



Cognitive and Linguistic Effects of Working Memory Training in Children With Corresponding Deficits

Laura J. Pauls and Lisa M. D. Archibald*

The University of Western Ontario, School of Communication Sciences and Disorders, London, ON, Canada

Working memory training has been found to result in improvements on tasks similar to those involved in the training (near transfer) but show limited impact on everyday skills such as language or academic abilities (far transfer). Previous research has largely focused on those with a broad range of skills, and examined group level responses. This study uses a single subject design to examine the effect of working memory training on both working memory and related domains in children with working memory impairment. Seven children (8–11 years old) with a working memory impairment completed 20 sessions of computerized working memory intervention. Data revealed near transfer for all participants. Evidence of far transfer to improvements in language, reading, or math was observed for approximately half of the participants on individual measures. Three participants showed convincing but modest training effects across multiple measures. A combination of factors appeared to influence far transfer including age, training intensity, and baseline measures.

OPEN ACCESS

Edited by:

Susana Castro-Kemp,
University of Roehampton London,
United Kingdom

Reviewed by:

Angelliki Mouzaki,
University of Crete, Greece
Vanessa Arán Filippetti,
Consejo Nacional de Investigaciones
Científicas y Técnicas (CONICET),
Argentina

*Correspondence:

Lisa M. D. Archibald
larchiba@uwo.ca

Specialty section:

This article was submitted to
Special Educational Needs,
a section of the journal
Frontiers in Education

Received: 10 November 2021

Accepted: 28 December 2021

Published: 28 January 2022

Citation:

Pauls LJ and Archibald LMD (2022)
Cognitive and Linguistic Effects of
Working Memory Training in Children
With Corresponding Deficits.
Front. Educ. 6:812760.
doi: 10.3389/feduc.2021.812760

Keywords: working memory impairment, working memory training, language, single subject design, near transfer, far transfer

1 INTRODUCTION

Working memory training results in improvements on tasks bearing a strong resemblance to those employed in training, but effects on related abilities, such as language, reading, and math, are unreliable (Melby-Lervåg et al., 2016). Two possible reasons for these inconsistent findings are the complexities of cognitive impairments and the limitations of the large-group study designs that make up the majority of studies in this literature. Children with cognitive impairments present with a variety of profiles that may be lost in large group comparisons. The present study addresses these limitations by testing working memory training on children with working memory impairments in a single-subject design.

Working Memory, a capacity limited cognitive resource for retaining task-relevant information in a highly accessible state, has been found to be related to language abilities (see Baddeley, 2003 for review). One particularly robust finding is the association of verbal short term memory with the acquisition of new vocabulary (Gathercole, 2006) with converging evidence of poor verbal short term memory ability among children with language impairment (Archibald and Gathercole, 2006; Graf Estes et al., 2007). Other research has demonstrated an association between verbal working memory ability and comprehension of complex syntax (Montgomery et al., 2008; Roberts and Gibson, 2002), suggesting that processing complex language relies on adequate working memory support. Recent theories regarding sentence comprehension suggest that a subset of language knowledge is held in working memory during processing (Montgomery et al., 2021; Majerus, 2013). Strong associations

have also been reported between working memory and academic abilities in reading and math (Gathercole et al., 2004; Nouwens et al., 2017).

Deficits in working memory have been noted in a number of populations, including children with Attention Deficit Hyperactivity Disorder (ADHD; e.g., Martinussen et al., 2005), and those with academic difficulties (Swanson and Jerman, 2006; Dawes et al., 2015). Nevertheless, the presentation of working memory impairment can be difficult to pinpoint due to the concomitant impairments in many populations and the close associations between working memory and other cognitive abilities.

Working memory training has been shown to improve performance on working memory tasks that are similar or identical to those used in the training program (Melby-Lervåg et al., 2016), otherwise known as near transfer. A major area of interest surrounding working memory training is far transfer, that is, the degree to which training gains in working memory can transfer to other tasks. Unfortunately, far transfer has been difficult to demonstrate (Melby-Lervåg et al., 2016). Even for studies involving children with measured working memory impairment, far transfer has been limited (Alloway et al., 2013; Dunning et al., 2013; Holmes et al., 2015). Similarly mixed findings have been reported for reading (Egeland et al., 2013; Holmes and Gathercole, 2014) and math gains (Holmes et al., 2009; Dahlin, 2013). Taken together, existing research offers no convincing evidence of reliable far transfer effects of working memory training among children with confirmed or possible low working memory.

One limitation of the previous research on working memory training is the over reliance on group-based studies, which ignore change at the level of the individual. Those with working memory impairments are a heterogeneous group, with all members unlikely to respond to intervention in the same way. These issues were addressed in the present study by employing a single-subject design. Such a design is ideal for studying heterogeneous populations because intervention can be tailored to individual abilities, and change is measured at the individual (rather than group) level, allowing for further investigation of participant characteristics that may influence intervention effects (Barlow & Hersen, 1973; McReynolds and Thompson, 1986). Importantly, single-subject designs offer sufficient design strength to establish a causal relationship between the intervention in question and the outcome, even with only a few subjects (Bordens and Abbott, 2011; Horner et al., 2012; Perdices and Tate, 2009).

The purpose of this study was to test the effectiveness of a working memory training program for individual children with working memory impairment. Of particular interest were the near and far transfer effects of working memory training to tasks tapping working memory or other skills (language, reading, math). To address these questions, a working memory training program, Cogmed™ (Klingberg et al., 2005), was offered to children with specific impairment in working memory and children with impairments in both language and working memory. Effects of the training on both near and far transfer tasks were measured using probes, which were collected

throughout baseline, intervention, and follow-up phases, and standardized measures, which were administered before, immediately after, and 3 months after the training was completed. Improvement on working memory measures only would be consistent with near transfer, whereas improvement on measures of language, reading, or math would be indicative of far transfer. In order to examine possible factors affecting transfer, response to the working memory training was evaluated relative to participant characteristics, including baseline abilities in working memory, language, math, and improvement on training tasks.

2 MATERIALS AND METHODS

2.1 Participants

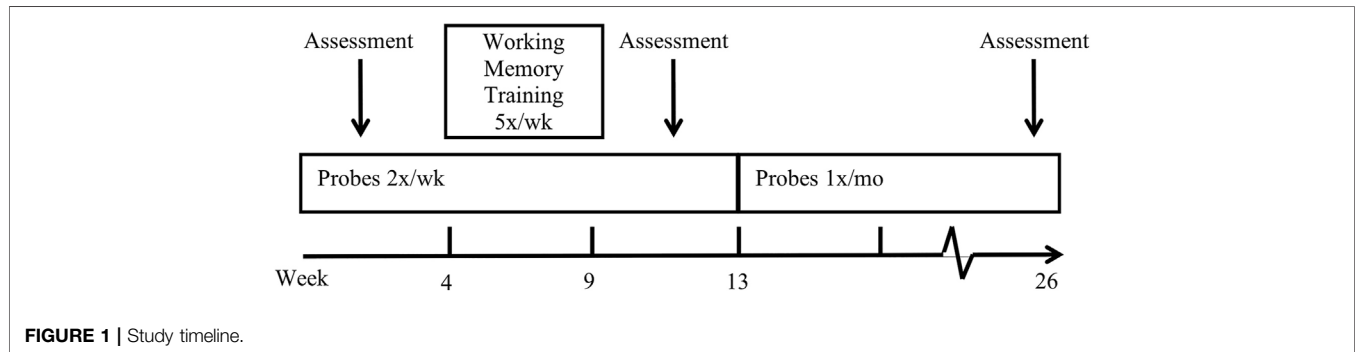
Participants were 7 children (5 male; M age 10.13 years), 6 of whom were recruited from the database of a study (Archibald et al., 2013) for which children completed an assessment battery on two occasions approximately 1 year apart. The battery included standardized measures of working memory, language, nonverbal intelligence, reading, and math at both time points, and parent and teacher reports at the first time point.

