HHS Public Access
Author manuscript
J Learn Disabil. Author manuscript; available in PMC 2017 January 01.

Published in final edited form as:
J Learn Disabil. 2017 January ; 50(1): 62–80. doi:10.1177/0022219415594609.

A Randomized Control Trial of Working Memory Training With and Without Strategy Instruction: Effects on Young Children's Working Memory and Comprehension

Peng Peng, Ph.D. and
Department of Special Education and Disability Studies, George Washington University

Douglas Fuchs, Ph.D.
Department of Special Education, Peabody College at Vanderbilt University

Abstract

Researchers are increasingly interested in working memory (WM) training. However, it is unclear whether it strengthens comprehension in young children who are at risk for learning difficulties. We conducted a modest study of whether the training of verbal WM would improve verbal WM and passage listening comprehension, and whether training effects differed between two approaches: training with and without strategy instruction. A total of 58 first-grade children were randomly assigned to 3 groups: WM training with a rehearsal strategy, WM training without strategy instruction, and controls. Every member of the 2 training groups received a one-to-one, 35-minute session of verbal WM training on each of 10 consecutive school days, totaling 5.8 hours. Both training groups improved on trained verbal WM tasks, with the rehearsal group making greater gains. Without correction for multiple group comparisons, the rehearsal group made reliable improvements over controls on an untrained verbal WM task and on passage listening comprehension and listening retell measures. The no-strategy- instruction group outperformed controls on passage listening comprehension. When corrected for multiple contrasts, these group differences disappeared, but were associated with moderate-to-large effect sizes. Findings suggest—however tentatively—that brief but intensive verbal WM training may strengthen the verbal WM and comprehension performance of young children at risk. Necessary caveats and possible implications for theory and future research are discussed.

Keywords

verbal working memory; working memory training; rehearsal strategy; listening comprehension; young at-risk children

A Randomized Control Trial of Working Memory Training With and Without Strategy Instruction

Working memory (WM) refers to the capacity to store information temporarily when engaging in cognitively demanding activities (Baddeley, 1986). Compared to short-term

Correspondence to: Douglas Fuchs.

memory, WM plays a more influential role in children's academic performance (Baddeley, 1986). This is because many academic tasks involve multiple steps with intermediate solutions that must be remembered for a short time to accomplish the task at hand (Shah & Miyake, 1996). For example, when reading a passage for comprehension, children must remember previously learned information while simultaneously integrating incoming information as they progress through a text (Cain, Oakhill, & Bryant, 2004).

In recent years, increasing numbers of researchers have explored whether training children's WM indeed strengthens this cognitive ability as well as improves academic performance. Findings are mixed. Investigators of several studies reported that their training improved children's WM and academic skills, like reading comprehension and mathematics reasoning (e.g., Dahlin, 2011; Holmes, Gathercole, & Dunning, 2009). But most researchers failed to find such effects. The authors of two recent reviews of WM training (Melby-Lervag & Hulme, 2012; Shipstead, Redick, & Engle, 2012) concluded that, for children between the ages of 8–15, WM training involving visual-spatial tasks, or a combination of visual-spatial and verbal tasks, can improve visual-spatial WM. But these reviews also found small or no transfer of this effect to verbal WM or academic performance.

Several issues should be considered in connection with these inconsistent results. First, the WM training of children has most often involved visual-spatial tasks, or a combination of visual-spatial and verbal tasks. Training focused solely on verbal tasks has occurred much less often (Melby-Lervag & Hulme, 2012). So, one may reasonably ask whether relatively intensive training of children's verbal WM might improve their verbal WM and academic skills. Second, it has only infrequently been the case that WM training has involved strategy instruction. So, a second pertinent question is whether strategy (e.g., rehearsal) use may be a more beneficial training method. Third, previous studies have centered almost exclusively on typically developing children in the intermediate grades. The importance of WM training for younger children who are at risk for learning problems is largely unknown. Each of these issues will be discussed in turn to provide proper background for the aims of this study.

Training Tasks and Domain-General vs. Domain-Specific Models of WM

One factor that may contribute to WM training's inconsistent effects is ongoing disagreement about the proper content of the training; specifically, the nature of the training tasks. This lack of agreement reflects a longstanding debate about two competing WM models: domain general vs. domain specific (Shah & Miyake, 1996).

Many researchers believe WM is a domain-general construct. Baddeley (1996) has famously explained this construct in terms of a multicomponent model, which includes a "visuo-spatial sketchpad" component that stores and manipulates visual images; a "phonological loop" component that does the same for verbal information; and a central executive function that coordinates the visuo-spatial sketchpad and phonological loop components. According to Baddeley (1996), the central executive function is pivotal to WM because it not only coordinates the components, but also directs attention to relevant information, suppresses irrelevant information, and manages cognitive processes when multiple tasks must be accomplished simultaneously (e.g., Engle, 2002). Baddeley's (1996) multicomponent conceptualization of WM is domain general; it is meant to pertain to any and all domains.

Hence, the type of tasks (e.g., verbal or visual spatial) used for WM training should not influence training effects.

Others take a different view. They understand WM as closely related to skills and knowledge specific to a given domain (e.g., Ericsson & Kintsch, 1995; Unsworth & Engle, 2007). According to this alternate perspective, WM should be trained as part of domain-specific activities. Verbal WM training, as an example, would be expected to be more effective than visual-spatial WM training for improving performance on verbal WM tasks and verbal-related academic skills.

Although there is empirical evidence to support both domain-general and domain-specific models (e.g., Ericsson & Kintsch, 1995; Shah & Miyake, 1996), studies focusing on *children's* learning seem to favor the domain-specific perspective. Research, for example, has shown that children's visual-spatial WM fails to explain variance in their word reading and passage comprehension (e.g., Nation, Adams, Bowyer, Crane, & Snowling, 1999; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000). By contrast, verbal WM accounts for statistically significant variance in performance on these verbal tasks, even when relevant verbal skills (e.g., word reading) are controlled (Cain et al., 2004; Daneman & Carpenter, 1980).

Further support of a domain-specific view comes from scholarly reviews of WM deficits among children with learning difficulties (Swanson & Jerman, 2006; Swanson, Zheng, & Jerman, 2009). These reviews indicate that, although children with serious learning problems exhibit WM deficits across verbal and visual-spatial domains, verbal WM deficits appear more important to the children with reading difficulties (Swanson et al., 2009). Visual-spatial deficits, by contrast, seem more relevant for children with mathematics difficulties (Swanson & Jerman, 2006). Moreover, the researchers of most previous WM training studies with children used visual-spatial WM tasks. Few reported training effects that transferred to verbal WM or academic performance (Shipstead et al., 2012). Taken together, research suggests that training children's verbal WM might strengthen their verbal WM and verbal-related academic skills.

We know of only two studies that investigated the effects of verbal WM training on children (Kroesbergen, van't Noordende, & Kolkman, 2014; Swanson, Kehler, & Jerman, 2010). Swanson et al. (2010) randomly assigned children to two groups: verbal WM training and controls. Both groups were matched on chronological age, IQ, and reading skills. The children who practiced verbal WM for a total of 15 minutes reliably improved their performance relative to controls on trained, but not on untrained, tasks. Swanson et al. (2010) did not explore whether these training effects transferred to academic performance.

Kroesbergen et al. (2014) reported that children practicing verbal (i.e., numerical) WM tasks did not improve their verbal WM but did improve their numeracy skills in comparison to age-matched controls. This latter finding may reflect that Kroesbergen et al. (2014) also gave students intensive numeracy skills training. Thus, it is unclear whether the researchers' verbal WM training or numeracy training (or both) affected the children's academic performance. In sum, findings from this very small group of studies are not easily interpretable, principally because so little work has explored the value of verbal WM training among children. More research is needed that investigates whether training children's verbal WM improves their verbal WM and verbal-related academic skills.

Training With and Without Strategy Instruction

In addition to whether WM should be considered domain general or domain specific, there is a second perhaps less frequently discussed issue that also bears importantly on WM training; namely, whether training with strategy instruction (e.g., rehearsing stimuli to be remembered; Swanson et al., 2010) is more efficacious than training without it (e.g., Klingberg, 2010). These two approaches, like domain general and domain specific perspectives, reflect different ideas about the nature of WM. We refer, here, to Strategy Mediation Theory (e.g., Bailey, Dunlosky, & Kane, 2008; McNamara & Scott, 2001) and Capacity Theory (e.g., Engle & Kane, 2004).

With strategy instruction—Strategy Mediation Theory recognizes WM as a finite and relatively fixed cognitive capacity. This view holds that differences in WM performance are determined by the efficiency with which the finite WM capacity is used (e.g., Daneman & Carpenter, 1980; Engle & Marshall, 1983). The efficient use of strategies can free up, or make available, more cognitive resources for the higher-level central executive (e.g., McNamara & Scott, 2001), which in term can strengthen WM performance. "Strategies" often discussed in this context are those (like rehearsal) that are applied to WM subsystems such as the short-term storage of visual or verbal information (Baddeley, 1996; McNamara & Scott, 2001; Turley-Ames & Whitfield, 2003).

Research on WM development and individual differences and WM training provides support for Strategy Mediation Theory. Research on WM development suggests that age differences in performance on WM tasks are the result of older children's more active application of strategies (Hagen, Jongeward, & Kail, 1975). Research on individual differences suggests that strategy use accounts for a reliable proportion of variance in WM performance, with stronger performance associated with more frequent use of strategies, or use of more effective strategies (Dunlosky & Kane, 2007; Friedman & Miyake, 2004; McNamara & Scott, 2001; Turley-Ames & Whitfield, 2003). Finally, studies involving the training of adults on WM tasks indicate that strategy instruction significantly improves their performance, whereas training without it does not (e.g., McNamara & Scott, 2001).

