Developmental Review 65 (2022) 101035

Contents lists available at ScienceDirect

Developmental Review

journal homepage: www.elsevier.com/locate/dr

The domain-specific approach of working memory training

Peng Peng [a,*], H. Lee Swanson [b]

[a] *The University of Texas at Austin, USA*
[b] *The University of New Mexico, USA*

A R T I C L E I N F O

Keywords:
Working Memory
Domain-specific
Strategy
Transfer

A B S T R A C T

Converging evidence suggests that traditional domain-general working memory (WM) training does not have reliable far-transfer effects, but produces reliable, modest near-transfer effects on structurally similar untrained tasks. Given the critical role of WM in academic development, WM training that incorporates task-specific features may maximize training effects on academic outcomes. In this theory paper, we discuss the training to emphasize the domain-specific function of WM highlighted by recent WM models. That is, WM should be better attuned to the materials being learned through enhancing strategies of linking together WM with the long-term memory knowledge, rather than only the enhancement of a "domain-general" attentional control overall. We provided two example training routes that emphasize explicit instruction and practice on WM-academic tasks (i.e., academic tasks that can be performed using a WM training paradigm) and task-linking strategies (i.e., strategies that can be used in both academic tasks and WM tasks to improve performance efficiency). We also review recent relevant intervention studies that are in line with this approach and report promising effects on academic outcomes. Implications for future studies are also discussed.

Introduction

Working memory (WM) is a system for the active maintenance and manipulation of information that is relevant for current task goals (Baddeley, 1992), different from short-term memory that is a passive storage system. WM plays an important role in children's academic performance (Daneman & Merikle, 1996; Friso-van den Bos et al., 2013; Peng et al., 2016, 2018; Raghubar et al., 2010; Swanson & Alloway, 2012). This is because many academic tasks involve multiple steps with intermediate information that must be remembered for a short period and simultaneously manipulated to accomplish the task at hand. Not surprisingly, WM deficits present the core of most severe cognitive deficits among children with varying types of learning disabilities, including difficulties in mathematics (Peng et al., 2018; Swanson & Jerman, 2006) and in reading (Gathercole et al., 2006; Swanson & Jerman, 2006; Wang & Gathercole, 2013).

Given the important role of WM in learning, over the past several decades, many studies have been conducted to understand how WM can be used to guide academic curriculum design and instruction for both typically developing and at-risk learners (Cowan, 2014; Gathercole et al., 2006; Paas et al., 2003, Swanson, 2020). One important line of research focuses on WM training. The fundamental hypothesis for the majority of WM training studies is a domain-general approach that emphasizes WM capacity is domain general and

* Corresponding author at: Department of Special Education, College of Education, The University of Texas at Austin, SZB 408D, 1912 Speedway STOP D5300, Austin, USA.
E-mail address: kevpp2004@hotmail.com (P. Peng).

https://doi.org/10.1016/j.dr.2022.101035
Received 10 November 2021; Received in revised form 29 June 2022;

malleable, so that it can be increased through training just as muscles can be trained to boost strength; improved WM capacity is therefore expected to have a long-lasting impact on skills (e.g., academic skills) that are closely related to WM (Holmes et al., 2009; Jaeggi et al., 2008; Melby-Lervåg & Hulme, 2013). Early optimistic reports in support of this analogy were later shadowed by *meta*-analyses indicating near-zero far-transfer effects (effects on untrained academic tasks with different task-paradigm/structure from the WM training tasks, Melby-Lervåg et al., 2016; Sala & Gobet, 2020; Shipstead et al., 2012).

Nevertheless, the majority of studies and *meta*-analyses have demonstrated robust, replicable near-transfer effects on similar WM tasks/variants that usually have the same materials (e.g., both training and near-transfer tasks have identical verbal materials), task paradigms (e.g., both training and near-transfer tasks have the same complex span paradigm), and strategies (e.g., both training and near-transfer tasks can be performed with the same strategy) (Fellman et al., 2020; Melby-Lervåg et al., 2016; Schwaighofer et al., 2015; Soveri et al., 2017). These near-transfer effects, largely driven by task-specific features in WM training, suggest a domain-specific approach to think about how to maximize WM training transfer effects on academic skills.

Specifically, based on the cognitive load theory, language/verbal skills, WM, and some cognitive strategies (e.g., sub-vocal rehearsal) can be considered as primary knowledge/skills that are often naturally acquired in a typical developmental trajectory. The major purpose of school is to teach students how to utilize the primary knowledge/skills to acquire evolutionarily novel knowledge such as academic learning (Geary, 1995; Sweller et al., 2019, 2021, 2022). However, because of individual differences in the primary knowledge/skills, students may utilize the primary knowledge/skills to varying degrees during academic learning. In addition, the cognitive load theory emphasizes that WM is limited and can be overloaded when instructional materials are novel and complicated, but WM can be unlimited when knowledge of instructional materials can be fluently retrieved from long-term memory (Sweller et al., 2019). Thus, instruction should be designed to help students efficiently utilize WM (not to overload WM) by linking learning materials to long-term memory or primary knowledge (language and rehearsal strategy) to the greatest extent.