For the present study, children were recruited based on the presence of a working memory impairment and reported teacher concern in any area. The criteria for working memory impairment included a low score (≤ 87) on a working memory composite, which was calculated by averaging performance on 3 subtests from the *Automated Working Memory Assessment* (AWMA; Alloway, 2007): 2 visuospatial working memory tasks (Odd One Out, Spatial Recall) and 1 verbal working memory task (Listening Span). As well, all participants demonstrated at least one indicator of working memory impairment at the original time point of the study (e.g., low working memory score; teacher or parent concern; academic difficulties). Nonverbal intelligence standard scores were 85 or greater for all participants as measured by either the *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler, 1999) or the *Wechsler Preschool and Primary Scale of Intelligence—Third edition* (WPPSI-III; Wechsler, 2002) according to participant age. Language ability was measured in the earlier study using the *Composite Language Score* (CLS) from the *Clinical Evaluations of Language Fundamentals—Fourth edition* (CELF-4; Semel et al., 2003), but no requirements were placed on language ability for participation in this study. One additional participant was self-recruited to the study based on parent report and performance on standardized measures of working memory and nonverbal intelligence. Of the 7 children completing the study, 2 were considered to have a language impairment in addition to a working memory impairment (LWMI) while the remaining 5 had an impairment in working memory alone (specific working memory impairment; SWMI). Baseline measures for all participants are shown in **Table 1**. Approval for this study was obtained from the University of Western Ontario Research Ethics Board (file no. 101971).

TABLE 1 | Descriptive statistics for participant demographics.

| | <i>n</i> | Male | Age (yrs) | WM comp | CLS | PIQ |
|------------------|----------|------|--------------|--------------|---------------|---------------|
| SWMI | 5 | 4 | 9.65 (0.98) | 84.25 (3.30) | 98.50 (4.12) | 103.60 (4.28) |
| LWMI | 2 | 1 | 11.34 (1.65) | 81.67 (1.89) | 77.00 (2.82) | 100.50 (0.71) |
| All participants | 7 | 5 | 10.12 (1.33) | 83.51 (3.07) | 91.33 (11.62) | 102.71 (3.82) |

SWMI, Specific working memory impairment; LWMI, Language and working memory impairment; WM comp, Working Memory composite; CLS, Composite Language Score on the CELF-4; PIQ, performance intelligence quotient.



2.2 General Procedure

This study was one in a pair of concurrent studies designed to evaluate language intervention (Pauls and Archibald, 2021) and working memory training. Both studies consisted of three phases: Baseline, intervention, and follow up (Figure 1). Participants completed all intervention and assessment sessions individually in a quiet room in their school or home. An assessment battery consisting of standardized tests of language, short-term and working memory, reading and math was completed at the beginning of the study, immediately following completion of the intervention phase and at the end of the follow up phase. Four probe measures were completed 2 times per week throughout the baseline phase, intervention phase, and for the first 4 weeks of the follow-up phase. For the final 3 months of the follow-up phase, probe measures were administered monthly. For the intervention, children completed 20 to 25 training sessions over 5 weeks. All research sessions were completed by trained research assistants. Different research assistants completed the assessment, probe measures, and intervention sessions. All research assistants were blinded to the working memory status of the participant, and those administering the assessment and probe measures were blinded additionally to the purpose of the study.

2.3 Intervention Procedure

The Cogmed RM training program (Klingberg et al., 2005) required the completion of 20–25 sessions of approximately 40 min. In each session, participants completed 8 of 11 possible games (15 trials each) targeting visuospatial short-term memory (3 games), visuospatial working memory (4 games), verbal short-term memory (2 games), or verbal working memory (2 games). The trial difficulty increased with the child's successes and decreased after failed trials so that the child was always working at capacity. In the visuospatial tasks,

sequences of lights or space creatures were presented, and children were required to click on the locations of the items in the order they were shown. In the verbal short-term tasks, letters or digits were presented for serial, reverse, or numerical order recall.

2.4 Outcome Measures

2.4.1 Probe Measures

The probes were designed to measure changes in working memory and language. Tasks were adapted from those in literature or standardized tests and designed to be reliably and consistently administered and scored. The tasks measured verbal short-term memory (nonword repetition), visuospatial working memory (puzzle completion), or language (sentence combining). A control probe (number comparison) was expected to place no demands on either working memory or language, and was therefore expected to show no change. Improved performance on the control probe would suggest a generalized response to the intervention, not a domain-specific response.

The *Nonword Repetition probe* was designed to place demands primarily on verbal short-term memory. In this probe, children listened over personal headphones via an mp3 player to 3 audiorecorded trials of four 3-syllable nonwords (e.g., da-moy-cho, tay-chee-dow, tow-doy-foo, voo-ta-ye). Within each trial, 1 to 3 of the nonwords were spoken by a female voice and the others were spoken by a male voice. At the beginning of each session, children were instructed to listen for the nonwords spoken by the target voice for that session and recall those nonwords at the end of each trial. Between 4 and 8 nonwords, or 12 to 24 syllables, were recalled in each session. The Nonword Repetition score was the percent of target syllables correctly recalled.

The *Puzzle Completion probe* measured visuospatial working memory. In this probe, children were provided with 7 plastic

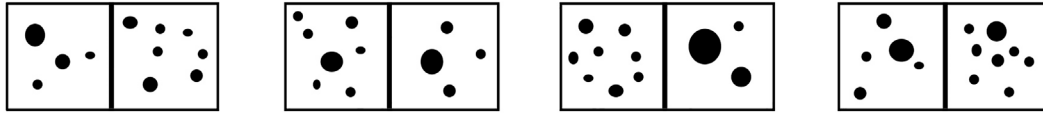


FIGURE 2 | Example of dot array pairs from the number comparison probe. Child draws a line through the array in each pair with more dots working as quickly as possible.

geometric shapes. For each trial, children were given 5 s to view a line drawing of a design, and then asked to recreate the design using 3 or 4 of the given shapes. Children were timed as they recreated the design. The score for each session was calculated by dividing the total number of shapes selected correctly by the total time required to recreate all three designs.

The *Sentence Combining probe* was a measure of language. In this probe, the research assistant read out two simple sentences (e.g., “Selena flies her kite”. “It is not very windy.”) and gave children two attempts to combine those sentences into one sentence (e.g., “Selena flies her kite even though it is not very windy”. “It is not very windy but Selena flies her kite anyway”). Three trials were completed each session using unique sentences. Sentences were transcribed and scored by calculating the ratio of propositions to words (proposition density; Kintsch and Keenan, 1973; Covington, 2009) in addition to propositions and words per sentence. Briefly, propositions are basic units of ideas in a text, and correspond approximately to verbs, adverbs, adjectives, prepositions, and subordinate conjunctions (Brown et al., 2008). For example, “Jason cleans up his toys at lunchtime” has 3 propositions (cleans up, his, at) and a propositional density of 0.43. In contrast “Her favourite dress is the one that looks like it has big polka dots” has 8 propositions (her, favourite, is, that, looks, like, has, big) and a propositional density of 0.57. The Sentence Combining probe was designed to place demands on language, specifically syntax. To minimize demands on memory, research assistants provided repetitions of the sentences when necessary.