Strategy use may mediate relationships between WM and higher-order cognitive activities, such as listening and reading comprehension, when identical strategies can be applied to both the WM and the cognitive tasks (Bailey, Dunlosky, & Kane, 2008). Rehearsal strategies in particular appear to facilitate WM and comprehension. Rehearsal is considered an important subcomponent of verbal working memory (Baddeley, 1986). Adults are often observed to rehearse to prevent forgetting information on WM tasks (McNamara & Scott, 2001); and children have been taught to rehearse to improve their comprehension of written passages, or passages read to them, and their retelling of text (e.g., Gersten, Fuchs, Williams, & Baker, 2001; Rose, Cundick, & Higbee, 1983). In addition, rehearsal is easier to learn and less demanding of cognitive resources than other strategies (e.g., semantics associations;

Turley-Ames & Whitfield, 2003), which suggests that rehearsal training on verbal WM tasks may improve children's verbal WM and passage comprehension. That said, in comparison to adults, WM tasks are more challenging to children, especially young children (Pickering & Gathercole, 2001). Thus, children may need more explicit instructions and more time to fluently use a rehearsal strategy when engaging in WM tasks.

Two studies investigated the effects of rehearsal training on children. St. Clair-Thompson, Stevens, Hunt, and Bolder (2010) trained children to use multiple strategies, including rehearsal, when performing short-term memory tasks. Compared to no-treatment controls, children who were taught multiple strategies improved their scores on a verbal WM task. There was no transfer to standardized reading or math measures. Swanson et al. (2010) reported that their rehearsal training group showed significantly greater improvement than matched controls on trained, but not on untrained, verbal WM tasks.

Although findings from these studies suggest rehearsal training may improve performance on verbal WM tasks, this possibility is complicated by several considerations. First, St. Clair-Thompson et al.'s (2010) strategy training involved mostly short-term memory tasks. Second, it is unclear from their study whether rehearsal was an "active ingredient" in the training since it was only one of several strategies taught. Third, Swanson et al.'s (2010) training regimen, as described earlier, was only 15 minutes in duration. This is considerably briefer than other efficacy studies of WM training involving children and verbal WM tasks (cf. Wass, Scerif, & Johnson's [2012] training regimen of 400–600 minutes). One can legitimately speculate that relatively short training programs do not provide study participants—and especially children—the opportunity to become fluent in their application of rehearsal to verbal WM tasks. Thus, there is need for more research on children's use of rehearsal on verbal WM tasks with researchers using these tasks and academic tasks as outcomes.

Without strategy instruction—Many researchers who train WM without strategy instruction base their approach on Capacity Theory, which says WM is a "mental space" (Engle & Kane, 2004) that can be expanded. Accordingly, the purpose of the training is to increase the size of WM, rather than to improve its efficiency. In other words, WM is like a muscle and WM training is the equivalent of repeated exercise that increases the capacity, or the strength, of the muscle. Repeated exercise, it is believed, produces long-term plasticity in the brain regions serving WM, which should benefit any activity that calls on the same underlying brain networks (e,g., Dahlin, Bäckman, Stigsdotter Neely, & Nyberg, 2009).

However, as indicated, research on WM training without strategy instruction has produced inconsistent results. Few studies have reported transfer to performance on untrained WM tasks or on measures of academic skills (Shipstead et al., 2012). There is more to understand about whether and how the WM training of children with and without strategy instruction may strengthen their WM and academic performance.

Training Young At-Risk Children

The last issue we address is whether WM training is efficacious for young children at risk for learning problems. The vast majority of prior research has focused on adults or typically-developing children in the intermediate grades (Melby-Lervåg & Hulme, 2012). Few investigations have involved younger at-risk children (e.g., children in early elementary grades), especially those in which investigators have attempted to train young children on verbal WM tasks and to explore whether such experience affects their performance on academic measures.

We believe that training the verbal WM of young at-risk children may be important in two respects. First, from a cognitive-developmental perspective, it may be important because young children's functional neural networks are relatively plastic and the training may more likely produce desired effects (Shipstead et al., 2012; Wass et al., 2012). Consonant with this view is a review by Melby-Lervåg and Hulme (2012) who found that the visual-spatial WM training of preschoolers produced larger effects than those associated with similar training of intermediate-grade children. Second, children's WM deficits can contribute to learning problems (Swanson & Jerman, 2006; Swanson et al., 2009), and can predict children's later academic performance and disability status (Alloway, 2009). Training young children's WM may strengthen academic-related skills such as comprehension. Consistent with this possibility are findings from an investigation conducted by Savage, Lavers, & Pillay (2007) who found that WM was more strongly connected to comprehension than to word reading among young children.

Study Aims

The purposes of this first-grade study were to investigate whether training verbal WM improves verbal WM; whether this presumed effect transfers to performance on a passage listening comprehension measure; and whether training effects differ when strategy (i.e., rehearsal) instruction is part of the training. The children participating in WM training without strategy instructions were presented with complex verbal WM span tasks on which they simultaneously practiced processing and storing verbal information. The children in rehearsal training were taught to apply an explicit strategy to the very same tasks.

Prior to and immediately following the training, we administered several trained and untrained verbal WM tasks to determine whether the training improved this area of cognitive functioning. Because rehearsal is closely related to articulation rate and verbal short-term memory (Baddeley, 1986), and because children may depend on visual-spatial short-term memory during the training (e.g., using the relative position of stimuli on a board), we assessed articulation, verbal short-term memory, and visual-spatial short-term memory on a pre- and post-training basis to explore possible relations between these cognitive functions and WM training. Our hypotheses were that young at-risk children's verbal WM would be strengthened through training, which, in turn, would improve their passage listening comprehension. Capacity Theory suggests training with or without strategy instruction will lead to greater improvement on verbal WM and listening comprehension tasks. Strategy Mediation Theory suggests rehearsal may be necessary for improvements in both verbal WM and comprehension.

Finally, study participants were identified by their teachers and by us as at risk for learning problems (see the Participants section below). Our interest in WM training reflects the view that cognitive training may prove an important *supplement* (not *substitute* for) skills-based instruction, and as such may be an important means of intensifying and strengthening instruction for children with serious learning difficulties. Because we involved a relatively small sample of children in a modest number of training sessions, we regard this study and its results as heuristic—suggestive rather than conclusive.

Method

Participants

Participants were 58 children from 13 elementary schools in a mid-sized city in the Southeastern United States. They were originally part of a larger group of children who had been identified by their teachers in fall of first grade as at risk for learning difficulties and appropriate for a reading and math intervention study that we were conducting. We individually tested this larger group of teacher-nominated children with a battery of reading measures that included timed and untimed tests of rapid letter naming, phonemic decoding, and word recognition. A factor score was derived for each child based on their performance on these measure tasks, and the children were rank ordered by their factor scores. The top 50% were eliminated from study participation as were children who performed below at T-score of 37 on both the Vocabulary and Matrix Reasoning subtests of the Wechsler Abbreviated Scale of Intelligence.

The remaining children, still in fall of first grade, were randomly assigned to two treatment groups or a control group in an intervention study to improve their reading and mathematics skills. In early spring, following the children's 20-week participation in the intervention, we asked their parents to consent to their continued participation in this second study. As part of this study, the children were randomly assigned to three groups: WM training with and without strategy instruction or no WM training (controls). These three study groups were comparable (ps > .15) in terms of *prior intervention status, demographics* (age, gender, race, and free/reduced lunch status), and *pre-training performance* (non-verbal IQ, listening comprehension, word reading, WM, articulation rate, short-term memory, passage retell, passage listening comprehension). There was a marginally significant group difference on pre-training listening comprehension, $F(2, 55) = 3.00$, $p = .06$, with the no-strategy-instruction group showing significantly lower performance than controls ($p = .02$). Table 1 provides demographic information and the children's non-verbal IQ, pre-training listening comprehension, and pre-training word reading performance. The data show that the sample's non-verbal IQ (47th percentile) and word reading (63rd percentile) were in the average range (following the 20-week skills-based intervention), but their listening comprehension (30th percentile) was below average.

WM Training

The children who were assigned randomly to the rehearsal training group and no-strategy-training group participated in 10 sessions, one per day, on 10 consecutive school days. Each session lasted 35 minutes. The training occurred in the children's schools in the quietest locations available. Twenty-two psychology and education master's students were deployed as research assistants (RAs). They were randomly assigned to train 3–4 children representing both training groups. All training sessions were one-on-one. That is, the RAs worked with only one child in each session. Written scripts guided the RAs' interactions with the children during the training.

Training without strategy instructions—In each session, children worked on four complex verbal WM span tasks. Each lasted 8 minutes. The four verbal WM tasks were Counting Figures, Calculation Span, Operation Span, and Puzzles. For the *Counting Figures* task, children were presented with a 4×4 grid on a piece of paper with two or three types of stimuli (e.g., shapes, cartoon characters, animals) in contrasting colors. There were 36 pages of these grids, each with different stimuli. For every trial (or attempt to recall), children were asked to count one stimulus (e.g., stars). They were then told to count a second stimulus (e.g., blue triangles). Finally, they were asked to recall the sums of the various stimuli in the order they were counted. Depending on the level of their performance, the children could be asked to count and recall three sums or more.

The *Calculation Span* task directed the children to solve several simple addition or subtraction problems presented on flash cards with answers less than 10 (e.g., 2+1, 9-0), and then to recall their answers to the problems in order of their presentation. If they had difficulty solving a problem, correct answers were given. Depending on their performance, they could be asked to recall two or more correct responses.

