs of a struggling learner in middle school (i.e., grades 4 to 8; 8 to 13 years old). Rather, these children might need an alternate approach such as using discussion and critical thinking to understand the patterns and conventions of the written language they are learning (English in the present study). A reason-based approach focused on the structure of words ([Bowers and Kirby, 2010](#)) would necessarily target knowledge of morphemes, the smallest segment of language that has meaning (i.e., words or part words). The inclusion of morphological awareness as an intervention focus might be particularly important for struggling learners given reports of greater morphological awareness intervention benefits for those with language or learning disabilities compared to typical readers ([Goodwin and Ahn, 2010](#); [Goodwin and Ahn, 2013](#)). The purpose of the present study was to compare skills- and reason-based reading intervention programs in struggling readers or spellers. Specifically, we employed a randomized, cross-over trial to examine outcomes related to oral language, phonemic awareness, word reading, reading fluency, and spelling in 8- to 13-year-old children completing community-based skills- and reason-based short-term interventions sequentially in differing orders.

Canadians concerned with their children's academic success are turning to private education. There is considerable evidence that extracurricular reading interventions are effective ([Ritter et al., 2009](#)). Similarly, in a meta-analysis focused on low achieving students, out-of-school programs were associated with a positive effect on reading, especially for children in the lower elementary and high school grades ([Lauer et al., 2006](#)). The effect size of 0.13 based on a random-effects model was small, although considered nontrivial and in line with previous meta-analyses ([Cooper et al., 1996](#)). Several factors moderated the effects including program focus and duration, student grouping, and grade. Non-zero effects were observed for primarily academic programs, programs of 44–210 h (not more or less), and programs working with students one-on-one. Large positive effects were observed for students in the lower elementary and high school grades. The authors noted that further investigation was needed of program elements contributing to a positive effect for upper elementary students (grades 3 to 8), which is in line with the purpose of the present study.

[National Institute of Child Health and Human Development et al. \(2000\)](#) described five critical skills for literacy development with three supporting word recognition skills (phonemic awareness, phonics, reading fluency), and two relevant to language comprehension skills (vocabulary, text comprehension). Targeting these five skills in intervention can be expected to be beneficial to struggling readers. Indeed, [Kim et al. \(2017\)](#) designed a multicomponent reading program for adolescent children where both foundational skills (i.e., phonemic awareness, phonics, word recognition) and comprehension skills were targeted. This supplemental (within school) reading program (3–5 times per week for a school year) involved instruction to support word-reading, fluency, vocabulary, comprehension, and

peer talk to promote reading engagement and comprehension. Compared to a matched control group, students in grades 6–8 performed better on measures of word recognition, reading comprehension, and morphological awareness. Struggling readers, as well, benefit from a multicomponent approach: [Afacan et al. \(2017\)](#) reported significantly greater improvements in reading skills for students with intellectual disability who had participated in a multicomponent compared to single skill focused intervention. In a review of reading intervention research, [Donegan and Wanzek \(2021\)](#) reported participation in a multicomponent intervention had significant predicted effects for both foundational skills and comprehension skills. A multicomponent approach to reading instruction might be particularly important for students whose educational curriculum has focused on balanced literacy or lacked a systematic approach to phonics instructions, as was the case for the students from Ontario in the current study ([Ontario Human Rights Commission, 2022](#)).

It has been acknowledged, however, that despite timely and intensive interventions emphasizing phonemic awareness, phonics, and repeated practice, approximately 25% of struggling readers will continue to have difficulty learning to read ([Reed, 2008](#)). Rather than repeated skill practice, these children might benefit from an approach to literacy instruction that emphasizes understanding and reasoning. One candidate reason-based approach to literacy instruction is structured word inquiry, which focuses on morphemes, the smallest unit of meaning in a word ([Bowers and Kirby, 2010](#)) and teaches grapheme-phoneme correspondences in the context of morphology. Instruction in morphological awareness was not included in the [National Institute of Child Health and Human Development et al. \(2000\)](#) despite evidence that morphological instruction positively impacts reading, spelling, vocabulary, and morphological skills ([Bowers et al., 2010](#)) and brings particular benefits for those learning English as an additional language ([Goodwin and Ahn, 2013](#)) and for struggling readers including those with a learning disability ([Goodwin and Ahn, 2010, 2013](#)) or language disorder ([Goodwin and Ahn, 2010](#)). Morphological intervention focuses on sublexical features of language, which, in turn, influence literacy skills at the lexical level. As morphological awareness increases, children are able to use this knowledge to infer the meaning of complex words and understand their spelling. Indeed, morphological awareness has been associated with word reading accuracy ([Carlisle, 2000](#)), spelling accuracy ([Deacon et al., 2009](#)), and vocabulary knowledge ([Carlisle, 2007](#)). Although effective for younger children, morphological instruction may be particularly beneficial for children in the upper elementary grades as in the present study because morphological awareness compared to phonemic awareness has been found to be a better predictor of reading in this age range ([Singson et al., 2000](#); [Kirby et al., 2012](#)). Importantly, morphological intervention has been found to be more effective when it is part of a comprehensive literacy instruction ([Bowers and Kirby, 2010](#)) and incorporates writing ([Nunes et al., 2003](#)).

Structured word inquiry emphasizes a scientific approach to the investigation of words where a student and instructor work together to discover the deepest structure that represents meaning and explains spelling ([Bowers and Kirby, 2010](#)). A hypothesis-driven approach is used to gain understanding of how a word's historical roots connects to the word's morphological structure, which in turn connects not only to the phonological realization of the word but also

many other words sharing the same central meaning. To date, evidence regarding the effectiveness of structured word inquiry is limited. Three studies have compared business-as-usual classroom instruction to classroom instruction in structured word inquiry for children between the ages of 5 and 11 years and reported improvements in morphological knowledge and its use in understanding untaught words after 20 lessons (Bowers and Kirby, 2010), in word reading and spelling after daily instruction for 6 weeks (Devonshire et al., 2013), and in spelling after weekly sessions for 9 weeks (Devonshire and Fluck, 2010). Murphy and Diehm (2020) reported the use of a structured word inquiry approach with 10 8–10-year-old struggling readers or spellers delivered in a 6-week (11 sessions) summer camp and observed improved reading and/or spelling abilities with the largest gains in the spelling of affixes with polymorphemic words. Although preliminary, these findings indicate that a structured word inquiry approach has the potential to positively impact reading and spelling in 8–13-year-old struggling readers, which was the focus of the present study.

## 2. Methodology

The present study came about as a result of a partnership between two community-based organizations providing services to children with a variety of needs including learning disabilities. In order to meet the needs of individuals with learning disabilities that include reading challenges, two of our authors (DSR, PC) together with Kathy Clark, another colleague from the local chapter of the Ontario Learning Disabilities Association (LDA) designed a skills-based program called *Leap to Literacy*. The program involves explicit instruction and repeated practice in the five key components of literacy instruction (National Reading Panel Report (NICHHD), 2000): phonemic awareness, phonics, fluency, vocabulary, and text comprehension. As part of this program, skills are reinforced and developed further through guided and independent practice with a computer-based reading program that targets the five critical components of reading instruction (LexiaCore 5 or LexiaPowerUp) and has been found to improve reading skills when delivered either in a classroom setting (Wilkes et al., 2016) or through individual instruction (Schechter et al., 2015). In between Leap to Literacy sessions, children are encouraged to complete Lexia sessions for half an hour per day.

As part of services offered through the local Child and Youth Development Clinic, four of the authors (LA, SD, AK, CD) together with other graduate students in speech-language pathology at The University of Western Ontario created a reason-based program called *Wise Words*. Based on the work of Bowers and Kirby (2010), *Wise Words* uses a structured word inquiry approach to engage children as partners in developing an understanding of the morphological structure of a word both in terms of meaning and impact on spelling. It is important to note that although there were key differences in the underlying approach employed in the Leap to Literacy and *Wise Words* programs, both programs aimed to be as comprehensive as possible. For example, the Leap to Literacy program encouraged discussion of word meaning and relationships and the *Wise Words* program included grapheme-phoneme correspondences and phoneme-level skills in all word study. Given this overlap, we understood that it would be challenging to uncover differential

intervention effects. Nevertheless, given the different underlying motivations and potential differences in impact of our two programs, we undertook a preliminary study to compare the different or complementary effects of our Leap to Literacy and *Wise Words* programs.