Under the cognitive load theory framework, the present paper discussed a domain-specific approach of WM training to help students utilize WM in actively searching for and activating knowledge/strategies in long-term memory, which not only increases the efficiency of WM in a specific task (maximizing the domain-specific function of WM) but also directly serves to the purpose of academic instruction at schools (Ericsson & Kintsch, 1995; Logie et al., 2020; Sweller et al., 2019; Unsworth & Engle, 2007). Of note, we did not intend to propose a theory on WM training that is readily to be applied in the practice (as we only had limited preliminary evidence later discussed in this paper). Rather, this paper proposes an alternative way of thinking about WM training from a developmental perspective, contributing to the debate/development of WM training theories.

In the following sections, we first describe the theoretical foundations for the domain-specific approach of WM training, including theories/debates on the domain-specificity of WM, the mutualism theory between WM and academic skills, and the element theory of cognitive transfer. We then describe two training task features of the domain-specific approach of WM training: WM-academic tasks (academic tasks with embedded WM training paradigm) and task-linking strategies (strategies such as rehearsal of important information that can be used in both academic and WM tasks to facilitate task performance). Next, we review relevant intervention studies that have used WM-academic task training and/or task-linking strategy training. Finally, we discuss implications for future research in academic interventions based on the domain-specific approach of WM training.

Domain-Specific function of working memory

In Baddeley's (1986) componential WM model, WM consists of two "slave systems" responsible for short-term maintenance of domain-specific (verbal and visuospatial) information and a central executive that coordinates ongoing processing and storage of information in those slave systems. The central executive (i.e., attentional control) directs attention to relevant information, suppressing irrelevant information and inappropriate actions, and coordinates cognitive processes when more than one task must be accomplished at the same time. It also differentiates WM from short-term memory, such that the central executive in Baddeley's componential model and many other successor WM models is the core component of WM or represents WM as a construct (Engle, 2018; Engle & Kane, 2004; Oberauer, 2002). Thus, a person's WM capacity is determined mostly by the central executive (attentional control), which is considered domain general (Engle, 2018; Kane et al., 2005; Peng et al., 2018; Süß et al., 2002).

However, there is a merging consensus among WM theorists and researchers that a completely domain-general WM model does not hold (Doebel, 2020; Logie et al., 2020). The operation of WM also depends on domain knowledge and thus the function of WM has domain specificity that is closely linked to long-term memory (Baddeley, 2000; Cowan, 1999; Ericsson & Kintsch, 1995; Paas et al., 2004; Unsworth et al., 2013; but cf. Hambrick & Engle, 2002). Specifically, Baddeley (2000) updated their componential WM model by adding a new component called the episodic buffer. The major role of the episodic buffer is to provide temporary storage of information held in a phonological loop and visuospatial sketchpad, as to bind information from these subsidiary systems and from long-term memory into a unitary episodic representation. Thus, episodic buffer has the domain-specific function to link among long-term memory, the central executive, and the short-term storage system as to better facilitate the use of central executive and short-term storage (Baddeley et al., 2010). Similarly, in Ericsson and Kintsch's (1995) long-term WM model, long-term memory is suggested to supplement or facilitate WM: individuals knowledgeable in a particular domain can process (encode and retrieve) information in that domain more efficiently than information in domains they are less knowledgeable about. Based on these two models, the major function of WM includes the integration of domain-specific skills, knowledge, and procedures as to meet the particular demands of learning tasks within a particular domain.

Cowan's and Unsworth' WM models also specified how long-term memory works within/together the WM system. Specifically, in the embedded-processes model of WM proposed by Cowan (Cowan, 1995, and 1999), WM is determined by the focus of attention that is capacity limited. WM information comes from hierarchically arranged entities consisting of long-term memory, a subset of long-term memory currently activated (not in the focus of attention), and a subset of activated long-term memory in the focus of attention. In this model, task-relevant information in long-term memory needs to be activated and ready to be used in a task and new combinations of information can be formed within this active memory in the focus of the attention, which then may become a part of long-term memory (learning or acquisition of new knowledge). In a similar vein, Unsworth and Engle (2007) proposed that WM consists of primary memory and secondary memory. Primary memory serves to maintain a distinct number of separate representations active for ongoing processing. Secondary memory is responsible for retrieving information that is not maintained in primary memory, which usually requires a controlled and strategic search in long-term memory (Unsworth et al., 2013). Taken together, both Cowan's and Unsworth' WM models implicated that in order for learning to happen, WM (attentional control) needs to timely (moment-to-moment) and efficiently search for, activate, maintain, and process relevant information from long-term memory, and combine different information to create new knowledge in long-term memory.