For *Operation Span*, children named several sets of cards in each trial. First, they were asked to solve a simple addition or subtraction problem (with answers less than 10) presented on a flash card. Then they had to name a picture card (e.g., tree). They were asked to recall in order all the picture cards at the end of each trial. If they had difficulty solving the math problem or naming a picture card, correct answers were given. Depending on the level of their performance, the children could be asked to recall two or more picture names.

In the *Puzzles* activity, children were read six clues (presented in simple sentences consisting of 5 words or less) about a person, place, or thing. They were then told to solve the puzzle and use the answer and one or more clues to make a sentence. For example, the RA read: "I have four legs. I have fur. I have a tail. I like to chase cats. I love to bark. I like to eat bones." The child was asked, "What am I?" After answering "dog," the child was told to use the answer and at least one clue to make a sentence like, "A dog has four legs." If the children had difficulty constructing a sentence, the RA provided help. If they forgot the clues, the RAs showed them how to make sentences with the clues they did not recall.

Training with rehearsal—The just-described four tasks were also used in rehearsal training. The main difference between the two training procedures was that the rehearsal group was explicitly taught a strategy and was encouraged to use it during each trial of every task. For Counting Figures, Calculation Span, and Counting Span, the procedure was the same: When the children first encountered numbers or words to be remembered, they were told to say them aloud, repeatedly, and as fast as possible for 3 seconds. As more stimuli were added in a trial, the children were told to say the new stimulus, as well as the previously named stimuli as fast as possible for 3 seconds (or three times if there were more than four stimuli to rehearse).To illustrate, for the Counting Figures task, if the child first counted 3 figures, s/he would say "3, 3, 3…." for 3 seconds. If s/he then counted 5 figures, s/he would say "3, 5, 3, 5, 3, 5…" for 3 seconds. When children forgot to rehearse, or rehearsed incorrectly, the RA corrected them.

On the Puzzles task, the children were read each clue. They were then asked to identify its key word. In the aforementioned dog puzzle, the clue was, "I love to bark." The key word was "bark." If the children failed to identify it, the RA provided it. Each time the children identified a key word, they were required to say it aloud together with other key words previously identified. After solving the puzzle, the children were told to use the answer and at least one clue to make a sentence. If they had difficulty constructing a sentence, the RA provided help. If they forgot the clues, the RAs showed them how to make sentences with the clues they had not recalled.

For both training groups, and three activities (Counting Figures, Calculation Span, and Counting Span), the WM training was *adaptive*. Task difficulty was matched to the children's memory span performance on a trial-by-trial basis. For example, in Counting Figures, if a child remembered three sums in correct order, she was asked to remember four sums in the next trial. If she could not remember three sums in order, she was given another try in the following trial. If the child could not remember three sums in two consecutive trials, she was asked to remember two sums in the next trial. The children were encouraged to solve puzzles and recall as many clues as possible in 8 minutes in each session. Points and small prizes (e.g., cartoon stickers) were used to keep the children engaged.

Documentation of WM training—For both training groups, the RAs documented the children's performance on each trial of every task on a log form. This form was completed for all sessions. Specifically, for each trial in Counting Figures, Calculation Span, and Operation Span, the RAs recorded the span level (the number of target words/numbers) on which the children worked; whether they succeeded at this level (correctly recalled all the target words/numbers in order); and the kinds of strategies they used. For each trial of the Puzzles activity, the RAs recorded the number of clues the children recalled independently or used in a sentence, and the strategies they used for remembering them. Whereas the children in the no-strategy-instruction group were not instructed to use strategies, neither were they discouraged from using them. Their strategy use was also documented by RAs for each trial of each task across the 10 sessions.

Fidelity of Implementation

The first author attempted to ensure training fidelity in three ways. First, he conducted a two-day workshop, after which each RA met with him to role-play a training session (with the first author as the child) using a standard protocol. The RAs were required to achieve a fidelity score of 90% or greater on an implementation checklist before they began working with the children. Second, the first author observed each RA during one training session and provided corrective feedback immediately afterward. Third, he met twice with the RAs as a group during the 10-day training period to review training procedures, answer questions from the RAs, and provide support. All training sessions were audiotaped. The first author listened to the complete audio file of one session per child to document average fidelity across all of them for the two training groups. Fidelity was determined to be 96% (*SD* =

3.30%) and 98% (*SD* = 1.57%) for the no-strategy-instruction and the rehearsal groups, respectively. An RA listened to 20% of the audio files and inter-rater agreement between the RA and first author was 82%.

Measures of Children's IQ and Academic Performance

Non-verbal IQ—WASI Matrix Reasoning (Wechsler, 1999) is a measure of non-verbal IQ. The child looks at a matrix from which a section is missing and completes it by selecting among five options. The score is the total number of matrices answered correctly. Wechsler (1999) reported a test-retest reliability coefficient of .90 for 6- and 7-year-olds.

Listening comprehension—We used the Woodcock–Johnson Oral Comprehension subtest (Woodcock, McGrew, & Mather, 2001), for which the child listens to short sentences or short passages and provides a missing word. The score is the number of items answered correctly. The test-retest reliability coefficient has been reported as .80 for 6- and 7-year-olds (Woodcock et al., 2001).

Word reading—The Word Identification Subtest of the Woodcock Reading Mastery Test-Revised (Woodcock et al., 2001) asks the child to read 100 single words ordered in difficulty. The score is the number of words read correctly. Test-retest reliability for 6- and 7-year-olds has been reported to be .90 (Woodcock et al., 2001).

Pre-and Post-Training Measures

Counting recall—This task is an adaptation of the Counting Recall activity from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). It requires the child to count piles of dots; to remember these sums; and to later recall the sums in sequence. There are six trials at each set size (2 to 7 piles of dots per set). The score is the number of trials recalled correctly. Cronbach's alpha for the sample was .84.

Listening recall—This task is an adaptation of the Listening Recall activity from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). The child listens to a series of short sentences, judges the veracity of each sentence by responding "yes" or "no," and then recalls the final word of each of the sentences in sequence. There are six trials at each set size (1 to 6 sentences per set). The score is the number of trials recalled correctly. Cronbach's alpha for the sample was .78.

Digit recall—This task, too, is adapted from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001).The tester orally presents digits ranging from 1 to 9 at the rate of one digit per second. The child is asked to recall the digits in correct serial order. There are six trials at each set size, which range from 1 to 9 digits. The score is the number of trials recalled correctly. Cronbach's alpha for the sample was .82.

Block recall—This task is from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). We did not modify it. The child views nine cubes placed randomly on a board. The tester taps the blocks in a predetermined sequence, and the child is told to tap the same sequence. The score is the number of trials recalled (tapped) correctly. Cronbach's alpha for the sample was .84.

Articulation rate—This task was administered to assess speed of speech. It is adapted from a task developed by Kail (1997). The child repeats a pair of single-syllable words and digits as quickly as possible in 5 seconds. There are three trials of word pairs (fish-pig, book-set, car-spoon) and three trials of number pairs (2–5, 9-3, 1–8). There are two scores. One is the articulation-word rate score: The average number of word pairs the child says correctly in 5 seconds. The second is the articulation-number rate score: The average amount of number pairs the child says correctly in 5 seconds. Mean Cronbach's alpha for the sample for both the articulation-word and articulation-number scores was .84.

Passage listening comprehension—This test is part of the Qualitative Reading Inventory (QRI; Leslie & Caldwell, 2001).The tester reads aloud a story of about 250 words. The child retells as much of the story as possible and answers six open-ended comprehension questions. Two equally difficult-to-read stories at the first-grade level were administered at pre- and post-training, respectively. For each story, there are two scores. One is the QRI-Retell score, which reflects the number of things the child recalls about the story. The second is the QRI-Passage Listening Comprehension score, or the number of comprehension questions answered correctly. One RA scored each child's retell performance. Another RA independently repeated the scoring for 20% of the sample. Both were "blind" to the study's purposes and to membership in study groups. Inter-rater agreement on the retell score was 99%. The mean split-half reliability coefficient for comprehension questions of the two stories for the sample was .60.

Data Collection and Analyses

Data collection—The 22 RAs conducted all testing. The tests were administered to children individually in the quietest place available at their schools. The RAs were randomly assigned to test children, except that they were not allowed to conduct post-treatment testing of those whom they tutored. Thus, the RAs were blind to the children's group membership at both pre-and post-training testing. The children were tested prior to the training in one session and immediately following the training in another session. Each of the two sessions lasted 60 minutes.

Two project staff (including the first author) trained the RAs in multiple sessions during which different tests were introduced. Each training session began with project staff explaining the purpose and design of the tests, and then modeling their proper administration. The RAs next role-played as examiner and examinee, and obtained immediate corrective feedback from staff. Following this training, the RAs were required to find a partner and practice test administration for 5 hours prior to pre-treatment testing. Two days after training, each RA "tested" the project staff on all measures. Staff recorded RA performances on detailed checklists for each test. The RAs were required to achieve at least 90% accuracy when administering and scoring every test. If they performed below 90% on one or more test, they were required to complete additional training and try again to meet administration and scoring criteria. The RAs were not permitted to test children before they did so. Moreover, all testing sessions were audiotaped, and 31% of the audio files were used to calculate the fidelity with which the tests were administered. Averaged across all RAs and tests, fidelity was determined to be 99%.