Finally, for the *Number Comparison probe*, children were shown 56 to 60 pairs of dot arrays on a worksheet (Figure 2) and were timed as they crossed out the array in each pair that contained the greater number of dots. The score for each session was percent correct items.

2.4.2 Assessment Battery

The assessment battery included two language subtests from the CELF-4: *Concepts and Following Directions*, in which children pointed to objects as indicated by increasingly lengthy verbal instructions, and *Recalling Sentences*, in which children repeated sentences read aloud by the examiner. As measures of working memory, children completed 3 subtests from the AWMA: *Digit Recall*, *Counting Recall*, and *Spatial Recall*. In Digit Recall, children repeated lists of numbers of increasing length. In Counting Recall, children first counted red circles in arrays of mixed shapes, and at the end of the trial recalled their tallies. In Spatial Recall, children recalled locations of a red dot after first completing a mental rotation task on a shape associated with the red dot. Reading ability was assessed with the *Test of Word*

Reading Efficiency (TOWRE; Torgensen et al., 1999). In the *Phonemic Decoding Efficiency* (PDE) subtest, children were given 45 s to read as many nonwords as possible. In the *Sight Word Efficiency* (SWE) subtest, children were given 45 s to read as many words as possible. The *Reading Fluency* subtest from the *Woodcock-Johnson-III Tests of Achievement* was also completed (WJ-III; Woodcock et al., 2001), in which children read sentences and made truth judgments about them, completing as many as possible in 3 min. For math measures, the *Math Fluency* subtest from the WJ-III was completed in which children are given 3 min to solve simple addition, subtraction, and multiplication questions, as was the *Calculations* subtest from the WJ-III, in which children solved increasingly difficult arithmetic problems.

2.4.3 Analysis

For the probe data, statistically significant change was tested using the proportion/frequency approach (Bloom et al., 2006). Briefly, a 2 standard deviation band was calculated from baseline data points, which then served as a benchmark for determining whether data points in the intervention or follow-up phases were successes (exceeding the 2 *SD* band) or failures (falling below the 2 *SD* band). Using the principles of binomial probability, rates of success were compared across phases. For the Sentence Combining probe, a more lenient benchmark of 1 *SD* was used in order to capture the subtle changes commonly found following language intervention. Effect sizes were also calculated as standard mean differences (SMD; Busk and Serlin, 1992), an output broadly comparable to Cohen’s *d* (Cohen, 1988). An SMD of 0.8 or greater was interpreted as a clinically significant treatment effect. For measures standardized around a mean of 100, this translated to a minimum increase of 12 standard points. For scaled measures standardized around a mean of 10, a minimum increase of 3 points was required. Finally, qualitative analyses were conducted to examine possible associations between training effects and participant characteristics (age, baseline abilities, and improvement on training tasks).

3 RESULTS

All participants completed the required 20 sessions of working memory training ($M = 23.9$ days, range = 22–25 days). One participant (SWMI-5), however, was reported to show fatigue and low levels of interest in the training tasks. Therefore, the training aide for this participant decided to permit the child to complete the training tasks in two sittings each day. All other participants completed the required number of sessions in the customary timeframe.

TABLE 2 | Summary of results from probes and standardized measures of working memory, language, reading, and math.

| | Probes | | | | | Standardized measures | | |
|--------|-----------------|-----------------|-----------------|----------|---|-----------------------|--------------------------------------|-------------------|
| | Puzz comp | Nwd rep | Sent comb | Num comp | WM | Language | Reading | Math |
| LWMI-3 | ✓ ^{IF} | — | — | — | — | — | — | — |
| SWMI-5 | ✓ ^{IF} | — | — | — | SR ^I | — | — | — |
| SWMI-1 | ✓ ^{IF} | ✓ ^I | ✓ ^{IF} | — | SR ^I | — | — | — |
| SWMI-4 | ✓ ^{IF} | — | ✓ ^I | — | — | CFD ^F | — | — |
| SWMI-2 | ✓ ^{IF} | — | — | — | DR ^F , CR ^F , SR ^I | — | — | Calc ^F |
| SWMI-3 | ✓ ^{IF} | — | ✓ ^{IF} | — | DR ^I , CR ^I | CFD ^F | — | MF ^F |
| LWMI-4 | — | ✓ ^{IF} | ✓ ^{IF} | — | SR ^F | — | PDE ^{IF} , RF ^{IF} | — |

Note. ✓ Improvement in probes according to either 2 SD bandwidth or effect size calculations. ^I Improvement during or post-intervention. ^F Improvement during or at follow-up. Sent Comb, Sentence Combining probe; Nwd Rep, Nonword Repetition probe; Puzz Comp, Puzzle Completion probe; Num Comp, Number Comparison probe; CFD, Concepts and Following Directions; RS, Recalling Sentences; CR, Counting Recall; DR, Digit Recall; SR, Spatial Recall; PDE, Phonemic Decoding Efficiency; RF, Reading Fluency; MF, Math Fluency.

TABLE 3 | Effect sizes of probe measures.

| Participant | Sentence combining | | | | | | Nonword repetition | | Puzzle completion | | Number comparison | |
|-------------|--------------------|-------------|-------------|-------------|-------------|-------------|--------------------|-------------|-------------------|-------------|-------------------|-------|
| | Density | | Words | | Props | | I | F | I | F | I | F |
| | I | F | I | F | I | F | | | | | | |
| SWMI-1 | 0.73 | 1.05 | 2.03 | 3.61 | 1.85 | 3.25 | 1.99 | 0.72 | 1.04 | 2.64 | -1.25 | -1.91 |
| SWMI-2 | 0.08 | -0.21 | -0.18 | 0.65 | -0.10 | 0.56 | 0.74 | 0.76 | 1.14 | 1.94 | -0.29 | -0.31 |
| SWMI-3 | 0.66 | -0.21 | 0.72 | 3.27 | 0.68 | 1.71 | 0.45 | -0.61 | 1.29 | 5.01 | -1.57 | -1.90 |
| SWMI-4 | 0.63 | -0.25 | 0.05 | 0.15 | 0.64 | 0.02 | -0.48 | -0.10 | 0.98 | 1.98 | -0.02 | -0.08 |
| SWMI-5 | -0.70 | -0.79 | -0.25 | -0.15 | -0.75 | -0.78 | 0.72 | 0.31 | 2.15 | 1.86 | -1.42 | -0.44 |
| LWMI-3 | -0.64 | -0.94 | -0.26 | 0.06 | -0.43 | -0.41 | 0.08 | 0.12 | 2.38 | 3.53 | -1.55 | -0.80 |
| LWMI-4 | -0.22 | 0.52 | 0.56 | -0.52 | 0.53 | -0.27 | 2.62 | 1.35 | -0.35 | 0.13 | -0.15 | 0.12 |

Note. I = Intervention phase, F = Follow-up phase, Large effect sizes ($d \geq 0.8$) in bold.