In the present study, 20 children in grades 4 to 8 were randomly assigned to complete 8 weekly, virtual, 1.25-h sessions of either Leap to Literacy or *Wise Words*, and after an 8 week break, completed the opposite program. Access to the software program, Lexia, was provided to all participants in the study. One purpose of the study was to compare language, reading, and spelling outcomes for the skills-based Leap to Literacy and reason-based *Wise Words* program. Consistent benefits of one program over another across all participants and regardless of order would provide convincing evidence in favor of that program. An advantage for the Leap to Literacy program might indicate that participants were in need of a structured approach individualized to the skills they needed to practice whereas greater benefits for the *Wise Words* program would suggest that meaning and understanding were important scaffolds for furthering literacy skills. A second aim of the study was to examine complementary benefits of the programs. Improved results for children completing the skills-before reason-based program would indicate that developing mastery over foundational skills facilitates engagement in the discussion and problem-solving associated with developing a deep understanding of word structure. On the other hand, an advantage for completing the reason-based program first would suggest that having an explanation and understanding of word structure would promote skill acquisition through practice. We anticipated that potential complementary effects might be more complicated and individualized than captured by these group-level hypotheses. Although we expected that our small, preliminary study would not allow us to uncover more nuanced patterns in our data, we planned to examine individual profiles and responses for possible indications for future studies.

### 2.1. Study design

#### 2.1.1. Participants

The local branch of Learning Disabilities Association (LDA) of Ontario and the University's Child and Youth Development Clinic each advertised their respective extracurricular intervention programs for members or struggling readers and spellers, respectively. Programs were advertised via posting on their respective websites, sending electronic fliers to their members, and connecting with learning resource teachers at local school boards. Advertisements were aimed at children with learning disabilities, although all struggling learners could register for the programs. All registrants who were currently enrolled in grades 4 to 8 and had not previously completed a *Wise Words* program (i.e., participants who already completed *Wise Words* level 1 were not eligible as they had already completed part 1 of the intervention) were invited to be in the study upon registration. Recruitment continued until 20 children were recruited. Ages (years; months) ranged from 8;9 to 13;5 ( $M = 10.8$ ;  $SD = 1.4$ ). All but three of the participants had a diagnosis of Learning Disability ( $n = 14$ ) and/or Auditory Processing Disorder ( $n = 3$ ), Attention Deficit Hyperactivity Disorder ( $n = 4$ ), or Developmental Language Disorder ( $n = 1$ ). In all, seven participants had more than one diagnosis. The 17 participants with a diagnosis were on Individualized Education Plans at their

school. For the three children without diagnoses, one was reported to use assistive technology for reading and writing, and the parents of the remaining two both reported ongoing struggles with reading and writing. English was the first language for 17 of the 20 participants (85%), although all participants were considered proficient in English.

Randomization to determine the intervention order was completed in blocks of 4. We created 6 orders of 4 (i.e., AABB, ABBA, ABAB, BBAA, BABA, BAAB) using an electronic random number generator. For each set of 4 participants recruited, we randomly selected one of these order blocks and assigned participants to intervention order A or B based on the selected intervention order and recruitment order. Once the initial assessment was completed at time 1, we determined through random selection that Intervention Order A was Leap to Literacy first then Wise Words, and Intervention Order B was the opposite. Thus, 10 participants were assigned to complete Leap to Literacy in Term 1 and Wise Words in Term 2 (LLWW), and 10 participants completed the opposite order (WWLL).

Parents were asked to report their highest education level attained, child's first language, concerns regarding their child's language, reading, and math skills, and how much time the child spent at home on activities such as homework, parent created or software-based educational activities, and independent reading. When parents provided a range for activity minutes, the higher estimate was used.

No home activities were reported for one child in the WWLL group, and one child in the LLWW group had completed a previous extracurricular program similar to the Leap to Literacy program. The highest level of education attained was reported for each of 2 parents for 19 of 20 participants (for the remaining participant, no response was provided). In all cases either one ( $n=3$ ) or both parents completed at least some postsecondary education (i.e., college or university). The child's first language was reported to be English for 17 participants, and Polish, Arabic, or Spanish for the remaining three. Table 1 shows the participant demographics for the LLWW and WWLL groups as well as results on the initial study (i.e., baseline) testing described below. Independent sample t-tests comparing the groups at baseline indicated no significant differences on any variable,  $t < 1.7$ ,  $p > 0.05$ , all cases. Notably, however, Cohen's  $d$  effect sizes were larger than 0.7 for home activities, recalling sentences (raw and scaled score), and percent words read correctly in a passage. Standard or scaled scores were available for one oral language (recalling sentences) and one reading fluency (word/nonword/total) measures, and indicated that the participants had low reading scores of about 1  $SD$  below the standardized mean but average oral language scores of within 1  $SD$  of the scaled mean. Seven participants (3 in the LLWW and 4 in the WWLL) scored above 87 on the available standardized measures. Of these, 3 were reading at least 2 grade levels below their grade on the

TABLE 1 Mean (standard deviation) demographic and baseline data for the groups completing the Leap to Literacy (LLWW) or Wise Words (WWLL) programs first.

Measure	Score type	Maximum score	Intervention groups	
			LLWW	WWLL
Age in years	Years	n/a	11.0 (1.3)	10.7 (1.5)
Home activities	Minutes/week	n/a	103 (97.1)	192.4 (152.4)
<b>Oral language measure</b>				
Recalling sentences	Raw	78	59.5 (8.1)	52.5 (8.4)
	Scaled		12.1 (2.7)	10.2 (2.5)
	Scaled Score Range		8–15	8–14
<b>Word recognition measures</b>				
Phonemic awareness probe	Raw	29	20.2 (4.3)	18.5 (3.9)
San Diego grade level	Grade	n/a	3.3 (1.6)	2.9 (2.3)
<b>Reading fluency measures</b>				
Sight word efficiency (TOWRE)	Raw	108	52.2 (11.5)	45.7 (22.4)
	Standard	n/a	86.3 (10.0)	80.1 (14.2)
Phonemic decoding (nonword) efficiency (TOWRE)	Raw	66	21.5 (8.1)	24.2 (12.5)
	Standard	n/a	86.7 (8.9)	88.8 (11.3)
TOWRE total score	Standard	n/a	83.8 (10.2)	81.6 (14.5)
	Standard score range		70–99	56–101
Passage accuracy (DIBELS)	Words correct per minute	n/a	79.6 (30.8)	61.5 (48.4)
	% accuracy	100	93.9 (4.0)	79.3 (27.6)
<b>Spelling measures</b>				
PSI words	Raw	26	15.2 (4.0)	15.2 (7.7)
PSI features	Raw	56	46.0 (4.9)	43.3 (10.6)
Affix identification		23	9.1 (4.5)	8.2 (3.9)

TOWRE, Test of Word Reading Efficiency-Second Edition; DIBELS, Dynamic Indicators of Basic Early Literacy Skills; PSI, Primary Spelling Inventory.

*San Diego Quick Assessment* (described below). Although a reading disability could not be confirmed for the remaining 4 participants based on our limited testing, 3 of them had a confirmed diagnosis of learning disability or auditory processing disorder. For all participants, parents were concerned about their child's reading, the majority were concerned about their child's language ( $n = 14$ ), and about half were concerned about their child's math skills ( $n = 11$ ).

## 2.2. Procedures

The study employed a randomized cross-over trial design in which children completed each of two 8-week interventions in random order (see [Figure 1](#)). The study was approved by the Nonmedical Research Ethics Board at The University of Western Ontario, and pre-registered on the Open Science Framework at <https://osf.io/59ew8>. For each participant, a parent virtually signed an online consent form and the child completed an assent form.