Data analyses—We first plotted the children's performance across the 10-session training on each WM task, and we summarized their strategy use. Because we drew our sample from 13 schools, we calculated intra-class correlation coefficients to evaluate school effects on each post-training measure. To account for medium or large school variance, we used multilevel modeling (Raudenbush & Bryk, 2002) with Level-1 representing a child level and Level-2 a school level. Then, we used hierarchical regression-based analysis to examine the treatment effects (i.e. no-strategy-instruction vs. control; rehearsal vs. control; no-strategy-instruction vs. rehearsal) on outcome measures, controlling for pretreatment performance on the same measures. Moreover, we examined whether children's pre-training non-verbal IQ, reading skills, and WM performance moderated the training effects. We describe these data-analytic steps in greater detail below.

Results

Performance During Training and Strategy Use

Based on information from the training logs, we plotted children's performance across the 10 sessions on each WM training task, and we summarized their strategy use. Figure 1 displays the rehearsal group's and no-strategy-instruction group's improvement in terms of the highest span achieved (i.e., the highest number of words/numbers/clues recalled correctly) on each of the four verbal WM tasks (Counting Figures, Calculation Span, Operation Span, and Puzzles). The rehearsal group demonstrated a statistically significant improvement rate (or slope) on each task across 10 sessions, slope = .15~.46, $ps < .01$. Their most impressive rate of improvement was on Counting Figures, slope =.46, $p < .001$; their least impressive rate of improvement was on Puzzles, slope = .15, $p = .001$. The no-strategy-instruction group showed statistically significant improvement on Calculation Span, slope = .11, $p = .02$, and Puzzles, slope = .10, $p = .02$. The rehearsal group exhibited a reliably greater improvement rate than the no-strategy-instruction group on all of the WM training tasks, $F(1, 36) = 19.63 \sim 36.25$, $ps < .001$, except on Puzzles, $F(1, 36) = .98$, $p = .33$.

On average, the children in the rehearsal group used strategies during 99% of all training trials. Among these trials, 89% (of the 99%) involved rehearsal, 5% reflected a counting strategy (i.e., children used their fingers to track the number of words/numbers), another 5% showed evidence of a visual strategy (i.e., children memorized/pointed to the position of the word/number flash cards), 1% involved a semantic strategy (i.e., children put the to-be-remembered words/numbers in a sentence), and 0.2% indicated use of other strategies (e.g., children chunked words/numbers).

Because we did not prevent the children in the no-strategy-instruction group from making use of strategies, we observed them on average using strategies in 28% of all training trials. Among these, 59% (of the 28%) involved rehearsal, 32% showed evidence of a counting strategy, 6% included a visual strategy, 3% involved a semantic strategy, and 0.4% reflected use of other strategies. Therefore, the no-strategy-instruction group's average use of rehearsal across trials was 17% (59% of 28%). The corresponding percentage for the rehearsal group was 88% (89% of 99%). In other words, although some in the no-strategy-instruction group used rehearsal, they did so considerably less frequently than the rehearsal group.

Training Effects on Working Memory and Comprehension

Preliminary analyses—We first explored distributions of performance on each measure (e.g., SD, skewness, kurtosis). Generally, performance was normally distributed at pre- and post-training (see Table 2). Because we drew our sample from 13 elementary schools, we calculated intra-class correlation coefficients (ICCs) to evaluate school effects on each post-training measure. Schools explained a small-to-large proportion of the variance (ICCs = 0.1% ~ 27%). To account for this variance, we used multilevel modeling (Raudenbush & Bryk, 2002) with Level 1 and Level 2 indicating child and school levels, respectively.

We created two sets of dummy variables to examine three group comparisons: (a) rehearsal vs. control, (b) no-strategy-instruction vs. control, and (c) rehearsal vs. no-strategy-instruction (see Stanovich & Siegel, 1994, for a rationale). The first set included dummy variables that compared the rehearsal group to controls (rehearsal = 1; no-strategy-instruction = 0; control = 0), and the no-strategy-instruction group to controls (rehearsal = 0; no-strategy-instruction = 1; control = 0) (Cohen et al., 2003). The second set subsumed two more dummy variables; one comparing the rehearsal group to the no-strategy-instruction group (rehearsal = 1; no-strategy-instruction = 0; control = 0); the second comparing controls to the no-strategy-instruction group (rehearsal = 0; no-strategy-instruction = 0; control = 1) (Cohen et al., 2003).

We compared the two training groups on the post-training measures of WM, short-term memory, articulation rate, QRI-Retell, and QRI-Passage Listening Comprehension, controlling for pre-training performance on each of them. There were marginally statistically significant between-group differences on listening comprehension at pre-training ($p = .06$), so we also controlled pre-training listening comprehension for group comparisons on post-training QRI-Retell and QRI-Passage Listening Comprehension.

HLM Version 7.0 (Raudenbush & Bryk, 2002) was used to fit the two-level models. Dichotomous variables (i.e., group comparisons) were entered uncentered; continuous variables were grand-mean centered. We chose grand-mean (rather than group-mean) centering because we were interested in how schools influenced individual students relative to the average student, rather than to the school average. All variables at the student level were tested for randomly varying slopes (Raudenbush & Bryk, 2002), but few were associated with significant variability at the school level. Thus, we only had intercepts vary at the school level in each model. In addition to presenting unstandardized coefficients, standard errors, and *p*-values in the HLM models, we present Hedges g to indicate training effects (See Table 3). We chose Hedges g because it provides a better estimate of effect size than Cohen’s d with small sample sizes (Grissom & Kim, 2005). As suggested by What Works Clearinghouse (2008), the formula to calculate Hedges g in HLM is as follows,

$$\mathrm{Hedges g} = \frac{\gamma}{\sqrt{\frac{(N_1-1)\mathrm{SD}_1^2+(N_2-1)\mathrm{SD}_2^2}{(N_1+N_2-2)}}}$$

Where **γ** is the HLM coefficient for the intervention's effect, which represents the group mean difference adjusted for both level-1 and level-2 covariates; N_1 and N_2 are the student sample sizes, and SD_1 and SD_2 the unadjusted student-level SDs for the intervention group and the comparison group, respectively.

Table 3 shows the fixed and random effects of group comparisons between the two training groups and controls on each outcome (See Appendix A, B, and C for HLM full models). In the following, we first present *p*-values *uncorrected* for multiple comparisons. We present them for heuristic purposes only. We then provide *corrected p*-values.

Uncorrected *p* values—The rehearsal group outperformed controls on an untrained verbal WM task (i.e., Listening Recall), Hedges g = .47, uncorrected $p = .03$; on QRI-Retell, Hedges g = .65, uncorrected $p = .04$; and on QRI-Passage Listening Comprehension, Hedges g = .63, uncorrected $p = .03$. The rehearsal group also (marginally) outperformed controls on the articulation-number rate measure, Hedges g = .45, uncorrected $p = .06$. The no-strategy-instruction group's performance was stronger than controls on QRI-Passage Listening Comprehension, Hedges g = .65, uncorrected $p = .02$, but not on any other measure. Although there were no statistically significant differences between the two training groups, the rehearsal group's performance on the untrained verbal WM task (i.e., Listening Recall) and QRI-Retell were associated with moderate effect sizes favoring the rehearsal group (Hedges g = .33 and .43, respectively; see Table 3).

Corrected *p* values—We corrected our *p* values using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) for group comparisons on each outcome. As reported in Table 3, after these *p*-value adjustments, all statistically significant group differences disappeared. Because our study participants varied with respect to their pre-training performance on measures of non-verbal IQ and academic and WM skills, we examined whether their performance on these measures moderated training effects. It did not.

Discussion

Few studies of WM training have explored the consequences of verbal WM training on verbal WM tasks and passage listening comprehension measures; or enlisted young children at risk for learning difficulties as study participants; or compared different approaches to train verbal WM in the same study. In this randomized control trial, we did all three. We examined whether intensive (2-week, 10-session, 5.8 hours) training of verbal WM would strengthen verbal WM and passage listening comprehension in first-grade children, and we explored whether no-strategy training and rehearsal training exerted differential effects.

Moreover, we pursued these issues with as much experimental control and rigor as we could muster. As mentioned, we randomly assigned our young participants to the three study groups and documented their pre-training comparability on many demographic, cognitive, and academic measures. We promoted instructional intensity by requiring tutors to work with the children one at a time; by challenging the children who met one performance criterion to then meet a second, more rigorous criterion; and by equipping tutors with meaningful training and scripted lessons to promote fidelity of implementation. And, following the advice of Shipstead, Reddick, and Engle (2012), our pre- and post-training measures included those like counting recall that reflected the WM training and others like listening recall that did not.

Yet, our sample ($N = 58$) was small for three study groups, and our analyses were underpowered. This contributed to difficult-to-interpret results, which we presented two ways: with p values *uncorrected* and *corrected* for multiple group comparisons. Whereas some will view this as unconventional, we have presented our findings this way in hopes that readers will see them as heuristic.

With an uncorrected p value set at .05, both training groups showed improvements on WM training tasks. The rehearsal group also strengthened its performance in contrast to controls on an untrained verbal WM task (i.e., Listening Recall; Hedges g =.47). For the no-strategy-instruction and rehearsal groups, training effects appeared to transfer to one or more measure of listening comprehension. Rehearsal training seemed to strengthen children's QRI-Retell (Hedges g = .65) and QRI-Passage Listening Comprehension (Hedges g = .63). Training without strategy instructions appeared to improve group members' QRI-Passage Listening Comprehension (Hedges g = .65). The superior performances of the two training groups versus controls on WM tasks and listening comprehension measures are not likely attributable to articulation speed or verbal short-term memory because the training groups did not show significant gains in these functions when compared to controls. Although there were no statistically significant differences between the two training groups (with or without adjusted p values) on any WM or comprehension measure, there were moderately large effect size differences favoring the rehearsal group on an untrained verbal WM task (i.e., Listening Recall) and QRI-Retell (see Table 3). When we used an adjusted p value to control for multiple group comparisons, all of the statistically significant between-group differences disappeared.