3.1 Probe Measures

Results from the probes indicating improvement according to the proportion/frequency approach and effect size calculations are summarized in **Tables 2, 3**, respectively. Studying the results of the Puzzle Completion probe (**Figure 3**) reveals intervention effects for 6 participants as measured by both effect size and the proportion/frequency approach (SWMI-1, SWMI-2, SWMI-3, SWMI-4, SWMI-5, LWMI-3). Of these, 3 participants (SWMI-1, SWMI-5, LWMI-3) showed large significant effects at both intervention and follow-up phases. The remaining three (SWMI-2, SWMI-3, SWMI-4) showed large effects at both intervention and follow-up but only significant results at follow-up. Results of the Nonword Repetition probe (**Figure 4**) showed improvements for two of seven participants. SWMI-1 demonstrated significant gains during intervention with a large effect, and LWMI-4 showed large significant effects for both intervention and follow-up phases. On the Sentence Combining probe (**Figures 5, 6**), four participants showed treatment effects (SWMI-1, SWMI-3, SWMI-4, LWMI-4). SWMI-1 showed large significant effects for words and propositions per sentence in intervention and follow-up and a large effect for propositional density at follow-up. SWMI-3 showed large significant effects for words and propositions per sentence at follow-up and significantly fewer shorter sentences in intervention. SWMI-4 demonstrated significant gains in intervention only for propositional density and propositions with moderate effect

sizes ($d = 0.63, 0.64$, respectively). LWMI-4 showed increases in words during the intervention and increases in propositional density during the follow-up phases with moderate effect sizes ($d = 0.56, 0.52$ respectively). On the Number Comparison probe (**Figure 7**), all participants achieved relatively high accuracy, which resulted in the 2 SD cut-off exceeding the limits of the task. Therefore, the 2 SD limit was set to 100% for all participants. Despite high accuracy and a lenient cut-off, no participant demonstrated ceiling effects or improvement on this probe according to either the proportion/frequency method or effect size calculations.

3.2 Standardized Measures

Results of standardized measures of working memory, language, reading, and math are summarized in **Table 2**, which shows the subtests for which clinically significant changes were observed for the assessment at the immediate post intervention (I) or follow up phase (F). Standardized measures of working memory improved for five participants (SWMI-1, SWMI-2, SWMI-3, SWMI-5, LWMI-4). Three participants showed increases on Spatial Recall only with two improving immediately after the intervention (SWMI-1, SWMI-5) and the third improving at follow-up (LWMI-4). One participant (SWMI-2) improved on all measures, showing gains on Spatial Recall immediately after the intervention and gains on both verbal tasks at follow-up. One final participant

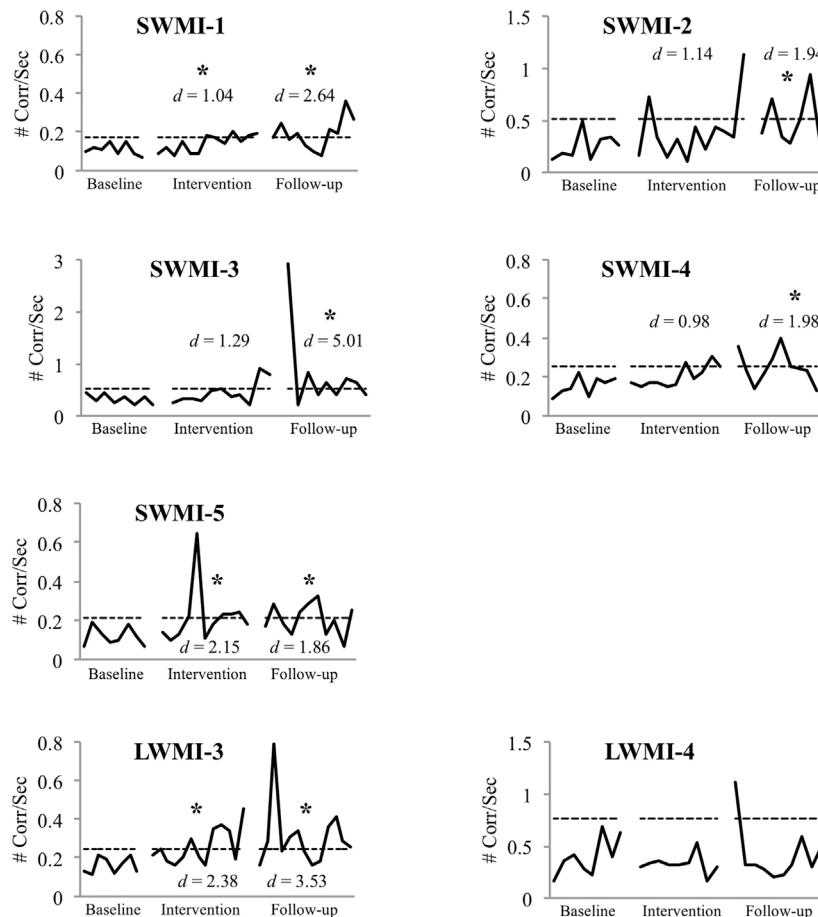


FIGURE 3 | Puzzle completion probe. Graphs present the correct number of shapes selected per second averaged over all three trials for each session. Dashed line represents 2 *SD* above mean score at baseline. Asterisks indicate significant improvement using 2 *SD* limit. All unmarked effect sizes $d < 0.8$.

(SWMI-3) demonstrated increases on the verbal measures immediately following intervention only.

Testing on standardized language measures revealed improvements for only two participants (SWMI-3, SWMI-4). In both cases, gains were seen at follow-up testing only. Results of standardized tests in reading and math showed limited treatment effects. Only LWMI-4 improved on reading measures (Phonemic Decoding Efficiency, Reading Fluency), but in both cases, gains were achieved post-training and at follow-up. Improvements on math measures were seen at follow-up only for two participants (SWMI-2, SWMI-3).

3.3 Responder Analysis

In examining responders, consider first the participants (LWMI-3, SWMI-5) who improved on only working memory measures (i.e., near transfer). LWMI-3 was differentiated by a markedly low Digit Recall score at baseline, scoring 2 *SD* below average for a child this age. This score was low according to standardized norms and relative to the scores of the other participants in the present study. On the other hand, SWMI-5 was set apart from other participants by age (youngest participant), training

intensity (completed 2 sessions per day), and performance on the training program. Having enrolled in the study at 8.1 years old, SWMI-5 was the youngest child to participate in this study; the other participants were between 1 and 4 years old. In addition, SWMI-5 was the only participant to complete the daily training in two shorter sessions rather than one longer session, and the only participant to demonstrate a steep downward trajectory in training performance.

Consider next the participants who improved on measures beyond working memory measures. This includes the two participants who improved on language but not academic measures (SWMI-1, SWMI-4), those who improved on both language and academic measures (LWMI-4, SWMI-3), and one who improved on only academic measures (SWMI-2). Review of baseline scores for these 5 participants revealed a possible effect of baseline working memory on far transfer. The 2 participants who made gains in multiple domains outside of working memory (LWMI-4, SWMI-3) also had some of the highest working memory scores at baseline. These participants additionally showed the greatest progress according to training data generated during performance (**Table 4**).