The domain-specific function view of WM well explains experts' performance differences between WM-demanding tasks that require their expertise and WM-demanding tasks for which they lack expertise. For example, chess experts can simultaneously process and remember many chess patterns during chess games that novice players cannot, whereas chess experts and novices perform equally well on verbal and visuospatial WM tasks that do not require chess knowledge (Unterrainer et al., 2006; Waters et al., 2002). These domain-specific function views of WM have strongly influenced research on expertise and decision making in various domains, such as firefighting, medical diagnosis, and aviation, where experts/professionals are more likely to make timely, correct decisions in WM-demanding, complex problem-solving situations (e.g., Zsambok & Klein, 2014). Thus, based on the domain-specific function of WM, one's WM capacity is at least jointly determined by two collaborating core components: the domain-general central executive (attentional control), and the ability to retrieve domain-specific knowledge from long-term memory.

Mutualism in the development of working memory and academic Skills.

The evidence for the domain-specific function of WM usually involves comparing verbal and visuo-spatial WM tasks in their correlations with performance on complex verbal or spatial tasks (Friedman & Miyake, 2000). Although there is evidence to support both domain-general and domain-specific WM (e.g., Conway et al., 2007; Kane et al., 2004; Shah & Miyake, 1996; Friedman & Miyake, 2000), many studies focusing on children's learning seem to favor the domain-specific perspective (also see Cowan et al., 2005; Gray et al., 2017). Research has shown, for example, that children's visuo-spatial WM fails to explain variance in their word reading and comprehension of written passages (e.g., Nation et al., 1999; Seigneuric et al., 2000), whereas verbal WM accounts for statistically significant variance in performance on those verbal tasks, even when relevant verbal skills (e.g., word reading) are controlled for (e.g., Arrington et al., 2014; Cain et al., 2004). Meta-analyses on children with specific learning difficulties also suggests a pattern of domain-specificity of WM in relation to learning. That is, although children with various types of learning difficulties showed WM deficits across (verbal, numerical, and visuospatial) domains (Peng et al., 2016; Swanson & Jerman, 2006), those with specific mathematics difficulties showed more severe numerical WM deficits than their verbal WM deficits (Peng et al., 2016). Children with mathematics difficulties could be differentiated from children with reading difficulties on visuo-spatial WM (Swanson & Jerman, 2006), likely reflecting the salient role of visuo-spatial WM in mathematics learning and mathematics difficulties (Mammarella et al., 2017).

Recently, emerging evidence suggests that the development and practice of domain-specific skills may influence and shape the domain-specific function of WM and its role in academic development. For example, van de Weijer-Bergsma et al. (2015) found that both visuospatial and verbal WM were associated with calculations early on, because young children relied on both visuospatial and verbal long-term memory knowledge (e.g., representations and strategies) to perform WM-demanding calculation problems. However, as children developed through the elementary grades, the effects of visuospatial WM on calculations waned while the effects of verbal WM increased. This developmental pattern is due primarily to the shift from using both visuospatial and verbal representations and strategies during calculations early on to the use of verbal representations and strategies during calculations in later grades (De Smedt et al., 2009; Geary et al., 2004).

Similar patterns have also been reported for the relation between WM and reading (Miller-Cotto & Byrnes, 2020; Peng et al., 2018). Peng et al.'s (2018) *meta*-analysis of 197 studies found that WM is constantly related to reading through development. Although domains of WM did not influence its relations to reading early on (e.g., before the 4th grade), the relation between verbal WM (in comparison with visuospatial WM) and reading became stronger with development (e.g., after the 4th grade). Such effects were robust even when other confounding variables (e.g., publication type, types of reading tasks, and bilingual status) were controlled for. Peng et al. explained these findings within a WM–reading development model as follows: The domain-general central executive of WM is heavily involved in reading's early stages. As reading experience accumulates, lexical and verbal knowledge are consolidated in long-term memory, and readers come to rely more on direct retrieval of lexical and verbal knowledge from long-term memory to perform a variety of reading tasks. As students develop foundational reading skills and attempt to read for understanding, WM resources are allocated to integrate verbal knowledge and procedures to meet the demands of reading tasks, strengthening verbal WM and the impact of verbal WM on reading in the process. In the WM–reading development model, the relation between reading and WM is reciprocal: WM primarily exerts an impact on reading early on, with reading also shaping the further development of verbal WM.