Caveats and Admonitions

Whereas we are suggesting that training verbal WM— especially with rehearsal — strengthens young children's verbal WM and passage listening comprehension, this suggestion is encumbered by more than our underpowered analyses. It should also be seen through a prism of caveats of a more substantive nature. These caveats are of two kinds. The first is a set of study limitations, the most important of which may be that we did not explore possible changes in children's attention that might have mediated group differences on verbal WM and comprehension. Children in the two training groups were frequently required to work with long lists of words and numbers. Research suggests that as one increases the load on short-term memory one is also requiring greater amounts of attention (e.g., Unsworth & Engle, 2007). Thus, our WM training may have simultaneously and inadvertently involved attention training; and stronger attention is a reasonable and competing explanation for the training groups' (presumably) superior performance versus controls.

Similarly, because children in the rehearsal group were likely to encounter longer lists of words and numbers than those in the no-strategy-instruction group, we may have been strengthening the rehearsal group's attention relative to the no-strategy-instruction children as well. Because attention is closely related to WM and comprehension, and attention training can improve reading comprehension (Solan, Shelley-Tremblay, Ficarra, Silverman, & Larson, 2003), future studies of WM training might attempt to parse WM by distinguishing it from attention. Such an effort, in principle, may eventually strengthen the WM construct as well as underscore the potential importance of related cognitive abilities, like attention, that are too infrequently recognized as related to WM.

A second study limitation is that the amount of WM training (one 35-minute session per day on 10 consecutive school days) may have been insufficient for children in the no-strategy-instruction group. In comparison to the rehearsal group, and using an uncorrected *p* value, these children demonstrated smaller but statistically significant improvement on WM training tasks (i.e., Calculation Span and Puzzles). Also, although the no-strategy-instruction group did not show statistically significant post-training improvement on untrained verbal WM tasks (i.e., Listening Recall and Counting Recall), the effect sizes (Hedges g = .17 ~. 36) reflected small-to-moderate improvement in comparison to controls. Thus, training verbal WM without strategy instructions may prove effective for improving young children's verbal WM if the training program is of longer duration.

Some may view our strategy instruction as a third limitation. Shipstead et al. (2012), for example, wrote that WM training "should not teach specific strategies for simply remembering more information (e.g., rehearsal techniques or mnemonic devices). Strategies might improve a person's score on a WM test; however, this is not the same as changing the underlying ability" (p. 5). Unlike St. Clair-Thompson et al.'s (2010) strategy training, which required children to apply rehearsal to short-term memory tasks, our study participants were expected to rehearse and process verbal information at the same time. That is, while members of the rehearsal group were trained to use a strategy to remember verbal information, they also had to process verbal information as they were using it.

Implications for Theory

Caveats notwithstanding, we believe our results are consistent with Strategy Mediation Theory. Many study participants became fluent in rehearsing important information while simultaneously managing distractions like calculation and counting. Because rehearsal appears to be an effective strategy for performing well on verbal WM tasks and comprehension tasks (e.g., Gersten et al., 2001; Rose et al., 1983; Turley-Ames & Whitfield, 2003)--perhaps because both draw on similar cognitive processes (Cain et al., 2004)-- our results suggest (however preliminarily) that the children who rehearse information fluently as they listen to (or read) a passage, may be more likely to remember it and integrate it with previous information, producing stronger comprehension. In short, study participants' rehearsal training may have improved the efficiency with which they used their WM—on both the verbal WM tasks and listening comprehension tasks.

We remind readers in this regard that we observed children in the no-strategy-instruction group also using strategies—for 28% of total trials, 59% of which involved rehearsal. That is, at least some young children use rehearsal during verbal WM training without explicit instruction to do so, presumably because it comes "naturally" to them and is less cognitively demanding than other strategies (Turley-Ames & Whitfield, 2003). Dunning and Holmes (2014) also observed the voluntary and spontaneous use of strategies among young adults involved with WM training. This, together with our observations on strategy use in the no-strategy-instruction group and the rehearsal group's relatively strong showing, suggests that at least some children and young adults perceive strategy use as sensible and helpful, and combining rehearsal with verbal WM tasks may lead to better comprehension among young children.

A second implication for theory building connects to the long-running discussion about whether WM is better understood as domain-general or domain-specific. Previous reviews of WM training indicate that attempts to strengthen visual-spatial WM have had little or no effect on improving children's verbal WM or their verbal-related academic performance (Shipstead et al., 2012). Together with our results, these prior reviews suggest that the efficacy of WM training may be influenced by the nature of the training task. Verbal WM training may be more fruitful than visual-spatial WM training when desired outcomes for children include verbal WM and comprehension. This is consistent with a domain-specific model of WM (Ericsson & Kintsch, 1995; Unsworth & Engle, 2007). Verbal WM training, therefore, should be more effective than visual-spatial WM training for improving performance on verbal WM tasks and verbal-related academic skills. Moving from expectations to fact will require more empirical work.

Future Research

We have already discussed a need for research to distinguish attention from WM. An equally important and difficult task for researchers can be understood in the context of responsiveness-to-intervention (RTI). Many view RTI as a fundamental reorganization of service delivery—a promising reconfiguring of general and special education into one unified set of multiple and increasingly intensive tiers of skills-based instruction (cf. D. Fuchs, Fuchs, & Stecker, 2010). *Skills-based instruction* refers to an attempt to strengthen academic skills (e.g., letter-sound correspondence and math problem solving) and to enhance knowledge in areas such as social studies and science.

A belief in the efficacy of skills-based instruction seems well founded. When implemented with fidelity, carefully scripted programs in reading, writing, and math have benefited numerous at-risk students (e.g., L. Fuchs, Fuchs, & Compton, 2012; Graham & Perin, 2007; Kroesbergen & VanLuit, 2003). Additionally, when researchers use a skills-based approach at Tier 1 or Tier 2 in an RTI framework, they often accelerate the academic progress of many children (Al Otaiba & Fuchs, 2006; McMaster, Fuchs, Fuchs, & Compton, 2005; Vaughn, Linan-Thompson, & Hickman, 2003), and decrease the likelihood that they will be wrongly identified as requiring special education.

As importantly, however, a skills-based approach fails to advance the progress of all students. Multiple research teams grappling with school-based implementations of RTI have independently demonstrated the veracity of this claim (e.g., D. Fuchs, Fuchs, & Compton, 2004; Vaughn et al., 2010). Extrapolating from their respective study samples, researchers have estimated that from 2% to 6% of the general population will not benefit from a skills-based approach when implemented by researchers (rather than by practitioners), suggesting these percentages are a conservative estimate. Thus, research (and common sense) promotes a view that if a child has not responded sufficiently to skills-based instruction at Tier 1, nor to a more intensive version at Tier 2, it makes little sense to “triple down” on the same approach at Tier 3. This raises the important question: If not a skills-based approach, then what?

Cognitively focused instruction is a well-known alternative (cf. Learning Disabilities Association, 2010). Arguably, the most popular variant targets the putative cognitive processes responsible for academic problems. Low-achieving students with working memory difficulties, for example, are trained to become more proficient on working memory tasks with the expectation that this increased proficiency will lead to stronger academic achievement (e.g., Holmes, Gathercole, & Dunning, 2009).

But there are alternate ways of thinking about skills-based and cognitive focused approaches. For example, they needn’t be mutually exclusive. That is, there can be a variety of combined, mixed, or hybrid approaches. One such approach incorporates task-relevant cognitive processes so that they are not taught in isolation. Self-regulated strategy development (SRSD), for example, is a skills-based writing intervention that requires students to use meta-cognition during their writing (e.g., Graham & Harris, 1989). Another example is mnemonics instruction, which aims to improve both meta-memory and academic skills (e.g., Mastropieri & Scruggs, 1998).

A second approach calls for accommodating students’ cognitive deficits by modifying the learning context. Montgomery (2004), for example, showed that by slowing by 25% the rate at which speech was directed at students with speech and language disorders, the children could comprehend at the same level as syntax-matched, typically-achieving peers. A third approach explores whether cognitive characteristics moderate instruction such that students with cognitive characteristic A improve more than students with characteristic B in the same skills-based program; or, whether students with cognitive characteristic A generally outperform those with characteristic B in one academic program while the reverse is obtained in a second academic program. Might cognitive characteristics, in short, cause differential responses to the same or different instructional programs? More generally, do cognitive attributes interact with features of instruction? The general point here is that cognitive moderators may be potentially important not because they can become targets of remediation but because they may suggest ways to tailor instruction for those not benefitting from it in its current form, Implicit is the suggestion that skills-based and cognitively focused approaches are not mutually exclusive. Researchers and practitioners may be able to use both to develop more effective programs for a greater number of children with serious learning problems.

Acknowledgments

This research was supported by R01 HD056109 and Core Grant #HD15052 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to Vanderbilt University and by R324A090052 from the U.S. Department of Education's Institute on Education Sciences to Vanderbilt University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. We wish to thank Lynn Fuchs for her comments on an earlier draft, and Sam Patton, Wenjuan Zhang, and especially Loulee Yen for their logistical support that made the research possible.