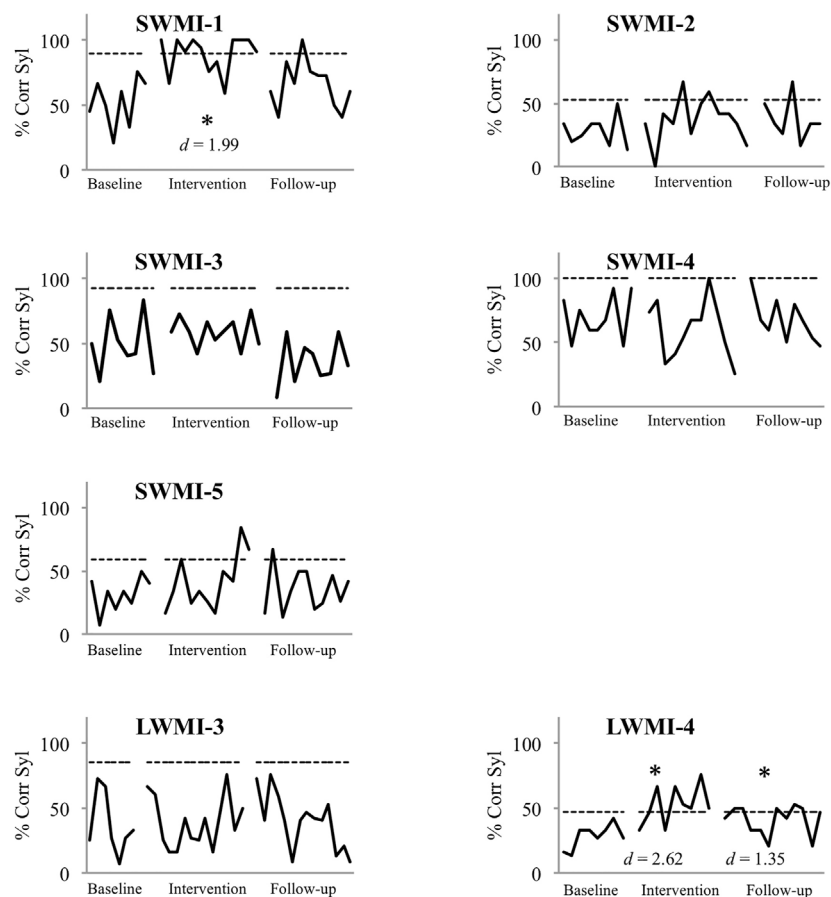


FIGURE 4 | Nonword repetition probe. Dashed line represents 2 *SD* above mean baseline score. Asterisks indicate significant improvement from baseline using 2 *SD* limit. All unmarked effect sizes $d < 0.8$.

Specifically, LWMI-4 and SWMI-3 demonstrated higher scores in Max Index and Improvement Index along with a positive slope. In comparison, those participants who improved on only one domain outside of working memory showed more modest training scores.

Lastly, three participants showed transfer effects on academic measures (LWMI-4, SWMI-3, SWMI-2). LWMI-4 was the only participant to improve on reading scores, and earned some of the lowest reading scores at baseline. Notably, compared with LWMI-3, LWMI-4 earned similar baseline reading scores, but substantially higher verbal short term memory scores. Both SWMI-2 and SWMI-3 showed gains in math at follow-up. They were also the only participants to improve on multiple standardized measures of working memory, suggesting an association between broad working memory growth and improvement in math. Interestingly, baseline math scores did not appear to differentiate SWMI-2 and SWMI-3 from other participants.

In summary, it appears that far transfer is more likely among participants with some minimum short term memory span and for those who completed the training program with the required intensity. Working memory ability seems to be linked to far

transfer in that transfer to multiple domains outside of working memory was associated with higher working memory scores at baseline and greater gains on training tasks. Reading gains appeared to be associated with lower baseline reading scores whereas math gains seemed to be associated with broad working memory improvement but not baseline math scores.

4 DISCUSSION

The purpose of this study was first to test the effectiveness of working memory training on children with working memory impairment, and second, to examine the effect of working memory training on related domains including language, reading, and math. Results of this single-subject design showed near transfer effects for all participants according to improved performance on either the visuospatial working memory probe or a standardized measure of working memory. As well, over half of the participants showed far transfer effects, that is, improvements in language, reading, or math performance. In addition to making working memory gains, two participants improved on language measures, one improved on a math measure, and two others

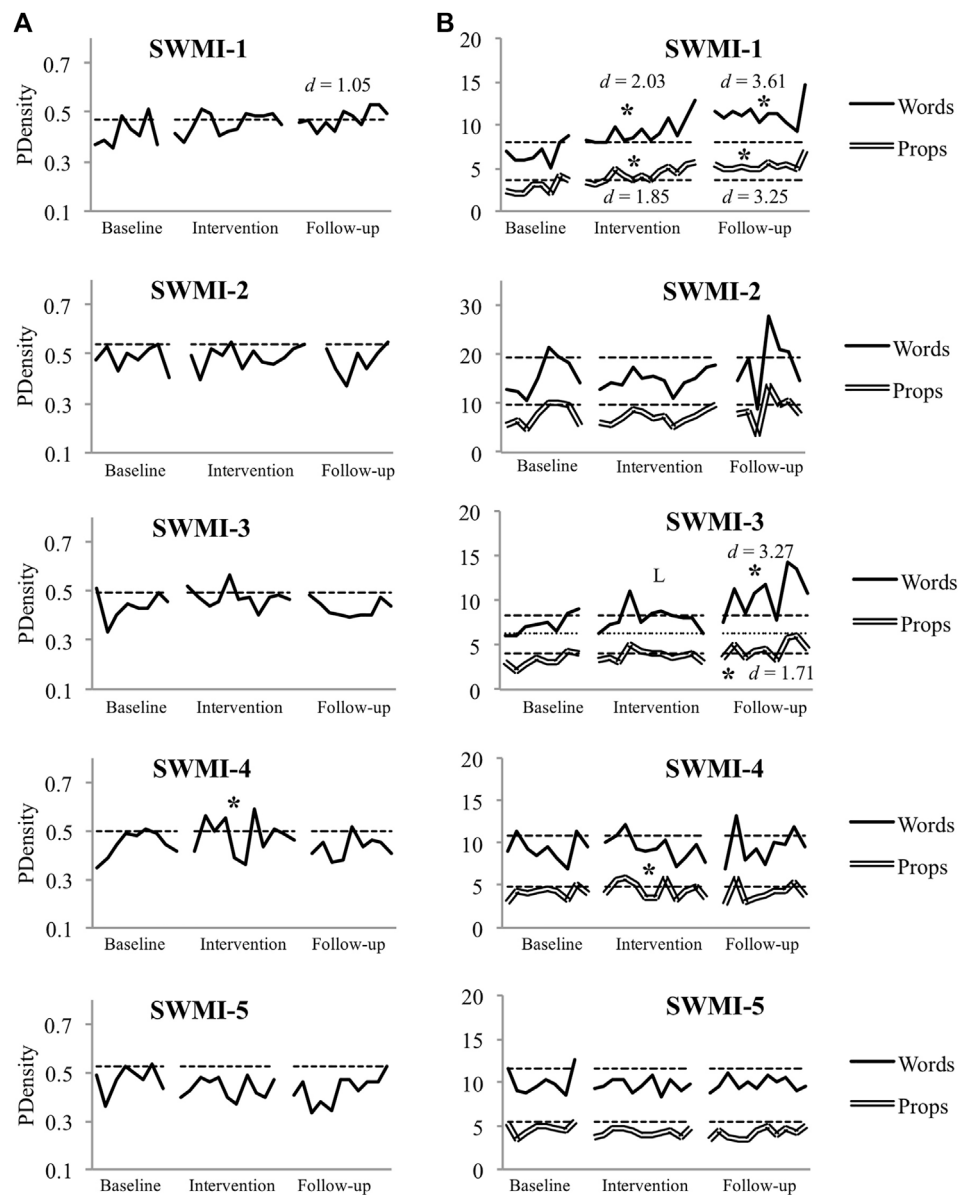


FIGURE 5 | Sentence Combining probe for participants with working memory impairment only (SWMI). Graphs represent three scores averaged over each session: the ratio of propositions to words [propositional density; **(A)**], words per trial, and propositions per trial **(B)**. Dashed line represents 1 *SD* above mean baseline performance (+1 *SD*). Where included, dotted line represents 1 *SD* below mean baseline performance (-1 *SD*). Asterisks indicate significance according to +1 *SD* limit. L indicates significance according to -1 *SD* limit. All unmarked effect sizes $d < 0.8$.