Children were seen once per week in 10, individual, virtual (Zoom), 1.25 h, weekly sessions for each of 2 study terms, which were separated by 8 weeks. In each term, assessment sessions in which all outcome measures were administered were completed in sessions 1 and 10, and the intervention sessions were conducted from weeks 2 to 9. Sessions were conducted by 1 or 2 speech-language pathology students completing a graduate course in school age child language disorders. Children were seen by the same speech-language pathology students for the first study term and then were paired with different students for the second study term. Approximately 10 min of each session was observed by one of the authors (either LA or CD) or a doctoral student working under the supervision of the first author. The observation schedule was rotated such that one of the authors (either LA or CD) observed at least three sessions for each child. The observer met with the graduate students after the session to review and plan for subsequent sessions. A new student team was assigned for each assessment session (i.e., weeks 1 and 10 of each term), and the team was blinded to the intervention programs to which the child was assigned during the study. For the week 1 assessment session, the student team was not told the child's intervention program until after the assessment was completed. As well, another graduate student not otherwise involved with the child observed two intervention sessions to complete a fidelity check by tracking that all activities were completed as per the session lesson plan. All graduate students completed a 2-h training session in administering the assessment materials at the start of the study, and a 2-h training session in the relevant intervention program at the start of each term. Additional session and term end activities including

participant engagement ratings and parent observations are described in [Supplementary materials 2, 3](#).

## 2.3. Assessment/outcome measures

The child's spoken responses were collected in the virtual session for all outcome measures except for (1) the affix identification task for which the child described their response and the session leader underlined the indicated letters on the screen, and (2) the spelling task for which the child wrote down the words on a sheet of paper, the parent took a picture of the paper and emailed it to the session leader. Standard or scaled scores based on published test norms were calculated for available tests for descriptive purposes, however they must be interpreted with caution given that the tests were administered in a virtual environment in the present study.

### 2.3.1. Oral language measures

#### 2.3.1.1. Recalling sentences

For the *Recalling Sentences* subtest of the *Clinical Evaluation of Language Fundamentals, 5th edition* (CELF-5; [Wiig et al., 2013](#)), the child was asked to repeat sentences immediately after hearing them until a cut off of 4 consecutive error or no response scores were reached. Items were scored on a 4-point scale as correct (3), or having 1 (2), 2 to 3 errors (1), or 4 or more errors (0), and scaled scores calculated.

### 2.3.2. Word recognition measures

#### 2.3.2.1. Phonemic awareness

The *phonemic awareness probe* was developed for the study and had 7 sets of 4 trials, with the exception of the first set, which had 5 trials. Each set started with 1 practice item, for which a correct response was provided if the child's response was incorrect. The 7 sets corresponded to asking the child (1) to identify all sounds in a 3–4 sound word (e.g., for the word “shop,” the child would identify the three sounds, sh—o—p), (2) to identify the middle sound (always a vowel) in a 3 sound word (e.g., “oo” in boot), (3) to delete the initial or (4) final sound in a 3–4 sound word (e.g., say “late,” say it again but do not say “l” → ate), (5) to substitute a sound for the initial, (6) final, or (7) middle sound of a 3–4 sound word (e.g., say cut, now change “t” to “f” → cuff). All items were administered. Each item was scored as correct or incorrect for a maximum score of 29.

#### 2.3.2.2. Single word reading accuracy

Grade level single word reading accuracy was assessed using the *San Diego Quick Assessment* ([La Pray and Ross, 1969](#)). The child was



FIGURE 1  
Flow chart outlining the study protocol.

asked to read a list of 10 words at two to three sets below their grade level, and continued on to additional lists until the child made three or more errors within a list. Grade level was determined as the highest grade level list in which the child read eight or more words correctly.

### 2.3.3. Reading fluency measures

#### 2.3.3.1. Word and nonword reading

The *Test of Word Reading Efficiency, 2nd edition* (TOWRE-2; Torgesen et al., 2012) was administered. In the *Sight Word Efficiency* (SWE) subtest, children read as many printed words as possible in 45 s. In the *Phonemic Decoding Efficiency* (PDE) subtest, children read as many pronounceable printed nonwords as possible within 45 s. The total number of words or nonwords read correctly was counted for each subtest, and standardized scores calculated.

#### 2.3.3.2. Passage reading

Each child read a middle benchmark passage from the Oral Reading Fluency subtest of the *Dynamic Indicators of Basic Early Literacy Skills, 8th edition* (DIBELS8; University of Oregon, 2018) corresponding to the child's current school grade. In this test, the child is asked to read from the first word of the passage and to stop after 1 min. Any word to which the child does not respond after 3 s is provided by the examiner and scored incorrect, and the child is asked to continue reading. The number of words read correctly was counted, and the percent accuracy calculated.

### 2.3.4. Spelling measures

#### 2.3.4.1. Spelling words

The *Primary Spelling Inventory* (PSI) from the *Words Their Way* (Bear, 1996) word study intervention program was completed. The examiner read each word in the list, used it in a sentence, and then repeated the word, and the child was asked to write the word down. All 16 words were administered. The word list was designed to test 56 spelling features including 7 each of various initial and final consonants, consonant blends, short vowels, digraphs, common long vowels, other vowel spellings, and inflected endings. The total number of words and features spelled correct were counted.

#### 2.3.4.2. Identifying affixes

The *identifying affix test* was developed for the study. A list of 23 words with and without affixes was shown to the child, and the child was asked to identify any prefixes or suffixes they saw in the words. There were a total 22 affixes in the list, and the number of affixes correctly identified was counted.

## 2.4. Intervention programs

The Leap to Literacy and Wise Words program were each 8-week programs involving weekly 1.25 h sessions. Core program elements are compared in [Supplementary material 1](#). In its original version, the Leap to Literacy program provided each participant with access to an online software program aimed at improving literacy called Lexia™, and included an additional 15 min per session in which the child and tutor worked together on Lexia activities. Our partners from the

Learning Disabilities Association wanted to retain access to Lexia as part of their program. A decision was made to provide all study participants with access to Lexia, but to discontinue the 15 min of shared Lexia work during the session due to concerns regarding internet connectivity for all participants for all sessions. Instead, all study intervention sessions ended with a summary, reflection and check in with both the child and parent on how much time had been spent on Lexia in the past week, what activities had been completed, and if any help was needed. Children had access to their Lexia program at home and parents were encouraged to complete this work with their child. On two occasions, a staff member from the Learning Disabilities Association (author DR who was not otherwise involved with the intervention) contacted parents to answer questions and provide additional practice activities available through the Lexia program.

### 2.4.1. Leap to literacy

Leap to Literacy is a skills-based reading program targeting the five core components of reading instruction including phonological awareness, phonics reading fluency, vocabulary knowledge, and text comprehension. Each session was structured around 4 activities (approximately 15 min each): fluency, comprehension, word study, and phonemic awareness or sight words. For the fluency activity, the grade level at which the child could read with 95%–97% accuracy was determined based on the assessment results. A corresponding grade level passage was selected from the DIBELS8 (University of Oregon, 2018) practice passages (or from other grade level reading passages made available from other sources). From the passage, 4 to 6 challenging words with high utility (i.e., commonly used in other texts) or interest were chosen. During the session, the selected words were reviewed individually, and the child was encouraged to sound out the word, identify parts of the word, underline parts that needed to be “learned by heart” (i.e., memorized due to low letter-sound correspondence), and read the word aloud three times. The child and session leaders worked together to understand the word, relate it to something the child already knew, and use the word in a sentence. After the word review, the child read the passage independently for 1 min, and percent correct words read per minute was calculated. For the remainder of the time, the child and session leader worked together to identify and read tricky words, use expression and pausing during reading, and re-read challenging sentences. Decisions regarding which passage to select for the next session were based on word reading accuracy: (1) if 90% or lower, a passage from a lower grade level was selected, (2) if 91%–97%, the same passage was selected, and (3) if 98%–100% and reading rate and expression were considered improved, a new passage was selected.

For the comprehension activity, a reading passage and question guide was selected from Reading A-Z (Learning A-Z, 2022) or other resources at the child's school grade level. Approximately 4–5 words considered Tier 2 (Beck et al., 2002) were selected from the passage for review. Initially, the passage topic was introduced and discussed in terms of how the topic connected to the child's own experiences (activating background knowledge). Each of the selected Tier 2 words was reviewed individually following a robust vocabulary instructional approach (Beck et al., 2002): print the word, say the word and have student repeat, define the word using a student-friendly definition and use in a sentence, ask student to use the word in a sentence, ask the

student to repeat the word. Then, the session leader read the passage aloud while the child followed along in the text. During and after the reading, the session leader would ask questions about the passage and facilitate comprehension by encouraging visualization, re-reading parts of the passage, highlighting key parts of the text, repeating questions, allowing ample time for response, asking follow-up questions, and providing a model answer. Students were encouraged to visualize, ask questions, judge their own comprehension, and make connections with information they already knew. A new passage at the same grade level was selected for the following session, and chosen to match the child's interests, if expressed.