Most recently, mutualism theory has been empirically tested to reveal possible mutualism among different cognitive skills and academic skills (e.g., Cunningham et al., 2021; Kail et al., 2016; Melby-Lervåg et al., 2012; although there is still debate on this theory). In a review, Peng and Kievit (2020) explored the potential mutualism between cognition (including WM) and academic achievement based on (1) developmental effects on concurrent relations between cognition and academic performance, (2) longitudinal relations between cognition and academic performance (especially between verbal WM and language skills), and (3) the effects of explicit academic instruction on cognition. Peng and Kievit (2020)'s findings, together with other recent empirical studies not included in their review (e.g., Ritchie et al., 2015; Ritchie & Tucker-Drob, 2018; Willoughby et al., 2019; Zhang & Peng, 2022), suggest a possible cognitive–academic bidirectional model. That is, in the early developmental stage, the accumulation and automatic retrieval of domain-specific knowledge are not well established, and students are more likely to invest their cognitive abilities in the acquisition of learning. Improvements in concrete knowledge then benefit more abstract cognitive abilities through mechanisms such as semantic bootstrapping. Students with more advanced verbal skills and knowledge (e.g., vocabulary, comprehension) can decompose abstract cognitive problems into constituent rules more efficiently (Kievit et al., 2017). Thus, the long-lasting learning and practicing of academic skills/tasks serve as an "intervention" to improve cognition with development. Moreover, such cognitive–academic bi-directionality is greatly influenced by environmental stimulation, such that sustained, high-quality academic instruction directly fosters children's academic and cognitive development, indirectly affecting academic and cognitive development by triggering cognitive–academic bi-directionality. Taken together, the mutualism suggests that the mutual impact between WM and academic skills is not likely to happen with a short period of time and requires a long-lasting and high-quality instruction input, especially among children with disadvantages.

Based on the mutualism between WM and academic development (how environmental stimulation can facilitate that mutualism), it is important to provide academic instruction and WM training simultaneously to produce synergistic, sustained effects on academic outcomes. However, prior WM training research barely linked WM training with academic instruction. Many existing WM training tasks and programs require children to practice WM tasks outside of the context of education or learning—that is, tasks that are not related to the practice of reading or of mathematics skills in depth. Such training is task-paradigm and material fixed, produced little effects on academic outcomes (even when considering publication bias such that sigfnicant training effects are more likely to be published, Sala & Gobet, 2017; Shipstead et al., 2012), did not seem relevant in an educational setting, and was unlikely to be adopted by schools or families especially when academic instructional time is already limited (Peng & Fuchs, 2017).

For example, traditional WM training studies have often used the *n*-back task (Shipstead et al., 2012, even though there are arguments against using *n*-back tasks as WM tasks Kane et al., 2007; Redick & Lindsey, 2013; Schmiedek et al., 2014). In *n*-back tasks, the participant is presented with a sequence of stimuli (e.g., figures and numbers) and must decide for each stimulus whether it is the same as a stimulus presented *n* trials earlier. In a single *n*-back task, the participant is required to attend to one stream of stimuli, and in a dual *n*-back task, two streams of stimuli are presented simultaneously, such that the participant must respond to both an auditory-verbal and a visuospatial stream of stimuli. The training follows an adaptive rule, such that as a person meets specific performance criteria, the task's difficulty increases. When these criteria are not met, task difficulty decreases. Such training paradigm purely focused on the practice of *n*-back tasks paradigm with very simple and limited types of materials (e.g., simple figures and numbers).

For most WM training studies with children, the CogMed WM training program has been adopted (Sala & Gobet, 2017). This program includes several visuospatial and verbal memory training tasks, in which children are trained with the adaptive rule. For example, in a visuospatial task called "Asteroids," a field of several free-floating asteroids is presented on the screen, and a subset of them lights up one at a time. The children reproduce the sequence of asteroids via mouse clicks. The length of the sequence to be reproduced changes based on children's performance trial-by-trial. In a verbal task called "Input Module," a sequence of auditory digits is played. The children reproduce the digits' sequence in reverse/forward order by clicking on a number pad displayed on a robot's arm. The length of the sequence to be reproduced changes based on children's performance trial-by-trial. Because these tasks are embedded in videogames where children must click on a screen, CogMed's activities are dominated by the simple manipulation of simple visuospatial stimuli even for verbal memory tasks.

Adding traditional WM training to academic instruction also does not seem to produce synergistic effects on academic skills as hypothesized by the mutualism model. Many WM intervention studies with school-aged children have been conducted during school time. Children in most of those studies, whether in WM training groups or control groups, were receiving regular academic instruction at school during the period of research (e.g., they were pulled out of the classroom for training or received training after school). With this design, studies have seldom shown that the addition of WM training leads to greater improvement (far-transfer) in academic outcomes among participants than among controls (Sala & Gobet, 2017; Shipstead et al., 2012).