improved on both language and academic measures. A qualitative responder analysis revealed that likelihood of any kind of far transfer may be influenced by age, training intensity, and baseline verbal short term memory span. In addition, transfer to reading appeared to be more likely for children with lower reading abilities at baseline, provided verbal short term memory abilities were not severely impaired. In contrast, transfer to math appeared to be associated with broad gains in working memory rather than to baseline math abilities.

Near transfer effects found in the present study replicate those seen elsewhere both among children with low working memory

ability (Holmes et al., 2009; Gray et al., 2012; Dunning et al., 2013) and among children with other ability levels (Karbach et al., 2015; Peng and Fuchs, 2015). These results were not surprising because, as in previous studies, outcome measures of working memory probed skills similar to those targeted in the training tasks.

The second aim of this study was to test the effect of working memory training on related domains, including language. The influence on language was examined by including participants with language impairment, and by measuring language gains on probes and standardized testing. The four participants who showed language gains did so on at least 2 measures (i.e., 2

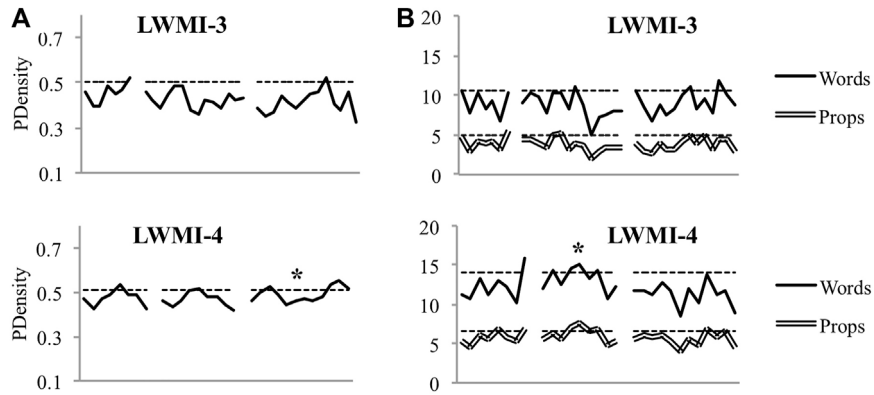


FIGURE 6 | Sentence Combining probe for participants with language and working memory impairments (LWMI). Graphs represent three scores averaged over each session: the ratio of propositions to words [propositional density; (A) left column], words per trial, and propositions per trial (B). Dashed line represents 1 SD above mean baseline performance (+1 SD). Where included, dotted line represents 1 SD below mean baseline performance (-1 SD). Asterisks indicate significance according to +1 SD limit. L indicates significance according to -1 SD limit. All unmarked effect sizes $d < 0.8$.

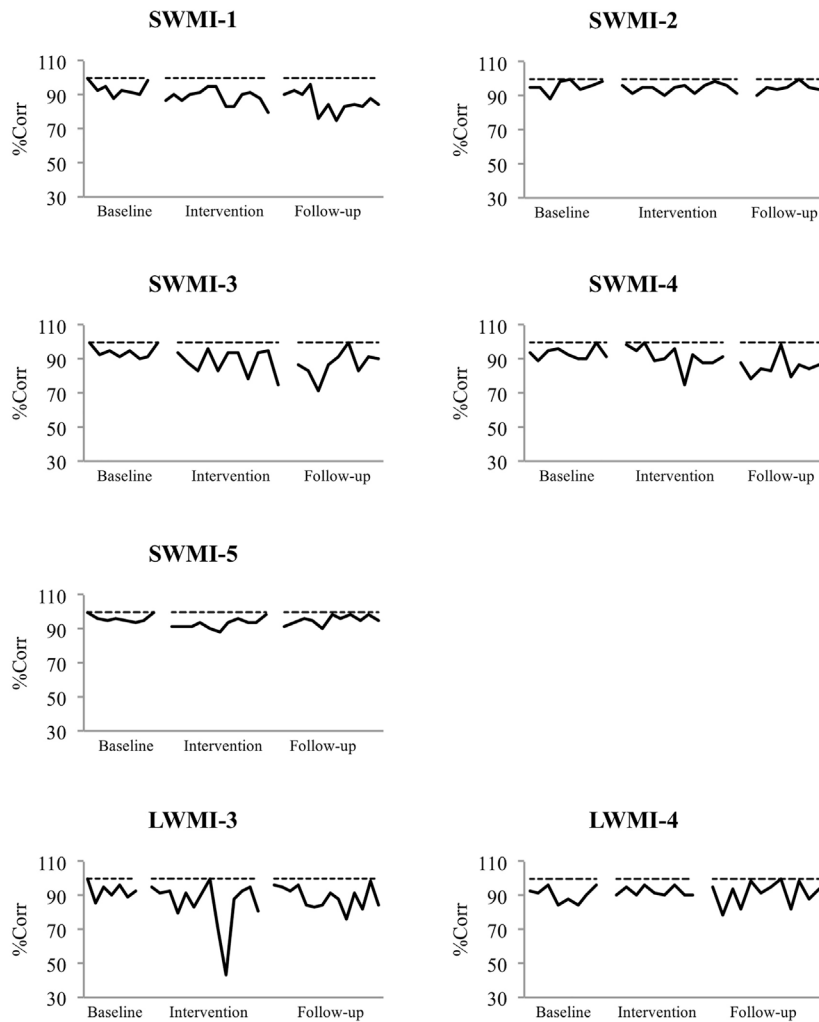


FIGURE 7 | Number Comparison probe. Graphs present percent items correct from each session. Dashed line indicates 100% items correct in place of 2 SD limit.