The word study activity involved word sorts based on the spelling features with which the child had difficulty in the Primary Spelling Inventory. The first features selected were those for which the child required some review (4–5 correct out of 7 at initial assessment), and then moved on to features for which the child needed specific instructions and practice (less than 4 out of 7 at initial assessment). Word sorts were selected from the Words Their Way program (Bear, 1996) to contrast spelling feature patterns (e.g., short and long vowels). Word sorts were completed by showing column headers for each of the categories to be sorted (e.g., short a vs. long a), having the child read or listen to a word, and then decide which column the word fit under. In some cases, the word sorts focused on morphological features such as verb endings. Words were used in sentences and defined incidentally throughout the activity. If the child achieved less than 80% accuracy in sorting, the same word sort was repeated at the following session. If the child achieved 80%–90% accuracy, a new word sort on the same feature was completed. If the child achieved greater than 90% correct, a new feature was targeted. The previous week's word sort was reviewed before starting the new sort, where applicable.

Either a phonemic awareness or sight word activity was completed as the last activity of the session. Phonemic awareness activities were created to correspond to any skills not mastered (i.e., at least 1 error) on the Phonemic Awareness Probe (e.g., initial sound deletion: say flip, say it again but do not say “f”). Sight word lists were from the Dolch (1948) or Fry (1999) sight word lists corresponding to the child's assessed grade level on the San Diego Quick Assessment. The activity used a turn-taking game format (e.g., Connect 4; tic tac toe) in which the session leader or child took a turn after responding (e.g., answered 1 phonemic awareness question; read 3 sight words).

#### 2.4.2. Wise words

The Wise Words program is a reason-based program that seeks to improve children's reading, spelling, and vocabulary by increasing understanding of the connection between speech sounds, letters, and part and whole word meaning. There is a particular focus on understanding how words that share meaning also share orthographic patterns. Using an adaptation of Bowers and Kirby's (2010) Structured Word Inquiry, the approach focuses on four key questions about a word: (1) What does the word mean? (2) How is it built? (3) What are some morphologically (or etymologically) related words?, and (4) What are the sounds that matter? Although the emphasis of the program is on morphological knowledge, the connection to the “sounds that matter” in each word also targets phonemic awareness and phonics skills. Each session started with a review of concepts from the previous session (approximately 10 min) and progressed through 2 to 4 learning and reinforcement activity

cycles (approximately 40 min). Learning activities involved hypothesis testing and/or discussion of individual and related words often using the 4 questions. Reinforcement activities involved using the new concept as part of turn-taking in a game (e.g., Connect 4, tic tac toe). The session ended with a short reading activity such as a comic containing words using the patterns explored during the session (approximately 10 min). The eight sessions were described in detailed lesson plans beginning with the session objectives, terms used in the session, and a material list, then providing a guide for all session activities in point form with color coding used to denote the following sections: (1) notes for the session leader; (2) activity introduction and set up, (3) carrying out the activity, and (4) reviewing with the student. The intention was for each lesson plan to be completed in one session, and for all participants to complete all 8 lesson plans. The content did not vary across participants, although explanations and words elicited were tailored to the child's interest and understanding. If a full lesson plan was not completed in one session, the next session continued with the same lesson plan starting where the previous session had left off before continuing with the next lesson plan.

The key concepts taught across the 8 lessons were as follows: (1) Etymology and roots—observing connections between words that contribute shared meaning (e.g., superhero, superior, supervise). (2) Introduce 4 questions, bases and compound words—understanding the base of a word is the smallest part of a word that carries the central meaning, and compound words have two bases; learn conventions for writing word equations (e.g., fire + pit → firepit) and calling out word equations (e.g., F-I-R-E plus P-I-T is rewritten as F-I-R-E-P-I-T). (3) Learning tapping, calling out, and writing words, and learning about digraphs and trigraphs—the convention of tapping once for each grapheme or important marker in a word while saying (calling out) or writing a grapheme provides a multimodal encoding strategy to support learning; learning about digraphs and trigraphs supports the child in decoding more complex words. (4) Replaceable-e—learning of 2 rules for why words end in a silent, nonsyllabic-e. (5) Prefixes—learn the concept of a prefix, and study several words with common prefixes using the 4 questions to understand the meaning of each prefix. (6) Suffixes—learn the concept of a suffix, and study several words with common suffixes using the 4 questions to understand the meaning of each suffix; learn about tapping and calling out affixes. (7) Multiple suffixes, combining prefixes and suffixes, tricky suffixes—identify prefixes and suffixes, including multiple suffixes, in a word; tap and call out words, and examine how meanings overlap or differ across words using the 4 questions; examine assimilation of prefixes. (8) Bound and free bases, word webs, and word matrixes—learn about bound and free bases; use a word web to discover the base of a word family, and learn how to create and read a word matrix.

At the end of each intervention session for both intervention programs, the session leader summarized the session activities and engaged the child in a session reflection (i.e., what is something you learned this week). Homework was also reviewed at this time, and a check-in on Lexia use was completed. Afterwards, the session leaders completed a session record noting observations regarding the child's progress and rated the child's engagement on a 3-point scale (low, medium, high). Parents were invited to join the session or the end of session activities.

At the end of each term, parents were invited to fill out a short questionnaire of 4 questions scored on a 5-point Likert scale (1 = none;

5 = lots), and open ended questions asking the parent to describe any areas of improvement. The scaled questions asked parents to rate the following questions: their child’s engagement in the session, level of support their child needed from them during the sessions, how much they thought their child’s reading had changed as a result of the study, and how successful they found the online sessions (see [Supplementary material 3](#)).

### 3. Results

#### 3.1. Attendance, engagement, Lexia, and fidelity checks and responder analysis

Data regarding attendance, engagement, and average number of reported minutes spent using Lexia are available in [Supplementary material 2](#). On average, attendance (7.7 of 8 sessions completed), engagement (2.7 out of 3), fidelity to treatment (95%–100%) were high, and the average number of minutes spent weekly on Lexia varied widely. These factors did not differ across programs or terms.

#### 3.1.1. Intervention effects

[Table 2](#) shows the assessment results for all study times points, term 1 pre- and post-intervention, and term 2 pre- and post-intervention for the LLWW and WWLL groups. We focused our statistical analysis on raw scores only. Although scaled (Recalling Sentences), standard (TOWRE-2), grade level (San Diego), and percentage (DIBELS passage) scores are shown in [Table 2](#), none of these were included in the analyses because the scaled, standardized, and grade level scores were not expected to change during this short intervention, and all of these data are captured in their corresponding raw scores. For each of the oral language (Recalling Sentences), word recognition (Phonemic Awareness Probe), reading fluency (Sight Word Efficiency; Phonemic Decoding Efficiency; Passage Accuracy), and spelling scores (Spelling Words; Spelling Features; Affix Identification), a mixed 2 (order: LLWW; WWLL) by 2 (program: Leaps; Wise Words), by 2 (time: pre; post) ANOVA was completed. [Table 3](#) summarizes these analyses. The main effect of time was significant for all outcome measures,  $F(1, 18) > 8, p < 0.05$ , all cases, and effect sizes ( $\eta_p^2$ , where greater than 0.14 is considered large, [Cohen, 1988](#)) ranged from 0.24 to 0.68, indicating significantly higher scores at post- than pre-intervention with large effects in all cases,

TABLE 2 Descriptive statistics including mean (standard deviation) scores for each intervention group, term, and test time point.