Element theory of cognitive transfer

Thus, WM training with a more nuanced task design is needed to increase the educational relevance of WM training to tie WM training closely to academic content from a developmental perspective. Such thought is in line with the element theory of cognitive skills (Taatgen, 2013; Woodworth & Thorndike, 1901), which specifies how skills are acquired and how transfer between skills can be explained.

Specifically, the theory claims skills can be broken down into primitive information processing elements, with some that are task-specific (e.g., retrieval of task-specific information), and some task-independent (e.g., attentional monitoring of the task procedure). A large overlap in these elements between tasks is the foundation for task learning/training and transfer. Taatgen (2013) suggested that in the early learning stage, novice need to attend to each important element in a task, and the processing of each element consumes WM. With the progress of learning and accumulation of knowledge, experts can combine these elements into fewer but larger clusters (e.g., element 1 and 2 can be combined into cluster 1; element 3 and 4 can be combined into cluster 2), saving WM capacity from processing each element. More importantly, these large clusters also bear an important transfer feature. That is, through the combination of the elements into larger clusters, individuals form strategies such as iteration (the repetition of a process in order to generate an outcome) and rehearsal (the repetition of information for maintenance), which can be used in a different task that also require these strategies.

For example, during reading comprehension, one needs to at least process several fundamental elements, including (not limited to)

1) reading the word, 2) the retrieval of word meaning from long-term memory, 3) linking all words together to understand the sentence, 4) the retrieval of background knowledge relevant to the sentences/text topic, 5) the maintenance of information from previous read sentences, 6) integrating information from previously read sentences with incoming text information, and 7) integrating information retrieved from long-term memory with text information. Beginning readers (early instruction on reading comprehension) often need to use WM to process each of these above-mentioned elements. In contrast, experienced/knowledgeable readers, with practice, can combine some of these elements into fewer and larger clusters such as 1) reading and retrieving the meaning of words can be simultaneously done with the understanding of the sentence, 2) the retrieval of background knowledge can be simultaneously done with the maintaining of information extracted from the text and the integrating information from the text and long-term memory. These larger clusters also promote strategies such as iteration (e.g., repeating the integration process of information from long-term memory and from text) and rehearsal (e.g., maintaining relevant and important information), which can be applied to a different task such as mathematics word problem solving.

The Domain-Specific approach of working memory training

Taken together, based on the domain-specific function view of WM, the mutualism between WM and academic development, and the element theory of cognitive skills, to maximize the effects of WM training on academic outcomes, we argue for a domain-specific approach of WM training that emphasizes 1) training tasks should closely link the central executive (attentional control) with the use/retrieval practice of long-term memory in a specific academic domain, and 2) training tasks should promote strategy use that can be effectively applied in different academic tasks. Together with near-transfer effects of WM training on structurally similar tasks (Melby-Lervåg et al., 2016; Schwaighofer et al., 2015; Soveri et al., 2017), the domain-specific approach of WM training is more closely related to academic development, incorporates various materials and strategies, and thus is more likely to produce effects on academic outcomes in comparison to training WM or academic skills instruction alone.

WM-academic tasks. One important component to link the central executive (attentional control) with the use/retrieval practice of long-term memory in a specific academic domain is to combine the WM training paradigm with academic materials, which we labeled as WM-academic training task. Training on WM-academic tasks requires children to actively utilize WM and long-term memory (searching for and retrieval of relevant information) simultaneously on a moment-to-moment basis. Knowledge formed during such training can further shape new long-term academic knowledge, thus facilitating academic learning. This is a developmental approach such that central executive and long-term memory are closely collaborated to process information and form new knowledge (Cowan, 1999; Ericsson & Kintsch, 1995; Unsworth et al., 2013).

Specifically, traditional WM training tasks only focused on the WM paradigm, which is fixed from a developmental perspective. In comparison, WM-academic tasks are academic-relevant and thus have the potential to be embedded into educational practice, which is developmental (accumulative) in nature (Ericsson et al., 2018). Moreover, such long-lasting practice on these tasks are more likely to trigger WM-academic mutualism, which helps further enhance the domain-specific function of WM in learning within a certain domain (e.g., verbal WM for reading) through development (Peng et al., 2018; Peng & Kievit, 2020).

Moreover, because WM-academic tasks are academic in nature, instruction of these tasks follow explicit, evidence-based instructional principles. That is, children are taught or trained on these tasks with a series of supports or scaffolds, such that they are guided through the learning process with clear statements about the rationale for learning the academic skill, clear explanations and demonstrations of task procedures, and supported practice with feedback to reach independent mastery (Kirschner et al., 2006; Stockard et al., 2018). In addition to these benefits of the traditional academic explicit instruction, WM-academic tasks also *explicitly* requires students to *exercise* the WM process during academic learning. Such training is necessary for complex academic skills acquisition that inevitably requires children to process information under a high WM load condition (e.g., Kendeou et al., 2014).