References

Al Otaiba S, Fuchs D. Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. Journal of Learning Disabilities. 2006; 39:414–431. [PubMed: 17004674]

Alloway TP. Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. European Journal of Psychological Assessment. 2009; 25:92–98.

Baddeley, AD. Working memory. New York: Oxford Univ. Press; 1986.

Bailey H, Dunlosky J, Kane MJ. Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. Memory & Cognition. 2008; 36:1383–1390. [PubMed: 19015498]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. 1995; 57:289–300.

Cain K, Oakhill J, Bryant PE. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. Journal of Educational Psychology. 2004; 96:31–42.

Dahlin KIE. Effects of working memory training on reading in children with special needs. Reading and Writing. 2011; 24:479–491.

Dahlin E, Bäckman L, Stigsdotter Neely A, Nyberg L. Training of the executive component of working memory: Subcortical areas mediate transfer effects. Restorative Neurology and Neuroscience. 2009; 27(5):405–419. [PubMed: 19847067]

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19:450–466.

Dunlosky J, Kane MJ. The contributions of strategy use to working memory span: A comparison of strategy assessment methods. Quarterly Journal of Experimental Psychology. 2007; 60:1227–1245.

Dunning DL, Holmes J. Does working memory training promote the use of strategies on untrained working memory tasks? Memory & Cognition. 2014; 42:1–9. [PubMed: 23784742]

Engle RW. Working memory capacity as executive attention. Current Directions in Psychological Science. 2002; 11:19–23.

Engle, RW.; Kane, MJ. Executive attention, working memory capacity, and a two-factor theory of cognitive control. In: Ross, B., editor. The psychology of learning and motivation. New York: Academic Press; 2004. p. 145-199.

Engle R, Marshall K. Do developmental changes in digit span result from acquisition strategies? Journal of Experimental Child Psychology. 1983; 36:429–436.

Ericsson KA, Kintsch W. Long-term working memory. Psychological Review. 1995; 102:211–245. [PubMed: 7740089]

Friedman NP, Miyake A. The Reading span test and its predictive power for reading comprehension ability. Journal of Memory and Language. 2004; 51:136–158.

Fuchs D, Compton DL, Fuchs LS, Hamlett CL, Lambert W. First-grade cognitive abilities as long-term predictors of reading comprehension and disability status. Journal of Learning Disabilities. 2012; 45:217–231. [PubMed: 22539057]

Fuchs D, Fuchs LS, Compton DL. Identifying reading disability by responsiveness-to-instruction: Specifying measures and criteria. Learning Disability Quarterly. 2004; 27:216–227.

Fuchs D, Fuchs LS, Stecker PM. The "blurring" of special education in a new continuum of general education placements and services. Exceptional Children. 2010; 76(3):301–323.

Fuchs LS, Compton DL, Fuchs D, Paulsen K, Bryant JD, Hamlett CL. The prevention, identification, and cognitive determinants of math difficulty. Journal of Educational Psychology. 2005; 97:493–513.

Fuchs LS, Fuchs D, Compton DL. The early prevention of mathematics difficulty: Its power and limitations. Journal of Learning Disabilities. 2012; 45:257–269. [PubMed: 22491809]

Gersten R, Fuchs LS, Williams JP, Baker S. Teaching reading comprehension strategies to students with learning disabilities: A review of research. Review of Educational Research. 2001; 71:279–320.

Graham S, Harris KR. Improving learning disabled students' skills at composing essays: Self-instructional strategy training. Exceptional Children. 1989; 56:201–214. [PubMed: 2806360]

Graham S, Perin D. A meta-analysis of writing instruction for adolescent students. Journal of Educational Psychology. 2007; 99:445–476.

Grissom, RJ.; Kim, JJ. Effect sizes for research: A broad practical approach. Mahwah, NJ: Erlbaum; 2005.

Hagen JW, Jongeward RH, Kail RV. Cognitive perspectives on the development of memory. Advances in Child Development and Behavior. 1975; 10:57–101. [PubMed: 1101662]

Holmes J, Gathercole SE, Dunning DL. Adaptive training leads to sustained enhancement of poor working memory in children. Developmental Science. 2009; 12(4):F9–F15. [PubMed: 19635074]

Kail R. Phonological skill and articulation time independently contribute to the development of memory span. Journal of Experimental Child Psychology. 1997; 67(1):57–68. [PubMed: 9344487]

Klingberg T. Training and plasticity of working memory. Trends in Cognitive Sciences. 2010; 14:317–324. [PubMed: 20630350]

Kroesbergen E, Van Luit J. Mathematics interventions for children with special educational needs: A meta-analysis. Remedial and Special Education. 2003; 24:97–114.

Kroesbergen EH, Van't Noordende JE, Kolkman ME. Training working memory in kindergarten children: Effects on working memory and early numeracy. Child Neuropsychology. 2014; 20:23–37. [PubMed: 23098260]

Learning Disabilities Association. The Learning Disabilities Association of America's white paper on evaluation, identification, and eligibility criteria for students with specific learning disabilities. Pittsburgh, PA: Author; 2010.

Leslie, L.; Caldwell, J. Qualitative reading inventory −3. New York: Addison Wesley Longman; 2001.

Mastropieri MA, Scruggs TE. Enhancing school success with mnemonic strategies. Intervention in School & Clinic. 1998; 33:201–208.

Melby-Lervåg M, Hulme C. Is working memory training effective? A meta-analytic review. Developmental Psychology. 2012; 49:270–291. [PubMed: 22612437]

McMaster KN, Fuchs D, Fuchs LS, Compton DL. Responding to nonresponders: An experimental field trial of identification and intervention methods. Exceptional Children. 2005; 71:445–463.

McNamara DS, Scott JL. Working memory capacity and strategy use. Memory & Cognition. 2001; 29:10–17. [PubMed: 11277453]

Montgomery JW. Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. International Journal of Language and Communication Disorders. 2004; 39:115–133. [PubMed: 14660189]

Nation K, Adams JW, Bowyer-Crane CA, Snowling MJ. Working memory deficits in poor comprehenders reflect underlying language impairments. Journal of Experimental Child Psychology. 1999; 73:139–158. [PubMed: 10328862]

Pickering, S.; Gathercole, SE. Working Memory Test Battery for Children for children (WMTB-C). Psychological Corporation; 2001.

Raudenbush, SW.; Bryk, AS. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage; 2002.

Rose MC, Cundick BP, Higbee KL. Verbal rehearsal and visual imagery: Mnemonic aids for learning disabled children. Journal of Learning Disabilities. 1983; 16:352–354. [PubMed: 6886558]

Savage R, Lavers N, Pillay V. Working memory and reading difficulties: What we know and what we don't know about the relationship. Educational Psychology Review. 2007; 19:185–221.

Seigneuric A, Ehrlich M-F, Oakhill JV, Yuill NM. Working memory resources and children's reading comprehension. Reading and Writing. 2000; 13:81–103.

Shah P, Miyake A. The separability of working memory resources for spatial thinking and language processing: An individual differences approach. Journal of Experimental Psychology: General. 1996; 125:4–27. [PubMed: 8851737]

Shipstead Z, Redick TS, Engle RW. Is Working Memory Training Effective? Psychological Bulletin. 2012; 138:628–654. [PubMed: 22409508]

Solan HA, Shelley-Tremblay J, Ficarra A, Silverman M, Larson S. Effect of attention therapy on reading comprehension. Journal of Learning Disabilities. 2003; 36:556–563. [PubMed: 15493437]

Stanovich KE, Siegel LS. Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core difference model. Journal of Educational Psychology. 1994; 86:24–53.

St. Clair-Thompson HL, Stevens R, Hunt A, Bolder E. Improving children's working memory and classroom performance. Educational Psychology. 2010; 30:203–219.

Swanson HL, Kehler P, Jerman O. Working memory, strategy knowledge, and strategy instruction in children with reading disabilities. Journal of Learning Disabilities. 2010; 43(1):24–47. [PubMed: 19749089]

Swanson HL, Zheng X, Jerman O. Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. Journal of Learning Disabilities. 2009; 42:260–287. [PubMed: 19255286]

Turley-Ames KJ, Whitfield MM. Strategy training and working memory task performance. Journal of Memory and Language. 2003; 49(4):446–468.

Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. Psychological Review. 2007; 114:104–132. [PubMed: 17227183]

Vaughn S, Cirino PT, Wanzek J, Wexler J, Fletcher JM, Denton CD, Francis DJ. Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. School Psychology Review. 2010; 39:3–21. [PubMed: 21479079]

Vaughn S, Linan-Thompson S, Hickman P. Response to instruction as a means of identifying students with reading/learning disabilities. Exceptional Children. 2003; 69:391–409.

Wass SV, Scerif G, Johnson MH. Training attentional control and working memory - Is younger, better? Developmental Review. 2012; 32:360–387.

Wechsler, D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Psychological Corporation; 1999.

What Works Clearinghouse. What Works Clearinghouse evidence standards for reviewing studies, version 1.0. Washington, DC: US Department of Education; 2008.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III. Itasca, IL: Riverside Publishing; 2001.