TABLE 4 | Progress scores from working memory training performance.

| | Responder type | Start index | Max index | Improvement index | Slope |
|--------|----------------|-------------|-----------|-------------------|-------|
| LWMI-3 | WMO | 53 | 70 | 17 | 0.63 |
| SWMI-5 | WMO | 59 | 75 | 17 | -2.21 |
| SWMI-1 | +La | 59 | 69 | 11 | -0.19 |
| SWMI-4 | +La | 68 | 76 | 8 | -0.09 |
| LWMI-4 | +La,Re | 67 | 87 | 21 | 0.86 |
| SWMI-3 | +La,Ma | 71 | 88 | 16 | 0.49 |
| SWMI-2 | +Ma | 71 | 82 | 11 | -0.18 |

Note. WMO = improved on working memory measures only, +La = improved on working memory and language measures, +La,Re = improved on working memory, language, and reading measures, +La,Ma = improved on working memory, language, and math measures, +Ma = improved on working memory and math measures.

probes or 1 probe and 1 standardized measure) but none improved on all language measures. These mixed findings are in line with the divide in existing literature between those studies showing language gains (e.g., Holmes et al., 2009; Peng and Fuchs, 2015) and those that show no effect on language (Holmes et al., 2010; Dunning et al., 2013). As well, neither of the children with language impairment improved on any language measures and an additional participant (SWMI-2) who demonstrated ample evidence of working memory gains also did not improve on any of the language measures. This pattern of findings suggests that working memory training does not generally lead to reliable changes in language abilities, and is not an effective intervention for those with language impairment.

Far transfer effects to reading and math were not widespread in this study, falling in line with findings from a recent meta-analysis showing no reliable far transfer to either reading or math (Melby-Lervåg et al., 2016). Nevertheless, the limited far transfer in the present study replicates patterns seen in other research. First, the participant who improved on reading performance had some of the lowest reading scores at baseline (Dahlin, 2011; Karbach et al., 2015). Second, far transfer to math performance was seen only at follow-up testing (Holmes et al., 2009; Dahlin, 2013; Holmes and Gathercole, 2014). Despite these modest results, it could be argued that these three participants benefited from working memory training. Each showed improvements on 5-6 outcome measures in the present study including at least one far transfer measure with gains persisting or observed at follow up. These responders tended to be those with higher working memory scores at baseline.

Collectively, the results presented here bring to light the complexities of far transfer from working memory training and the many factors that influence it. Possible moderating factors were specific to participants and, in combination with other known influences on far transfer such as intensity of training and similarity of training and transfer task demands, are indicative of the complex nature of cognitive development and the connections between domains. Moreover, this complex interaction speaks to the inconsistent effects of working memory training in the literature. It may be that working memory training is better suited for participants with particular profiles. Future research is needed to examine in more detail the interaction of participant characteristics that are likely to predict response to working memory training.

This study employed a single-subject design in order to evaluate individual responses to working memory training in terms of both near and far transfer. Given the lack of reliable group level far transfer effects reported for working memory training (Melby-Lervåg et al., 2016), we argued that a systematic focus on individual training response is needed to understand outcomes. To this end, we employed three probe measures that were designed to respond to changes in either working memory (near transfer) or language (far transfer), and one additional control probe expected to show no response to training. One limitation is the variable baselines in these probes for many participants (Figures 3–7). This variability, however, results in large 2 SD bands, which decreases the likelihood of achieving “improbable improvements” in our proportion/frequency statistical analysis. Therefore, effects reported here are likely underestimated rather than overestimated. Given these limitations, caution is warranted in interpreting and generalizing the present results.

In summary, results of this study have revealed a number of findings. First, working memory training can lead to immediate and long term near transfer gains among children with working memory impairment. Second, working memory training was associated with modest far transfer effects to academic skills for some children with working memory impairment. Finally, whether or not a participant is likely to exhibit far transfer effects may be influenced by a number of participant-specific characteristics including baseline working memory and academic abilities.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because participants of this study did not agree for their data to be shared. Requests to access the datasets should be directed to LA, larchiba@uwo.ca.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The University of Western Ontario Nonmedical

Research Ethics Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Participants provided written assent to participate.

AUTHOR CONTRIBUTIONS

LP and LA contributed to conception and design of the study. LP organized the data, performed the data analysis, and wrote the first draft of the manuscript. LP and LA contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

- Alloway, T. P. (2007). *Automated Working Memory Assessment*. London: Pearson Assessment.
- Alloway, T. P., Bibile, V., and Lau, G. (2013). Computerized Working Memory Training: Can it lead to Gains in Cognitive Skills in Students. *Comput. Hum. Behav.* 29, 632–638. doi:10.1016/j.chb.2012.10.023
- Archibald, L. M., Oram Cardy, J., Joannis, M. F., and Ansari, D. (2013). Language, Reading, and Math Learning Profiles in an Epidemiological Sample of School Age Children. *PLoS ONE* 8 (10), e77463. doi:10.1371/journal.pone.0077463
- Archibald, L. M., and Gathercole, S. E. (2006). Short-term and Working Memory in Specific Language Impairment. *Int. J. Lang. Commun. Disord.* 41 (6), 675–693. doi:10.1080/13682820500442602
- Baddeley, A. (2003). Working Memory and Language: An Overview. *J. Commun. Disord.* 36, 189–208. doi:10.1016/S0021-9924(03)00019-4
- Barlow, D. H., and Hersen, M. (1973). Single-case Experimental Designs. Uses in Applied Clinical Research. *Arch. Gen. Psychiatry* 29, 319–325. doi:10.1001/archpsyc.1973.04200030017003
- Bloom, M., Fischer, J., and Orme, J. G. (2006). *Evaluating Practice: Guidelines for the Accountable Professional*. 5th ed. Boston, MA: Pearson.
- Bordens, K. S., and Abbott, B. B. (2011). *Research Design and Methods: A Process Approach*. 8th ed. Boston, MA: McGraw-Hill.
- Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., and Covington, M. A. (2008). Automatic Measurement of Propositional Idea Density from Part-of-Speech Tagging. *Behav. Res. Methods* 40 (2), 540–545. doi:10.3758/BRM.40.2.540
- Busk, P. L., and Serlin, R. C. (1992). "Meta-Analysis for Single-Case Research," in *Single-Case Research Designs and Analysis: New Directions for Psychology and Education*. Editor T. R. Kratochwill and J. R. Levin (Hillsdale, NJ: Lawrence Erlbaum Associates), 187–212.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Erlbaum.
- Covington, M. A. (2009). "Idea Density – A Potentially Informative Characteristic of Retrieved Documents," in *Proceeding of the Southeastcon, 2009 IEEE*, Atlanta, GA, USA, March 2009 (IEEE), 201–203.
- Dahlin, K. I. E. (2011). Effects of Working Memory Training on reading in Children with Special Needs. *Reading and Writing* 24, 479–491. doi:10.1007/s11145-010-9238-y
- Dahlin, K. I. E. (2013). Working Memory Training and the Effect on Mathematical Achievement in Children with Attention Deficits and Special Needs. *J. Edu. Learn.* 2 (1), 118–133. doi:10.5539/jel.v2n1p118
- Dawes, E., Leitão, S., Claessen, M., and Nayton, M. (2015). A Profile of Working Memory Ability in Poor Readers. *Aust. Psychol.* 50, 362–371. doi:10.1111/ap.12120
- Dunning, D. L., Holmes, J., and Gathercole, S. E. (2013). Does Working Memory Training lead to Generalized Improvements in Children with Low Working Memory? A Randomized Controlled Trial. *Dev. Sci.* 16 (6), 915–925. doi:10.1111/desc.12068

FUNDING

This work was supported by an early research award from the Ontario Ministry of Research and Innovation and institutional matching funds from The University of Western Ontario to the second author.

ACKNOWLEDGMENTS

The work of dedicated research assistants and the participation of children and schools is gratefully acknowledged. We thank CogMed for allowing access to the training program for the purposes of this study.