We can use the retell task based on complex span WM training as an example (WM-retell task). In an operation span task during WM training, students are first asked to solve a simple calculation problem (presented on a flashcard) and then name a picture card (e.g., flower). This process proceeds with additional calculations and picture cards. At the end of each trial, students are asked to recall all the picture cards in order. The number of picture cards in each trial is constantly adjusted to challenge students' memory span. Based on the domain-specific approach of WM training, such complex span tasks should be directly modified to include the practice of reading skills and use of knowledge. For example, in such a task, students can be asked to read a story with several paragraphs. Each time students finish reading a paragraph, they summarize the paragraph's main idea. At the end of the story, they recall in order the main ideas from each paragraph in order to get the story's general main idea. Depending on the level of their performance, students can be asked to practice with stories that have more or fewer paragraphs, or texts that vary in difficulty. The WM-retell task is based on the retell process, but it differs from traditional retell instruction, which has students practice how to retell *without* explicitly teaching students to practice the WM process during the retell. That is, the WM-retell task further decomposes the retell into more explicit teachable steps to practice reading paragraphs, summarize paragraphs, remember main ideas of each paragraph, and practice how to allocate attention to important information in passage comprehension and remembering main ideas of previous paragraphs.

The task-linking strategy. As emphasized by the element theory of cognitive skills, strategies that can be applied to the training tasks and the transfer tasks are important elements for cognitive transfer (Taatgen, 2013). Following this logic, training should emphasize strategies that can be applied to both WM tasks and academic tasks and can be used to improve task performance efficiency (e.g., facilitating the information retrieval from the long-term memory, maintaining information) under high WM load conditions. We call these strategies task-linking strategies.

Indeed, studies are emphasizing the important role of strategy use in WM performance and the relation between WM and other tasks (e.g., Gonthier & Roulin, 2020; Turley-Ames & Whitfield, 2003). Specifically, research suggests the use of a specific (effective)

Table 1
Summary of Reviewed Studies.

Study	Age/Grade	Status	Study Groups	Domain-specific Approach of WM training		Effect Sizes (Hedges g/Cohen's d) on Academic Outcomes	
				WM-Academic Task	Task-linking Strategy	Researcher-developed Academic Measures	Standardized Academic Measures
Fuchs et al., 2018	Grade 3 and 5	At-risk readers	Randomized control trial 1) Business-as-usual (BAU), n = 382) Reading (R) , n = 403) Domain-specific Approach of WM training (DS-WM) , n = 38	WM-comprehension tasks	N/A	DS-WM vs. BAU Grade 3, Knowledge acquisition = 1.90** Grade 3, Reading comprehension = 0.79* Grade 5, Knowledge acquisition = 1.60** Grade 5, Reading comprehension = 0.31 R vs. BAU Grade 3, Knowledge acquisition = 1.36** Grade 3, Reading comprehension = 0.05 Grade 5, Knowledge acquisition = 1.46** Grade 5, Reading comprehension = 0.52	DS-WM vs. BAU Grade 3, Reading comprehension = 0.08 Grade 5, Reading comprehension = 0.33 R vs. BAU: Grade 3, Reading comprehension = -0.17 Grade 5, Reading comprehension = 0.21
García-Madruga et al., 2013	8.45 years old/Grade 3	Non-selective	Randomized control trial1) Reading (R) , n = 162) Domain-specific Approach of WM training (DS-WM) , n = 15	WM/ Executive function -comprehension tasks	N/A	N/A	DS-WM vs. R Reading comprehension = 0.72
Goodrich et al., 2020;	4.38 years old/preschool	At-risk readers	Randomized control trial 1) Business-as-usual (BAU), n = 232) Reading (R) , n = 233) Domain-specific Approach of WM training (DS-WM) , n = 23	WM/ Executive function –reading/ language tasks	Rehearsal strategy	N/A	DS-WM vs. BAU Print knowledge = 0.25 Definitional vocabulary = 0.12 Phonological awareness = 0.49 Vocabulary = 0.76** Syntax = 1.30** Process = 0.13 DS-WM vs. R Print knowledge = 0.24 Definitional vocabulary = 0.15 Phonological awareness = 0.15 Vocabulary = 0.17 Syntax = 0.78** Process = -0.31

(continued on next page)

Table 1 (*continued*)