Measure	Score type	Maximum Score	Intervention groups							
			LLWW				WWLL			
			Term 1	Term 1	Term 2	Term 2	Term 1	Term 1	Term 2	Term 2
			LL	LL	WW	WW	WW	WW	LL	LL
			Pre	Post	Pre	Post	Pre	Post	Pre	Post
<b>Oral language measure</b>										
Recalling sentences	Raw	78	59.5 (8.1)	60.5 (7.3)	63.5 (6.2)	62.6 (4.0)	52.5 (8.4)	56.0 (8.4)	54.0 (12.2)	60.7 (8.2)
	Scaled		12.1 (2.7)	12.0 (2.8)	12.8 (2.9)	11.9 (2.4)	10.2 (2.5)	10.8 (2.5)	10.4 (3.9)	12.4 (3.3)
<b>Word recognition measures</b>										
Phonemic awareness	Raw	29	20.2 (4.3)	23.5 (2.5)	23.7 (3.5)	24.3 (3.9)	18.5 (3.9)	22.0 (3.7)	21.6 (3.7)	25.2 (2.6)
San Diego level	Grade	n/a	3.3 (1.6)	3.7 (1.4)	3.4 (2.7)	3.9 (2.5)	2.9 (2.3)	3.0 (2.2)	4.0 (1.5)	4.1 (1.2)
<b>Reading fluency measures</b>										
Sight word efficiency	Raw	108	52.2 (11.5)	56.6 (13.2)	58.6 (12.4)	58.2 (14.7)	45.7 (22.4)	49.2 (22.8)	51.4 (25.4)	53.3 (22.5)
	Standard	n/a	86.3 (10.0)	89.2 (11.8)	89.4 (11.9)	87.5 (11.8)	80.1 (14.2)	82.5 (14.1)	85.4 (16.6)	85.6 (16.1)
Phonemic decoding (nonword) efficiency	Raw	66	21.5 (8.1)	23.7 (8.9)	25.1 (9.7)	27.9 (11.7)	24.2 (12.5)	27.9 (14.5)	25.2 (15.2)	29.6 (16.1)
	Standard	n/a	86.7 (8.9)	88.0 (10.9)	88.2 (11.6)	90.1 (13.4)	88.8 (11.3)	92.9 (14.7)	88.6 (14.4)	94.3 (16.4)
TOWRE total score	Standard	n/a	83.8 (10.2)	86.3 (12.7)	86.7 (13.3)	86.5 (13.8)	81.6 (14.5)	85.2 (16.4)	84.1 (18.6)	88.0 (18.9)
Passage accuracy (DIBELS)	Words correct/min	n/a	79.6 (30.8)	92.4 (40.7)	100.1 (36.3)	112.2 (39.5)	61.5 (48.4)	75.2 (48.8)	76.8 (47.6)	82.0 (50.7)
	%accuracy	100	93.9 (4.0)	95.2 (7.1)	95.9 (4.2)	97.2 (3.4)	79.3 (27.6)	86.4 (13.6)	88.2 (14.5)	90.3 (12.1)
<b>Spelling measures</b>										
PSI words	Raw	26	15.2 (4.0)	17.6 (5.1)	17.2 (4.2)	18.3 (3.5)	15.2 (7.7)	16.5 (8.1)	16.8 (8.2)	18.2 (7.7)
PSI features	Raw	56	46.0 (4.9)	48.0 (5.7)	48.3 (4.3)	49.4 (3.2)	43.3 (10.6)	45.4 (10.3)	45.0 (11.1)	47.4 (8.8)
Affix identification	Raw	22	9.1 (4.5)	10.5 (4.7)	12.3 (3.7)	15.2 (2.4)	8.2 (3.9)	11.3 (2.8)	12.4 (2.5)	14.0 (2.8)

LLWW, group completing Leap to Literacy (LL) and then Wise Words (WW); WWLL, group completing Wise Words and then Leap to Literacy; Reading fluency measures are from the Test of Word Reading Efficiency (TOWRE) except where indicated; PSI, Primary Spelling Inventory.



TABLE 3 Results of ANOVAs ( $F$ ,  $p$ ,  $\eta_p^2$ ) including main (order, time, program) and interaction effects for all outcomes measures.

Measure	Order	Time	Program	Order × program	Order × time	Time × program	Order × program × time
Recalling sentences	3.1, 1.0, 0.15	5.8, 0.03*, 0.24	<0.001, 1.0, <0.0001	9.1, 0.008**, 0.33	5.6, 0.03*, 0.24	3.0, 0.1, 0.14	0.2, 0.7, 0.01
Phonemic awareness probe	0.8, 0.4, 0.04	32.6, <0.001**, 0.64	0.5, 0.5, 0.03	13.6, 0.002**, 0.43	2.8, 0.1, 0.13	1.7, 0.2, 0.08	1.4, 0.2, 0.07
Sight word efficiency (TOWRE)	0.6, 0.4, 0.03	2.2, 0.2, 0.11	0.1, 0.7, 0.01	12.6, 0.002**, 0.41	0.5, 0.8, 0.003	0.5, 0.5, 0.03	2.0, 0.2, 0.01
Phonemic decoding (nonword) efficiency (TOWRE)	0.2, 0.7, 0.01	8.1, 0.011*, 0.31	1.0, 0.3, 0.06	4.4, 0.050, 0.20	0.5, 0.5, 0.03	0.001, 1.0, 0.0001	0.2, 0.7, 0.01
Passage accuracy (DIBELS)	1.5, 0.2, 0.07	11.4, 0.003**, 0.39	1.4, 0.3, 0.07	16.2, <0.001**, 0.47	0.2, 0.7, 0.01	0.4, 0.5, 0.02	0.5, 0.5, 0.03
PSI words	0.02, 0.9, 0.001	37.4, <0.001*, 0.67	0.1, 0.7, 0.01	11.2, 0.004**, 0.38	0.6, 0.4, 0.03	1.0, 0.3, 0.05	0.8, 0.4, 0.04
PSI features	0.6, 0.5, 0.03	29.3, <0.001**, 0.62	<0.001, 1.0, <0.0001	9.1, 0.007**, 0.34	1.0, 0.3, 0.05	0.3, 0.6, 0.02	0.1, 0.8, 0.004
Affix identification	0.08, 0.8, 0.005	8.4, 0.01*, 0.32	0.2, 0.7, 0.009	36.9, <0.001*, 0.67	0.02, 0.9, <0.0001	1.4, 0.3, 0.07	<0.001, 1.0, <0.0001

Degrees of freedom = 1, 18 (all cases); \* $p < 0.05$ , \*\* $p < 0.01$ .

except Sight Word Efficiency,  $F(1, 18) = 2.2$ ,  $p = 0.11$ . There was also a significant interaction between program order and program for all outcomes measures except Phonemic Decoding Efficiency,  $F(1, 18) > 9$ ,  $p < 0.05$ . Effect sizes ( $\eta_p^2$ ) ranged from 0.20 to 0.67, indicating large effects in all cases. The one additional significant effect was the interaction between program order and time for Recalling Sentences,  $F(1, 18) = 5.6$ ,  $p = 0.03$ ,  $\eta_p^2 = 0.24$ . All remaining effects were not significant,  $F(1, 18) < 3.1$ ,  $p > 0.05$ ,  $\eta_p^2 < 0.15$ , all cases.

In order to unpack the interaction between program order and program, pairwise comparisons with Bonferroni corrections were completed, as shown in Table 4. No statistically significant effects survived the Bonferroni correction for Recalling Sentences, Sight Word Efficiency, Spelling Words, and Spelling Features, and so these interactions were not analyzed further. For the Phonemic Awareness Probe, there was one marginally significant and large effect due to higher scores in Leap to Literacy than Wise Words when Leap to Literacy was completed second (WWLL group),  $t = 3.1$ ,  $p = 0.050$ ,  $d = 0.9$ , whereas the programs did not differ when completed in the opposite order (LLWW group),  $t = 2.1$ ,  $p = 0.3$ ,  $d = 0.6$ . This pattern suggests a positive effect of completing Leap to Literacy second on the Phonemic Awareness Probe. For the Passage Accuracy measure, there was a significant and medium effect due to higher scores in Wise Words completed second compared to Leap to Literacy completed first (LLWW group),  $t = 3.7$ ,  $p = 0.01$ ,  $d = 0.5$ , whereas the programs did not differ when completed in the opposite order (WWLL group),  $t = 2.0$ ,  $p = 0.4$ ,  $d = 0.3$ . This pattern suggests a modest Passage Accuracy advantage for completing the Wise Words program second. Finally, in the case of Affix Identification, there were significant and large effects when comparing programs for both the LLWW and WWLL groups favoring the program completed in the second term,  $t > 3.9$ ,  $p < 0.01$ ,  $d > 0.8$ , both cases. Affix Identification was the only variable for which both between and within program comparisons (i.e., LLWW vs. WWLL for Leap to Literacy and LLWW vs. WWLL for Wise Words) showed higher scores in the second term regardless of order with large effects that were significant for Wise Words,  $t = 3.3$ ,  $p = 0.015$ ,  $d = 1.1$ , and marginal for Leap to Literacy,  $t = 2.8$ ,  $p = 0.053$ ,  $d = 1.0$ . It must be noted that despite these significant effects, no significant within-term,

cross-program differences were observed when comparing the LLWW and WWLL groups in term 1,  $t < 1.8$ ,  $p > 0.6$  (all cases), or term 2,  $t < 1.8$ ,  $p > 0.6$  (all cases), with small to vanishing effect sizes ( $d < 0.41$ , all cases) for all but Recalling Sentences ( $d = 0.7$ , both cases).