- Egeland, J., Aarlien, A. K., and Saunes, B. K. (2013). Few Effects of Far Transfer of Working Memory Training in ADHD: A Randomized Controlled Trial. *PLoS One* 8 (10), e75660. doi:10.1371/journal.pone.0075660
- Gathercole, S. E. (2006). Nonword Repetition and Word Learning: The Nature of the Relationship. *Appl. Psycholinguistics* 27 (4), 513–543. doi:10.1017/S0142716406060383
- Gathercole, S. E., Pickering, S. J., Knight, C., and Stegmann, Z. (2004). Working Memory Skills and Educational Attainment: Evidence from National Curriculum Assessments at 7 and 14 Years of Age. *Appl. Cognit. Psychol.* 18, 1–16. doi:10.1002/acp.934
- Graf Estes, K., Evans, J. L., and Else-Quest, N. M. (2007). Differences in the Nonword Repetition Performance of Children with and without Specific Language Impairment: A Meta-Analysis. *J. Speech Lang. Hear. Res.* 50, 177–195. doi:10.1044/1092-4388(2007)015
- Gray, S. A., Chaban, P., Martinussen, R., Goldberg, R., Gotlieb, H., Kronitz, R., et al. (2012). Effects of a Computerized Working Memory Training Program on Working Memory, Attention, and Academics in Adolescents with Severe LD and Comorbid ADHD: A Randomized Controlled Trial. *J. Child. Psychol. Psychiatry* 53 (12), 1277–1284. doi:10.1111/j.1469-7610.2012.02592.x
- Holmes, J., Butterfield, S., Cormack, F., van Loenhoud, A., Ruggero, L., Kashikar, L., et al. (2015). Improving Working Memory in Children with Low Language Abilities. *Front. Psychol.* 6, 519. doi:10.3389/fpsyg.2015.00519
- Holmes, J., Gathercole, S. E., and Dunning, D. L. (2009). Adaptive Training Leads to Sustained Enhancement of Poor Working Memory in Children. *Dev. Sci.* 12 (4), F9–F15. doi:10.1111/j.1467-7687.2009.00848.x
- Holmes, J., and Gathercole, S. E. (2014). Taking Working Memory Training from the Laboratory into Schools. *Educ. Psychol. (Lond)* 34 (4), 440–450. doi:10.1080/01443410.2013.797338
- Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., and Elliott, J. G. (2010). Working Memory Deficits Can Be Overcome: Impacts of Training and Medication on Working Memory in Children with ADHD. *Appl. Cognit. Psychol.* 24, 827–836. doi:10.1002/acp.1589
- Horner, R. H., Swaminathan, H., Sugai, G., and Smolkowski, K. (2012). Considerations for the Systematic Analysis and Use of Single-Case Research. *Edu. Treat. Child.* 35 (2), 269–290. doi:10.1353/etc.2012.0011
- Karbach, J., Strobach, T., and Schubert, T. (2015). Adaptive Working-Memory Training Benefits reading, but Not Mathematics in Middle Childhood. *Child. Neuropsychol.* 21 (3), 285–301. doi:10.1080/09297049.2014.899336
- Kintsch, W., and Keenan, J. (1973). Reading Rate and Retention as a Function of the Number of Propositions in the Base Structure of Sentences. *Cogn. Psychol.* 5, 257–274. doi:10.1016/0010-0285(73)90036-4
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., and Dahlström, K. (2005). Computerized Training of Working Memory in Children with ADHD—A Randomized, Controlled Trial. *J. Am. Acad. Child Adolesc. Psychiatry.* 44 (2), 177–186. doi:10.1097/00004583-200502000-00010
- Majerus, S. (2013). Language Repetition and Short-Term Memory: An Integrative Framework. *Front. Hum. Neurosci.* 7, 357. doi:10.3389/fnhum.2013.00357
- Martinussen, R., Hayden, J., Hogg-Johnson, S., and Tannock, R. (2005). A Meta-Analysis of Working Memory Impairments in Children with Attention-Deficit/

- hyperactivity Disorder. *J. Am. Acad. Child. Adolesc. Psychiatry* 44 (4), 377–384. doi:10.1097/01.chi.0000153228.72591.73
- McReynolds, L. V., and Thompson, C. K. (1986). Flexibility of Single-Subject Experimental Designs. Part I: Review of the Basics of Single-Subject Designs. *J. Speech Hear. Disord.* 51, 194–203. doi:10.1044/jshd.5103.194
- Melby-Lervåg, M., Redick, T. S., and Hulme, C. (2016). Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of "Far Transfer". *Perspect. Psychol. Sci.* 11 (4), 512–534. doi:10.1177/1745691616635612
- Montgomery, J. W., Magimairaj, B. M., and O'Malley, M. H. (2008). Role of Working Memory in Typically Developing Children's Complex Sentence Comprehension. *J. Psycholinguist Res.* 37, 331–354. doi:10.1007/s10936-008-9077-z
- Montgomery, J. W., Gillam, R. B., and Evans, J. L. (2021). A New Memory Perspective on the Sentence Comprehension Deficits of School-Age Children with Developmental Language Disorder: Implications for Theory, Assessment, and Intervention. *Lang. Speech, Hearing Serv. Schools* 52 (2), 449–466. doi:10.1044/2021_LSHSS-20-00128
- Nouwens, S., Groen, M. A., and Verhoeven, L. (2017). How Working Memory Relates to Children's Reading Comprehension: The Importance of Domain-Specificity in Storage and Processing. *Read. Writ* 30, 105–120. doi:10.1007/s11145-016-9665-5
- Pauls, L. J., and Archibald, L. M. (2021). Cognitive and Linguistic Effects of Narrative-Based Language Intervention in Children with Developmental Language Disorder. *Autism Develop. Lang. Impairments* 6, 239694152110158. doi:10.1177/23969415211015867
- Peng, P., and Fuchs, D. (2017). A Randomized Control Trial of Working Memory Training with and without Strategy Instruction: Effects on Young Children's Working Memory and Comprehension. *J. Learn. Disabil.* 50 (1), 62–80. doi:10.1177/0022219415594609
- Perdices, M., and Tate, R. L. (2009). Single-subject Designs as a Tool for Evidence-Based Clinical Practice: Are They Unrecognised and Undervalued. *Neuropsychol. Rehabil.* 19 (6), 904–927. doi:10.1080/09602010903040691
- Roberts, R., and Gibson, E. (2002). Individual Differences in Sentence Memory. *J. Psycholinguist Res.* 31 (6), 573–598. doi:10.1023/A:1021213004302
- Semel, E., Wiig, E. H., and Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals*. 4th ed. San Antonio, TX: The Psychological Corporation.
- Swanson, H. L., and Jerman, O. (2006). Math Disabilities: A Selective Meta-Analysis of the Literature. *Rev. Educ. Res.* 76 (2), 249–274. doi:10.3102/003465430760224910.3102/00346543076002249
- Torgensen, J. K., Wagner, R. K., and Rashotte, C. S. (1999). *Test of Word Reading Efficiency*. Austin, TX: AGS Publishing.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence*. 3rd ed. Sydney, NSW: Pearson Assessments.
- Woodcock, R. W., McGrew, K. S., and Mather, N. (2001). *Woodcock Johnson Tests of Achievement*. 3rd ed. Itasca, IL: Riverside Publishing.
- 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.
- 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.
- Copyright © 2022 Pauls and Archibald. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.