Study	Age/Grade	Status	Study Groups	Domain-specific Approach of WM training		Effect Sizes (Hedges g/Cohen's d) on Academic Outcomes	
				WM-Academic Task	Task-linking Strategy	Researcher-developed Academic Measures	Standardized Academic Measures
							R vs. BAU Print knowledge = 0.04 Definitional vocabulary = -0.01 Phonological awareness = 0.38 Vocabulary = 0.67** Syntax = 0.70** Process = 0.44
+Kroesbergen et al., 2014	5.87 years old/ kindergarten	Non-selective	Randomized control trial 1) Business-as-usual (BAU), n = 212) Domain-general WM training (DG-WM), n = 153) Domain-specific Approach of WM training (DS-WM), n = 15	WM/ Executive function –numerical tasks	N/A	DG-WM vs. BAU Dots comparison = 0.48 DS-WM vs. BAU Dots comparison = 0.97 DS-WM vs. DG-WM Dots comparison = 0.49	DG-WM vs. BAU Early numeracy = 0.42 DS-WM vs. BAU Early numeracy = 0.93* DS-WM vs. DG-WM Early numeracy = 0.53
Swanson, 2016	7.95 years old/Grade 3	Children with mathematics difficulties	Randomized control trial1) Word-problem verbal strategies (Verbal) n = 262) Word-problem visual strategies (Visual) = 27 3) Word-problem verbal + visual Strategies (VV) n = 39 5) Business-as-usual (BAU) n = 344) Domain-specific approach of WM training (DS-WM), n = 36	WM-word problem tasks	N/A	N/A	Verbal vs. BAU High WM group: Calculation = -0.06 Low WM group: Calculation = -0.05 High WM group: Word problem = 0.14 Low WM group: Word problem = -0.01 Visual vs. BAU High WM group: Calculation = 0.40 Low WM group: Calculation = 0.50 High WM group: Word problem = 0.31 Low WM group: Word problem = -0.29 Verbal + Visual vs. BAU High WM group: Calculation = 0.10 Low WM group: Calculation = 0.03 High WM group:

(*continued on next page*)

Table 1 (*continued*)

Study	Age/Grade	Status	Study Groups	Domain-specific Approach of WM training		Effect Sizes (Hedges g/Cohen's d) on Academic Outcomes	
				WM-Academic Task	Task-linking Strategy	Researcher-developed Academic Measures	Standardized Academic Measures
							Word problem = 0.14 Low WM group: Word problem = -0.22 <u>DS-WM vs. BAU</u> High WM group: Calculation = 0.80* Low WM group: Calculation = 0.89* High WM group: Word problem = 0.47* Low WM group: Word problem = -0.18
+Jones et al. (2020)	12.54 years old	Non-selective	Randomized control trial1) Visual search (V), n = 292) Domain-general WM training (DG-WM), n = 313) Domain-specific Approach of WM training (DS-WM), n = 32	N/A	Meta-cognitive strategies embedded into WM training	<u>DG-WM vs. V</u> Post-intervention, Reading comprehension = -0.04 Follow-up, Reading comprehension = -0.11 <u>DS-WM vs. V</u> Post-intervention, Reading comprehension = 0.09 Follow-up, Reading comprehension = 0.08 <u>DS-WM vs. DG-WM</u> Post-intervention, Reading comprehension = 0.12 Follow-up, Reading comprehension = 0.19	<u>DG-WM vs. V</u> Post-intervention, Math = 0.27 Follow-up, Math = 0.13 <u>DS-WM vs. V</u> Post-intervention, Math = 0.22 Follow-up, Math = 0.15 <u>DS-WM vs. DG-WM</u> Post-intervention, Math = -0.06 Follow-up, Math = 0.01
Partanen et al. (2015)	8.61 years old/ Grade 2 and 3	Children with special needs	Randomized control trial 1) Business-as-usual (BAU), n = 242) Domain-general WM training (DG-WM), n = 203) Domain-specific Approach of WM training (DS-WM), n = 20	N/A	Meta-cognitive strategies embedded into WM training	N/A	<u>DS-WM vs. BAU</u> Post-intervention, Arithmetic = 0.19 Follow-up, Arithmetic = -0.05 Post-intervention, Reading = -0.15 Follow-up, Reading = -0.15 Post-intervention, Phonological ability = 0.01 Follow-up,

(*continued on next page*)

Table 1 (*continued*)

Study	Age/Grade	Status	Study Groups	Domain-specific Approach of WM training		Effect Sizes (Hedges g/Cohen's d) on Academic Outcomes	
				WM-Academic Task	Task-linking Strategy	Researcher-developed Academic Measures	Standardized Academic Measures
							Phonological ability = -0.21 Post-intervention, Word comprehension = 0.07 Follow-up, Word comprehension = 0.15 Post-intervention, Spelling = 0.02 Follow-up, Spelling = 0.16 <u>DG-WM vs. BAU</u> Post-intervention, Arithmetic = -0.28 Follow-up, Arithmetic = 0.12 Post-intervention, Reading = -0.17 Follow-up, Reading = 0.02 Post-intervention, Phonological ability = -0.21 Follow-up, Phonological ability = -0.67 Post-intervention, Word comprehension = -0.19 Follow-up, Word comprehension = 0.13 Post-intervention, Spelling = -0.37 Follow-up, Spelling = 0.02 <u>DS-WM vs. DG-WM</u> Post-intervention, Arithmetic = 0.46 Follow-up, Arithmetic = -0.17 Post-intervention, Reading = 0.02 Follow-up, Reading = -0.17 Post-intervention, Phonological ability = 0.23 Follow-up, Phonological ability = 0.46 Post-intervention, Word comprehension = 0.26 Follow-up, Word