Appendix A

Fixed and Random Effects for Group Comparisons on WM and Short-Term Memory Tasks (HLM Full Models)

	Listening Recall					Counting Recall					Digit Recall					Block Recall			
Fixed Effects	Coefficient	SE	T-ratio	*p*	Fixed Effects	Coefficient	SE	T-ratio	*p*	Fixed Effects	Coefficient	SE	T-ratio	*p*	Fixed Effects	Coefficient	SE	T-ratio	*p*
Model 1a					Model 2a					Model 3a					Model 4a				
Intercept	7.39	0.56	13.15	0.00	Intercept	13.97	0.94	14.81	0.00	Intercept	23.73	0.56	42.34	0.00	Intercept	18.93	0.76	24.90	0.00
Strategy vs. Control	1.50	0.67	2.25	0.03	Strategy vs. Control	0.95	1.25	0.76	0 .46	Strategy vs. Control	−0.82	0.76	− 1.07	0.29	Strategy vs. Control	−0.01	1.08	−0.01	0.99
No - strategy vs. Control	0.49	0.67	0.74	0.47	No - strategy vs. Control	1.76	1.26	1.40	0 .17	No- strategy vs. Control	−0.57	0.77	− 0.74	0.46	N o- strategy vs. Control	0.79	1.10	0.71	0.48
Pre-Listening Recall	0.57	0.09	6.25	0.00	Pre-Counting Recall	0.29	0.12	2.31	0.03	Pre-Digit Recall	0.72	0.08	8.71	0.00	Pre-Block Recall	0.79	0.11	7.47	0.00
Model 1b					Model 2b					Model 3b					Model 4b				
Intercept	7.88	0.56	13.98	0.00	Intercept	15.73	0.96	16.45	0.00	Intercept	23.16	0.57	40.49	0.00	Intercept	19.72	0.78	25.36	0.00
Strategy vs. No-strategy	1.01	0.67	1.52	0.14	Strategy vs. No-strategy	−0.81	1.26	−0.64	0.52	Strategy vs. No-strategy	−0.25	0.77	−0.32	0.75	Strategy vs. No-strategy	−0.79	1.09	−0.73	0.47
Control vs. No-strategy	−0.49	0.67	−0.74	0.47	Control vs. No-strategy	−1.76	1.26	−1.40	0.17	Control vs. No-strategy	0.57	0.77	0.74	0.46	Control vs. No-strategy	−0.78	1.10	−0.71	0.48
Pre-Listening Recall	0.57	0.09	6.25	0.00	Pre-Counting Recall	0.29	0.12	2.31	0.03	Pre-Digit Recall	0.72	0.08	8.71	0.00	Pre-Block Recall	0.79	0.11	7.47	0.00
Random Effects	School level variance	Student level residual variance	χ^2	BIC	Random Effects	School level variance	Student level residual variance	χ^2	BIC	Random Effects	School level variance	Student level residual variance	χ^2	BIC	Random Effects	School level variance	Student level residual variance	χ^2	BIC
	1.15	4.15	26.14*	262.48		1.34	14.99	17.24	327.21		0.33	5.60	13.31	272.66		0.01	11.17	10.09	307.52

Note. Coefficient = Unstandardized coefficient in HLM models, the mean difference between study groups adjusted for level-1 covariates; SE = Standard Errors of the coefficient, BIC = Bayesian Information Criterion. Models 1a~4b represent the comparison of "Strategy vs. Control" and "No-strategy-instruction vs. Control". Because of Dummy coding, the Random Effects are the same for Model 1a and 1b, Model 2a and 2b, Model 3a and 3b, and Model 4a and 4b.

All coefficients in HLM models became non-significant after application of the Benjamini–Hochberg correction of *p* values for multiple comparisons (i.e., $p < .01$).

* $p < .05$,

\+ $p < .08$.

Appendix B

Fixed and Random Effects for Group Comparisons on Articulation Measures (HLM Full Models)

	Articulation-word rate					Articulation-number rate			
Fixed Effects	**Coefficient**	**SE**	**T-ratio**	***p***	**Fixed Effects**	**Coefficient**	**SE**	**T-ratio**	***p***
Model 5a					Model 6a				
Intercept	5.69	0.23	24.67	0.00	Intercept	7.50	0.25	30.49	0.00
Strategy vs. Control	0.21	0.29	0.72	0.47	Strategy vs. Control	0.63	0.33	1.91	0.06
No-strategy vs. Control	0.13	0.29	0.46	0.65	No-strategy vs. Control	0.03	0.33	0.08	0.94
Pre-Articulation-word rate	0.50	0.11	4.51	0.00	Pre-Articulation-number rate	0.56	0.10	5.62	0.00
Model 5b					Model 6b				
Intercept	5.82	0.23	24.82	0.00	Intercept	7.52	0.25	30.17	0.00
Strategy vs. No-strategy	0.07	0.29	0.25	0.80	Strategy vs. No-strategy	0.60	0.33	1.82	0.08
Control vs. No-strategy	−0.13	0.29	−0.46	0.65	Control vs. No-strategy	−0.03	0.33	−0.08	0.94
Pre-Articulation-word rate	0.50	0.11	4.51	0.00	Pre-Articulation-number rate	0.56	0.10	5.62	0.00
Random Effects	School level variance	Student level residual variance	χ^2	BIC	Random Effects	School level variance	Student level residual variance	χ^2	BIC
	0.14	0.79	22.18*	167.74		0.09	1.02	15.57	179.91

Note. Coefficient = Unstandardized coefficient in HLM models, the mean difference between study groups adjusted for level-1 covariates; SE = Standard Errors of the coefficient, BIC = Bayesian Information Criterion. Because of Dummy coding, the Random Effects are the same for Model 5a and 5b, and Model 6a and 6b. Models 5a~6b represent the comparison of “Strategy vs. Control” and “No-strategy-instruction vs. Control”.

All coefficients in HLM models became non-significant after application of the Benjamini–Hochberg correction of *p* values for multiple comparisons (i.e., $p < .01$).

* $p < .05$,

+ $p < .08$.

Appendix C

Fixed and Random Effects for Group Comparisons on QRI Retell and Listening Comprehension Measures (HLM Full Models)

	QRI-Retell					**QRI-Passage Listening Comprehension**			
Fixed Effects	**Coefficient**	**SE**	**T-ratio**	***p***	**Fixed Effects**	**Coefficient**	**SE**	**T-ratio**	***p***
Model 7a					Model 8a				
Intercept	3.10	0.29	10.65	0.00	Intercept	10.72	1.30	8.26	0.00
Strategy vs. Control	0.85	0.38	2.25	0.30	Strategy vs. Control	3.96	1.81	2.18	0.04
No-strategy vs. Control	0.95	0.40	2.38	0.02	No-strategy vs. Control	0.98	1.89	0.52	0.61
Pre-QRI-Retell	0.04	0.14	0.25	0.80	Pre-QRI-Passage Listening Comprehension	0.39	0.11	3.42	0.00
Pre - Listening Comprehension	0.19	0.05	4. 47	0.00	Pre - Listening Comprehension	0. 21	0.17	1.23	0.23
Model 7b					Model 8b				
Intercept	4.05	0.30	13.56	0.00	Intercept	11.70	1.32	8.89	0.00
Strategy vs. No-strategy	−0.10	0.39	−0.25	0.80	Strategy vs. No-strategy	2.98	1.84	1.62	0.11
Control vs. No-strategy	−0.95	0.40	−2.38	0.02	Control vs. No-strategy	−0.98	1.89	−0.52	0.61
Pre-QRI-Retell	0.04	0.14	0.25	0.80	Pre-QRI-Passage Listening Comprehension	0.39	0.11	3.42	0.00
Pre - Listening Comprehension	0.19	0.04	4. 47	0.00	Pre - Listening Comprehension	0. 21	0.17	1.23	0.23
Random Effects	School level variance	Student level residual variance	χ^2	BIC	Random Effects	School level Variance	Student level residual variance	χ^2	BIC
	0.02	30.93	9.17	357.77		0.14	1.34	17.61	196.75

Note. Coefficient = Unstandardized coefficient in HLM models, the mean difference between study groups adjusted for level-1 covariates; SE = Standard Errors of the coefficient, BIC = Bayesian Information Criterion. Models 7a~8b represent the comparison of “Strategy vs. Control” and “No-strategy-instruction vs. Control”. Because of Dummy coding, the Random Effects are the same for Model 7a and 7b, and Model 8a and 8b. Because there were marginally statistically significant between-group differences on listening comprehension at pre-training ($p = .06$), we also controlled pre-training listening comprehension for group comparisons on post-training QRI-Retell and QRI-Passage Listening Comprehension.

All coefficients in HLM models became non-significant after application of the Benjamini–Hochberg correction of p values for multiple comparisons (i.e., $p < .01$).

* $p < .05$,

+ $p < .08$.