The remaining significant interaction was between program order and time for the Recalling Sentences outcome measures. In post-hoc pairwise comparisons collapsed across program, the only significant difference was between pre- and post-intervention Recalling Sentences scores for the WWLL group,  $t = 3.4$ ,  $p = 0.02$ ,  $d = 0.6$ , which differed noticeably from the lack of significant difference and very small effect for the same comparison in the LLWW group,  $t = 0.03$ ,  $p = 1.0$ ,  $d = 0.006$ . Nevertheless, the comparisons between pre-intervention time points (LLWW vs. WWLL) and pre-intervention for WWLL and post-intervention for LLWW were not significant but had a large effect size,  $t = 2.4$ ,  $p = 0.1$ ,  $d = 1.0$ , both cases, suggesting that the lower pre-intervention scores of the WWLL group contributed to this pattern. All remaining comparisons were not significant and had small effects,  $t < 1.0$ ,  $p > 0.7$ ,  $d = 0.4$ .

## 4. Discussion

This study compared two community-based reading intervention programs across a group of individuals with learning disabilities including reading challenges. In a randomized cross-over trial, one group received a skills-based program, *Leap to Literacy*, focusing on explicit instruction and discrete skill practice, and the other received a reason-based program, *Wise Words*, focusing on discussion, critical thinking, and word meaning relationships (morphological awareness). Regardless of program order or program, positive changes from pre- to post-intervention were observed for word recognition (Phonemic Awareness Probe), reading fluency (Phonemic Decoding Efficiency; Passage Accuracy), and spelling scores (Spelling words; Spelling features; Affix identification). Only one measure, Sight Word Reading Efficiency did not show a study-wide score increase. A change from

TABLE 4 Mean (standard deviation) scores for each program within each term and order group, and results of pairwise comparisons with Bonferroni correction (t, p, Cohen's d).

Group/comparison → measure↓	Leap to literacy				Wise words				Pairwise comparisons					
	Term 1		Term 2		Term 1		Term 2		LLWW-LLWW		WWLL-WWLL		WWLL-LLWW	
	LLWW	WWLL	LLWW	WWLL	WWLL	LLWW	LLWW	LLWW	LLWW-LLWW	WWLL-WWLL	WWLL-LLWW	WWLL-LLWW	WWLL-LLWW	
	L1	L2	W1	W2	L1-L2	L1-W1	L1-W2	L2-W1	L2-W2	W1-W2				
Recalling sentences	60.0 (7.5)	57.4 (10.7)	54.2 (8.4)	63.0 (5.1)	0.8, 1.0, 0.3	1.7, 0.6, 0.7	-2.1, 0.3, 0.4	2.1, 0.3, 0.4	-1.7, 0.6, 0.7	-2.6, 0.1, 1.1				
Phonemic awareness probe	21.8 (3.8)	23.4 (3.6)	20.2 (4.1)	24.0 (3.6)	-1.1, 1.0, 0.4	1.1, 1.0, 0.4	-2.1, 0.3, 0.6	3.1, 0.05*, 0.9	-0.4, 1.0, 0.2	-2.6, 0.08, 1.1				
Sight word efficiency (TOWRE)	54.4 (12.3)	52.3 (23.4)	47.4 (22.0)	58.4 (13.2)	0.2, 1.0, 0.1	0.8, 1.0, 0.4	-2.3, 0.2, 0.2	2.8, 0.08, 0.3	-0.7, 1.0, 0.3	-1.3, 1.0, 0.6				
Passage accuracy (DIBELS)	86.0 (35.7)	79.4 (47.9)	68.3 (47.9)	106.0 (38.4)	0.3, 1.0, 0.2	0.9, 1.0, 0.4	-3.7, 0.01*, 0.5	2.0, 0.4, 0.3	-1.4, 1.0, 0.6	-2.0, 0.4, 0.9				
Spelling words	16.4 (4.6)	17.5 (7.8)	15.8 (7.1)	17.7 (3.9)	-0.4, 1.0, 0.2	0.2, 1.0, 0.1	-2.1, 0.3, 0.2	2.6, 0.1, 0.3	-0.09, 1.0, 0.04	-0.7, 1.0, 0.3				
Spelling features	47.0 (5.3)	46.2 (9.8)	44.3 (10.3)	48.9 (3.9)	0.2, 1.0, 0.1	0.8, 1.0, 0.3	-2.1, 0.3, 0.2	2.1, 0.3, 0.2	-0.8, 1.0, 0.3	-1.3, 1.0, 0.6				
Affix identification	9.8 (4.5)	13.2 (2.7)	9.7 (3.7)	13.7 (3.4)	-2.8, 0.053, 1.0	0.04, 1.0, 0.01	-4.6, 0.001**, 0.9	4.0, 0.005**, 1.0	-0.5, 1.0, 0.2	-3.3, 0.015*, 1.1				

LLWW, group completing Leap to Literacy then Wise Words; WWLL, group completing Wise Words then Leap to Literacy; L1, L2, indicates Leap to Literacy completed in terms 1 or 2; W1, W2, indicates Wise Words completed in terms 1 or 2; \*p < 0.05; \*\*p < 0.01.

pre- to post-intervention was modified by an interaction with program order for the oral language measure, Recalling Sentences. Specifically, only the group completing Wise Words first showed a significant change in Recalling Sentences, although they also started the program with the lowest scores on this measure. We observed an interaction between program order and program for nearly all variables, however, these effects were robust (significant and large) for only three variables. For the Phonemic Awareness Probe, the effect of Leap to Literacy was larger than Wise Words when completed in the second but not first term, whereas for the Passage Accuracy measure, the effect of Wise Words was larger than Leap to Literacy when completed in the second but not first term. For Affix Identification, completing both programs regardless of order resulted in large effects after the second term.

Overall, we found significant improvements on various reading and spelling measures in 8- to 13-year-old children (grades 4 to 8) who participated in two, 8-week, 1.25-h weekly, community-based, virtual, reading intervention programs. When completing both our skills-based, Leap to Literacy, and reason-based, Wise Words programs, regardless of order and program, children improved on measures of phonemic awareness, nonword reading, reading words in a grade-level passage, spelling words, spelling patterns and identifying affixes. Generally speaking, these measures tap phonics, orthographic mapping, and morphological knowledge, all of which were directly targeted in our intervention programs. Word reading fluency showed some improvement with higher accuracy on a grade-level passage measure, but not in the number of words read accurately on our sight word efficiency measure. The overall effects of time in intervention regardless of program completed is to be expected given the considerable overlap in the core elements of our reading programs (see [Supplementary material 1](#)).

In total, our middle school participants spent 20 h in intervention across the two intervention programs, and group level effect sizes expressed as partial eta-squared were large in all cases. These results compare favorably to the small but nontrivial effects reported in previous meta-analyses of out-of-school programs (Cooper et al., 1996) and for which effects for programs of less than 44 h were not found (Lauer et al., 2006). In their meta-analysis, Lauer et al. (2006) did not observe positive effects for upper elementary students, unlike the present study. Our programs incorporated two factors found to moderate intervention outcomes in the Lauer et al. meta-analysis, namely the positive effects of programs focusing primarily on academic skills and working with students one-on-one.

Nevertheless, we did observe one case in which the overall effect of intervention time was modified by program order. A significant change in recalling sentences was found when Wise Words was completed before Leap to Literacy, but not in the opposite order. Recalling sentences is an oral language measure (Klem et al., 2015). Why might completing Wise Words first positively influence oral language? The Wise Words program fostered ample opportunity for discussion and critical thinking regarding words and in complex one-to-one conversations with expert language users (i.e., the session leaders). It may be that the language facility gained in Wise Words when completed first was carried over into Leap to Literacy sessions and further fostered resulting in the observed order effect for recalling sentences. There are other possible explanations, however. The improved recalling sentences scores could be related to a practice

effect given that participants completed the task 4 times over 24 weeks, however, a practice effect would not explain improvements for only one group as observed for this order effect. A more probable possibility is that the order effect resulted from the lower term 1 (i.e., baseline) recalling sentences scores of the group completing Wise Words first. Although not significantly different from the Leap to Literacy first group, the score difference at baseline was associated with a large effect with the Wise Words first group scoring lower. It may be that the Wise Words first group had more “room to grow” in their scores given their lower start.