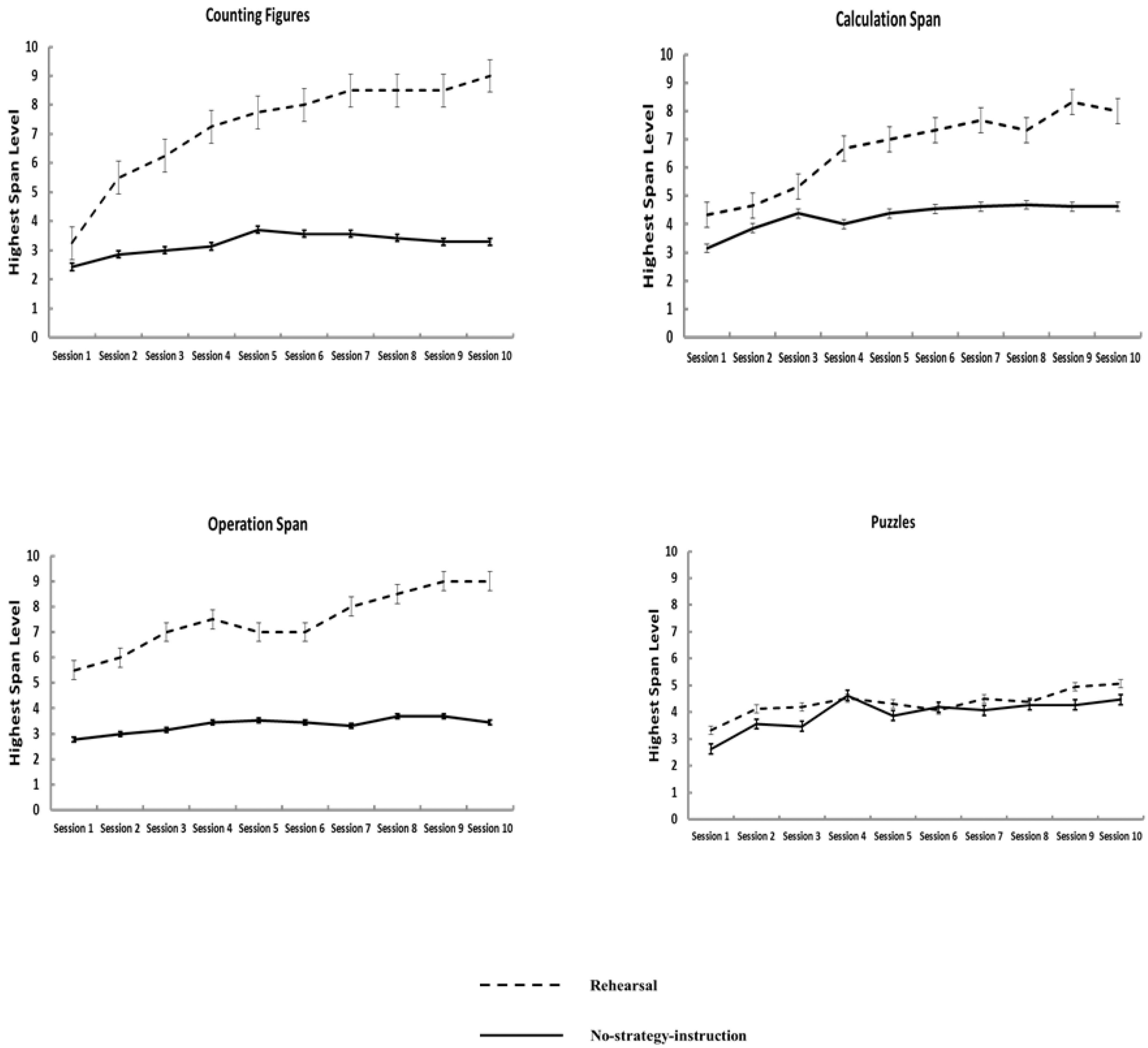

Figure 1.
Each data point in each of the four graphs represents an averaged highest performance score for a given training session. This holds for both rehearsal and no-strategy-instruction groups. Standard errors are represented by error bars attached to each point.

Table 1

Demographics and Pre-Training Non-verbal IQ, Word Reading, and Listening Comprehension Performance by Study Groups

	Rehearsal (n = 19)			No-strategy-instruction (n = 19)			Control (n = 20)		
Variables	***n***	**Mean**	**SD**	***n***	**Mean**	**SD**	***n***	**Mean**	**SD**
Age (years)		7.19	.41		7.09	.36		7.13	.33
Gender (female)	9 (47.3%)			7 (36.8%)			10 (50%)		
Race									
African American	9 (47.3%)			13 (68.4%)			5 (25%)		
Caucasian	6 (31.6%)			2 (10.5%)			10 (50%)		
Hispanic	2 (10.5%)			3 (15.7%)			3 (15%)		
Other	2 (10.5%)			1 (5.2%)			2 (10%)		
Free/Reduced Lunch	16 (84.2%)			16 (84.2%)			14 (70%)		
Non-Verbal IQ		12.47	6.86		9.16	4.49		10.55	4.86
Word Identification		37.95	11.94		34.68	10.07		38.05	9.59
Listening Comprehension		17.16	5.39		14.95	4.62		18.50	3.55

Note. Non-verbal IQ is WASI Matrix Reasoning (Wechsler, 1999); Word Identification is the Word Identification Subtest of the Woodcock Reading Mastery Test-Revised (Woodcock et al., 2001); Listening Comprehension is the Woodcock–Johnson Oral Comprehension subtest (Woodcock, McGrew, & Mather, 2001). There were no significant group differences on those variables.

Table 2

Pre-and Post-Training Performance of Study Groups

Measures	Rehearsal (*n* = 19)						No-strategy-instruction (*n* = 19)						Control (*n* = 20)					
	Pre-training			Post-training			Pre-training			Post-training			Pre-training			Post-training		
	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis	Mean (SD)	Skewness	Kurtosis
Listening Recall	7.05 (3.46)	.41	.37	9.00 (3.30)	−.08	.40	6.37 (3.39)	.32	−.22	7.79 (2.76)	.51	−.69	6.60 (3.32)	.51	−.98	7.60 (3.10)	.65	−.16
Counting Recall	12.37 (3.98)	−.01	−.77	14.84 (3.44)	−.47	1.44	12.26 (3.91)	−.26	−.64	15.58 (3.99)	1.12	1.16	13.10 (4.98)	.54	1.22	14.20 (4.88)	−.47	.22
Articulation-word rate	5.16 (1.17)	−.26	−.96	6.02 (1.21)	−.18	−1.41	4.66 (.98)	.36	−.21	5.68 (1.11)	−.35	.20	5.21 (1.19)	.29	−.09	5.90 (1.05)	.52	.42
Articulation-number rate	6.72 (1.72)	−.32	−1.24	8.14 (1.40)	−.32	−.76	6.65 (1.07)	−.67	−.69	7.47 (1.19)	−.97	1.33	6.83 (1.51)	−.72	−.37	7.58 (1.39)	.10	−.91
Digit Recall	24.05 (4.59)	1.37	4.04	23.16 (4.27)	.16	.05	23.11 (3.76)	−.03	−.70	22.68 (3.15)	1.08	3.64	23.95 (3.50)	−.62	.36	23.90 (3.57)	−.08	−.93
Block Recall	19.00 (4.82)	−.24	.46	18.79 (4.93)	−.12	1.12	18.00 (3.51)	−.49	.37	18.79 (4.78)	.06	.20	20.45 (4.27)	−.09	−.91	19.95 (4.47)	.09	−.56
QRI-Retell	10.95 (7.18)	.58	−.58	14.63 (7.44)	.23	.58	10.79 (5.72)	.13	−.37	11.11 (6.44)	.46	1.20	11.68 (7.34)	1.07	1.27	11.50 (4.56)	.12	−.35
QRI-Passage Listening Comprehension	4.26 (1.37)	−.53	.06	3.95 (1.54)	−.81	.96	4.26 (1.24)	−.37	−.65	3.63 (1.74)	−.35	−.58	4.40 (1.39)	−.29	−1.12	3.40 (1.14)	.51	−.19

Note. Listening Recall is adapted from the Listening Recall from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001); Counting Recall is adapted from the Counting Recall from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001); Articulation-word rate is the word rehearsal speed adapted from Kail (1997); Articulation-word rate is number rehearsal speed adapted from Kail (1997); Digit Recall is adapted from the Digit Recall from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001); Block Recall is from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001); QRI-Retell is the retell score from the Qualitative Reading Inventory (Leslie & Caldwell, 2001); QRI-Passage Listening Comprehension is the comprehension score from the Qualitative Reading Inventory (Leslie & Caldwell, 2001), and two different passages from the same level were used in pre- and post-testing, respectively.

Table 3

Fixed and Random Effects for Group Comparisons with Uncorrected p Values

	Fixed Effects												Random Effects			
	Rehearsal vs. Control				No-strategy-instruction vs. Control				Rehearsal vs. No-strategy-instruction							
Measures	***Coefficient***	***p***	**SE**	**ES**	***Coefficient***	***p***	**SE**	**ES**	***Coefficient***	***p***	**SE**	**ES**	**School level variance**	**Student level residual variance**	χ^2	**BIC**
Listening Recall	1.50	.03	.67	.47	.49	.47	.67	.17	1.01	.14	.67	.33	1.15	4.15	26.14*	262.48
Counting Recall	.95	.45	1.25	.22	1.6	.17	1.26	.36	−.81	.52	1.26	−.22	1.34	14.99	17.24	327.21
Articulation-word rate	.21	.47	.29	.19	.13	.65	.29	.12	.07	.80	.29	.06	.14	.79	22.18*	167.74
Articulation-number rate	.63	.06	.33	.45	.03	.94	.33	.02	.60+	.08	.33	.46	.09	1.02	15.57	179.91
Digit Recall	−.82	.29	.76	−.20	−.57	.46	.77	−.17	−.25	.75	.77	−.07	.33	5.60	13.31	272.66
Block Recall	0	.99	1.08	0	.78	.48	1.08	.17	−.79	.47	1.09	−.16	.01	11.17	10.09	307.52
QRI-Retell	3.96	.04	1.81	.65	.98	.61	1.89	.16	2.99	.11	1.84	.43	.02	30.93	9.17	357.77
QRI-Passage Listening Comprehension	.85	.03	.38	.63	.95	.02	.40	.65	−.10	.80	.39	−.06	.14	1.34	17.61	196.75

Note. Coefficients = Unstandardized coefficients in HLM models, the mean difference between study groups adjusted for level-1 covariates; SE = Standard Errors of the coefficient, ES = Effect Size (Hedges g). BIC = Bayesian Information Criterion.

All coefficients in HLM models became non-significant after application of the Benjamini–Hochberg correction of *p* values for multiple comparisons (i.e., $p < .01$).

* $p < .05$,

\+ $p < .08$.