Another interesting finding in the current study was fairly consistent interactions (across 7 of 8 measures) between intervention program and the order in which the programs were completed. These interactions are somewhat challenging to interpret given the lack of significant differences between groups completing the programs in opposite orders observed for the majority of measures across the study. That is, there was no evidence that either program order resulted in significantly different performance at the end of the study. Nevertheless, understanding the observed order effects on each of the intervention programs could point to future directions for further study. We took a conservative approach and analyzed three of the program by program order interactions for which significant and large effects were found after Bonferroni correction. In the case of phonemic awareness, Leap to Literacy scores were significantly higher than Wise Words only when Leap to Literacy was completed second. It may be that the morphological knowledge gained during Wise Words primed participants for improving phonemic awareness skills during Leap to Literacy. We know that morphological and phonological knowledge are closely related (Berninger et al., 2010), and so a carryover effect of morphological to phonological knowledge could be predicted. We also know that from about grade 4, morphological knowledge predicts more variability in reading performance than phonological knowledge (Singson et al., 2000; Kirby et al., 2012). Perhaps the focus on morphological awareness in Wise Words provided the older children in this study with the knowledge needed to further develop their phonemic awareness skills.

We also observed greater passage reading accuracy for the Wise Words compared to Leap to Literacy program when Wise Words was completed second but not first. One possible explanation for this could relate to the emphasis on the child reading aloud in the Leap to Literacy program, a strategy found to be important in repeated reading activities (Therrien, 2004). It may be that this focus in the skills-based Leap to Literacy program resulted in improved reading facilitating greater engagement in the reading activities in Wise Words. Unfortunately, we did not capture any session level data that would allow us to investigate this hypothesis. In Wise Words, children could choose to complete the passage reading activities as they wished, that is, they could ask for it to be read to them, they could take turns reading parts with the session leader, or they could do all the reading themselves. It is possible that after completing Leap to Literacy which involved more reading aloud by the child, they took on more independent reading in the Wise Words program. Future studies would need to record the amount of reading during the sessions in order to evaluate this hypothesis.

Finally, the ability to identify affixes showed greater gains after the second term of intervention regardless of intervention order. This finding was somewhat surprising given that teaching regarding affixes was a particular emphasis of the Wise Words program. If the effect was

driven by the Wise Words program, then we would expect program differences within terms, at least for the first term when only one group had completed the Wise Words program. However, there were no significant differences in identifying affixes at the end of the first term when comparing those who had completed Leap to Literacy vs. Wise Words. In fact, completing both programs always resulted in significantly higher scores when compared to having completed only one program, regardless of the program completed. The finding could also reflect a practice effect, however, there would be no reason to expect a practice effect on only this study measure.

Taken together, the findings of program order effects in the present study allow for the possibility of additive effects of skills- and reason-based community reading programs and/or particular benefits of one approach for different learners. There were indications of potential benefits of both program orders. First completing the reason-based Wise Words program with its emphasis on discussion, critical thinking, and word meaning relationships might have fostered more conversation and greater success with phonemic awareness activities in the skills-based Leap to Literacy program resulting in improved oral language (recalling sentences) and phonemic awareness scores. On the other hand, completing the Leap to Literacy program first with its emphasis on word and text reading fluency may have resulted in more engagement with the reading activities in the Wise Words program leading to better passage reading accuracy. It would be interesting to examine how potential benefits of early discussion, critical thinking, word meaning examination, and word reading fluency could be incorporated in a single, merged community reading program. Nevertheless, it may be that some learners are particularly responsive to either skills- or reason-based approaches to reading instruction. Further research is required to investigate possible additive effects of program orders or the profiles of children who respond best to a skills- or reason-based approach to literacy instruction.

One of the marked limitations of this study was the lack of comparison groups who had completed two terms of Leap to Literacy or two terms of Wise Words. For all of the order effects we observed, we do not know if two terms of the same program would have resulted in the same or even stronger order effects. For example, phonemic awareness scores were improved when Leap to Literacy was completed second. It may be that phonemic awareness scores would have been even more improved had two terms of Leap to Literacy been completed rather than one term of Wise Words and one term of Leap to Literacy. Similarly, for passage reading accuracy where completing Wise Words second resulted in higher scores, we do not know if it was crucial for Leap to Literacy to be completed prior to Wise Words or if two terms of Wise Words would have been more beneficial. In addition, a second term of Wise Words would have provided knowledge of many additional concepts related to why words are spelled the way they are, and how adding affixes impacts spelling. We do not know if completing two terms of Wise Words would have been particularly beneficial. Nevertheless, the findings of order effects in this small study do motivate a larger study with all possible program orders included in order to understand how and whether these two programs complement one another. Such a study would also assist in the evaluation of individual participant responses to each program.

Another limitation is the lower scores of the Wise Words first group at time point 1. Our randomization process resulted in no

statistical differences between our program order groups (Leap to Literacy first vs. Wise Words first) at the initial testing point. Although these results provide evidence of baseline equivalence, the Wise Words first group had somewhat lower means and/or higher standard deviations for all measures. There was a large effect difference at baseline for two of the outcome variables for which significant intervention effects were observed in the present study, namely, Recalling Sentences and Passage Accuracy. This pattern makes it challenging to interpret the importance of the findings regarding these measures. There was one case in which the Wise Words first group had higher results at baseline, and that was in the number of minutes spent on home reading-related activities. Given that the Wise Words first group did not systematically perform more strongly in the study, it is unlikely that this trend conveyed any intervention-relevant advantage. A larger randomized trial would not have the same vulnerability to baseline trends, and so would shed light on these findings.

A number of additional limitations relate to the small sample size coupled with the use of several different measures, limited training of intervention approaches (i.e., 2 h), the large number of session leaders, limited supervision of graduate students administering the intervention, and the lack of blinding regarding the program the child had completed in the first term. It was also difficult to assess the potential impact of Lexia on intervention findings as data collected regarding Lexia use was limited and inconsistently reported. It should also be noted that Lexia is better aligned with Leaps to Literacy compared to Wise Words and as a result these participants could have potentially received more exposure to skill-based interventions. As well, it was necessary for session leaders to complete the initial assessment each term in order to plan sessions, especially Leap to Literacy sessions, based on the assessment results. As a result, however, the session leaders were not blinded to the reading status of the participant. We also had many different session leaders involved in the study. Although we demonstrated high fidelity to our treatment plans, the inherent variability across session leaders likely contributed noise to the data. Finally, the affix measure used was designed for this study and as a result, psychometric data is not available. Future work may consider the development of this tool's psychometric properties.

This randomized, cross-over study compared two 8-week, 1.25-h weekly, community-based reading intervention programs for 8- to-13-year-old children. Leap to Literacy focused on skill learning through explicit instruction, and Wise Words emphasized discussion and critical thinking to understand how a word's morphological structure is connected to how it is spelled. Regardless of program and program order, improved scores were observed for measures of phonemic awareness, nonword reading, passage reading, spelling words, spelling patterns, and identifying affixes. These improvements in skills related to phonemic awareness, phonics, and orthographic mapping regardless of program are explained by the overlap of core elements across the programs. One oral language measure, the ability to recall sentences, improved only when Wise Words was completed first, suggesting a carry forward benefit of the emphasis on discussion in this reason-based approach. Although not modifying the overall results of the interventions, there were indications of an order effect in at least three cases: (1) a benefit for phonemic awareness when Leap to Literacy was completed after Wise Words, (2) a passage

reading accuracy advantage to completing Leap to Literacy before Wise Words, and (3) a greater increase in affix identification in the second term of the intervention regardless of program. These order effects might suggest a potential additive effect of combining reason- and skills-based approaches to intervention. Future research is needed to examine additive effects of a short-term, community-based skills-based program alone, reason-based program alone, or a combination. The results do, however, demonstrate the positive impact of community-based reading intervention programs delivered virtually.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by Western University's Research Ethics Board WesternREM (WREM). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## Author contributions

LA led the study design, data collection, and manuscript preparation. CD, AK, and SD supported data collection and manuscript preparation. CK, DS-R, and PC support study design and manuscript preparation. MV supported manuscript preparation. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by a Research Western Award given to LA (grant number: 48361).

## Acknowledgments

The authors are grateful to the many people who made this study possible. Danielle Ferguson was the lead writer and Jonathan Spence contributed significantly to the lesson plan design of the Wise Words program used in the present study. The materials for the virtual Wise Words sessions were created by SD. As well, Marina Bishay, Rachel Dube, Emily Patrick, and Theresa Pham assisted with the writing of previous versions of the Wise Words program that were used in smaller pilot studies. Kathy Clark contributed significantly to the design of the Leap to Literacy program. We are grateful to the many graduate students in Dr. Archibald's Developmental Language Disorder 2 (#WesternDLD2) class who acted as session leaders for the present and pilot studies, Dr. Archibald's doctoral students who acted as mentors (Meghan

Vollebregt, Taylor Bardell, Theresa Pham, and AK), and the staff at both the London Disabilities Association—London Region and the Child and Youth Developmental Clinic who assisted with the administration of the intervention programs. To the students and families who participated, we thank you for your willingness and engagement in this study. Thanks, also, to Deanne Friesen for helpful comments on an earlier version of this manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Afacan, K., Wilkerson, K. L., and Ruppert, A. (2017). Multicomponent reading interventions for students with intellectual disability. *Remedial Spec. Educ.* 39, 229–242. doi: 10.1177/0741932517702444
- Bear, D. R. (1996). *Words their way*. Iowa State University & University of Nevada, Reno.
- Beck, I. L., McKeown, M. G., and Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*, 2nd Edn. New York, NY: Guilford Press.
- Berninger, V. W., Abbott, R. D., Nagy, W., and Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *J. Psycholinguist. Res.* 39, 141–163. doi: 10.1007/s10936-009-9130-6
- Bowers, P. N., and Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Read Writ* 23, 515–537. doi: 10.1007/s11145-009-9172-z
- Bowers, P. N., Kirby, J. R., and Deacon, H. (2010). The effects of morphological instruction on literacy skills: a systematic review of the literature. *Rev. Educ. Res.* 80, 144–179. doi: 10.3102/0034654309359353
- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: impact on reading. *Read. Writ. Interdiscip. J.* 12, 169–190. doi: 10.1023/A:1008131926604
- Carlisle, J. F. (2007). “Fostering morphological processing, vocabulary development, and reading comprehension” in *Vocabulary acquisition: Implications for reading comprehension*. eds. R. K. Wagner, A. E. Muse and K. R. Tannenbaum (NY: Guilford Press), 78–103.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum.
- Cooper, H., Nye, B., Charlton, K., Lindsay, J., and Greathouse, S. (1996). The effects of summer vacation on achievement test scores: a narrative and Meta-analytic review. *Rev. Educ. Res.* 66, 227–268. doi: 10.3102/00346543066003227
- Deacon, H., Kirby, J. R., and Casselman-Bell, M. (2009). How robust is the contribution of morphological awareness to general spelling outcomes? *Read. Psychol.* 30, 301–318. doi: 10.1080/02702710802412057
- Devonshire, V., and Fluck, M. (2010). Spelling development: fine-tuning strategy-use and capitalising on the connections between words. *Learn. Instr.* 20, 361–371. doi: 10.1016/j.learninstruc.2009.02.025
- Devonshire, V., Morris, P., and Fluck, M. (2013). Spelling and reading development: the effect of teaching children multiple levels of representation in their orthography. *Learn. Instr.* 25, 85–94. doi: 10.1016/j.learninstruc.2012.11.007
- Dolch, E. W. (1948). *Problems in reading*. Champaign, IL: The Garrard Press.
- Donegan, R. E., and Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: a look at recent research. *Read Writ.* 34, 1943–1947. doi: 10.1007/s11145-021-10123-y
- Fry, E. B. (1999). *1000 instant words*. Westminister, CA: Teacher Created Materials.
- Goodwin, A. P., and Ahn, S. (2010). A meta-analysis of morphological interventions: effects on literacy achievement of children with literacy difficulties. *Ann. Dyslexia* 60, 183–208. doi: 10.1007/s11881-010-0041-x
- Goodwin, A. P., and Ahn, S. (2013). A meta-analysis of morphological interventions in English: effects on literacy outcomes for school-age children. *Sci. Stud. Read.* 17, 257–285. doi: 10.1080/10888438.2012.689791
- Kim, J. S., Hemphill, L., Troyer, M., Thomson, J. M., Jones, S. M., LaRusso, M. D., et al. (2017). Engaging struggling adolescent readers to improve reading skills. *Read. Res. Q.* 52, 357–382. doi: 10.1002/trq.171
- Kirby, J. R., Deacon, S. H., Powers, P. N., Izenberg, L., Wade-Woolley, L., and Parrila, R. (2012). Children’s morphological awareness and reading ability. *Read. Writ.* 25, 389–410. doi: 10.1007/s11145-010-9276-5
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S.-A. H., Gustafsson, J., and Hulme, C. (2015). Sentence repetition is a measure of children’s language skills rather than working memory limitations. *Dev. Sci.* 18, 146–154. doi: 10.1111/desc.12202
- La Pray, M., and Ross, R. R. (1969). San Diego Quick Assessment. *J. Read.* 12, 305–307.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., and Martin-Glenn, M. L. (2006). Out-of-school-time programs: a meta-analysis of effects for at-risk students. *Rev. Educ. Res.* 76, 275–313. doi: 10.3102/00346543076002275
- Learning A-Z (2022). *ReadingA-z.com*. Dallas, TX, United States: Lazel, Inc.
- Murphy, K. A., and Diehm, E. A. (2020). Collecting words: a clinical example of a morphology-focused orthographic intervention. *Lang. Speech Hear. Serv. Sch.* 51, 544–560. doi: 10.1044/2020\_LSHSS-19-00050
- National Reading Panel Report (NICHD). (2000). *Report of the National Reading Panel: Teaching children to read: Reports of the subgroups (00–4754)*. Washington, DC: U.S. Government Printing Office.
- Nunes, T., Bryant, P., and Olsson, J. (2003). Learning morphological and phonological spelling rules: an intervention study. *Sci. Stud. Read.* 7, 289–307. doi: 10.1207/S1532799XSSR0703\_6
- Ontario Human Rights Commission (2022). *Right to read: Public inquiry into human rights issues affecting students with reading disabilities*. Toronto, Canada: Government of Ontario.
- Reed, D. K. (2008). A synthesis of morphology interventions and effects on reading outcomes for students in grades K-12. *Learn. Disabil. Res. Pract.* 23, 36–49. doi: 10.1111/j.1540-5826.2007.00261.x
- Ritter, G. W., Barnett, J. H., Denny, G. S., and Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: a meta-analysis. *Rev. Educ. Res.* 79, 3–38. doi: 10.3102/0034654308325690
- Schechter, R., Macaruso, P., Kazakoff, E. R., and Brooke, E. (2015). Exploration of a blended learning approach to reading instruction for low SES students in early elementary grades. *Comput. Sch.* 32, 183–200. doi: 10.1080/07380569.2015.1100652
- Singson, M., Mahony, D., and Mann, V. (2000). The relation between reading ability and morphological skills: evidence from derivational suffixes. *Read. Writ. Interdiscip. J.* 12, 219–252. doi: 10.1023/A:1008196330239
- Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial Spec. Educ.* 25, 252–61.
- Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (2012). *Test of word Reading efficiency*, 2nd Edn. Dallas, TX: Pro-ed, Inc.
- University of Oregon (2018). *Dynamic indicators of basic early literacy skills*. 8th Edn. Oregon, USA: College of Education, University of Oregon.
- Wiig, E. H., Semel, E., and Secord, W. A. (2013). *Clinical evaluation of language fundamentals*, 5th Edn. Bloomington, MN: NCS Pearson.
- Wilkes, S., Macaruso, P., Kazakoff, E., and Albert, J. (2016). Exploration of a blended learning approach to Reading instruction in second grade. In *Proceedings of EdMedia: World Conference on Educational Media and Technology 2016* (pp. 791–796). Association for the Advancement of Computing in Education (AACE).

## Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2023.1104945/full#supplementary-material>