(*continued on next page*)

Table 1 (*continued*)

Study	Age/Grade	Status	Study Groups	Domain-specific Approach of WM training		Effect Sizes (Hedges g/Cohen's d) on Academic Outcomes	
				WM-Academic Task	Task-linking Strategy	Researcher-developed Academic Measures	Standardized Academic Measures
							comprehension = 0.03 Post-intervention, Spelling = 0.39 Follow-up, Spelling = 0.14
Peng & Fuchs, 2017	7.13 years old/ Grade 1	At-risk readers	Randomized control trial 1) Business-as-usual (BAU), n = 202) Domain-general WM training (DG-WM) , n = 193) Domain-specific Approach of WM training (DS-WM) , n = 19	N/A	Rehearsal strategy	DG-WM vs. BAU Retell = 0.16 Listening comprehension = 0.65* DS-WM vs. BAU Retell = 0.65* Listening comprehension = 0.63 DS-WM vs. DG-WM Retell = 0.43 Listening comprehension = -0.06	N/A

Note: +: effect sizes were calculated based on the descriptive statistics provided by the study. * $p <. 05$; ** $p < .01$.

strategy during WM tasks may reflect WM performance and partially underlie the relation between WM and other tasks (e.g., Gonthier & Thomassin, 2015; Robison & Unsworth, 2017; Unsworth & Spillers, 2010). There are two models on the role of specific strategy use in WM. The strategy mediation hypothesis suggests that the relation between WM and other tasks may be mediated by the effective strategies used in the tasks (Gonthier & Thomassin, 2015). That is, individuals with high WM tend to use effective strategies in a specific task. Having all individuals using an effective strategy in a specific task reduce the WM demand in the task performance and thus reduce the relation between WM and task performance. In contrast, the strategy affordance hypothesis suggests the relation between WM and other tasks is affected by the similarity of the strategy used in both tasks (Bailey et al., 2008). That is, if the strategies used to improve WM performance also can be used to improve performance on other tasks, then using these strategies in WM tasks is likely to link WM closely to performance on other tasks (Fellman et al., 2020; Laine et al., 2018; Logie, 2012; Peng & Fuchs, 2017). For example, for young at-risk children, teaching and intensively practicing rehearsal strategy in complex span tasks may facilitate the use of such strategy in comprehension tasks where important information can be efficiently rehearsed and remembered, which can help reduce the WM load during comprehension and achieve better comprehension outcomes (Peng & Fuchs, 2017).

Most previous WM training programs and studies did not explicitly teach strategies (Shipstead et al., 2012), but they often promoted the development of strategies spontaneously employed to complete WM tasks (Soveri et al., 2017). Research based on introspective reports and observations from children in WM training suggested that repeated practice on WM tasks promotes the development of idiosyncratic strategies. For example, when asked what they thought had helped them improve, 37% of children with low WM and 67% of children with ADHD reported using strategies after training that included rehearsal and visualization (Holmes et al., 2009; Holmes et al., 2010). Similarly, Peng and Fuchs (2017), based on observations of 1st graders at risk for reading difficulties in a verbal WM training condition (without strategy instruction), found that the children used strategies in 28% of all training trials on average. Of these, 59% involved rehearsal, 32% showed evidence of a counting strategy (such as remembering numbers in numerical WM tasks), 6% included a visual strategy (to air draw the to-be-remembered number or words), and 3% involved a semantic strategy (to semantically link to-be-remembered words together).

Moreover, prior WM training studies suggested that the near-transfer effects may be linked to task-specific strategy (Chooi & Logie, 2020; Forsberg et al., 2020). For example, in their *meta*-analysis, Soveri et al. (2017) focused on *n*-back task training and specifically separated WM outcomes with different task paradigms: *n*-back, simple span, and complex span. They found a moderate effect of task-specific transfer to untrained *n*-back tasks, $g = 0.62$, 95CI [0.44, 0.81], and smaller transfer effects to other untrained WM measures, $g = 0.24$, 95CI [0.16, 0.32]. The authors argued that task-specific WM training improvement can only enhance performance on tasks with a similar structure where the same strategies can be successfully employed. Such strategy mediation effects have been more directly examined and observed in most recent WM training studies among adults, where strategy instruction on WM tasks significantly improved performance on the same and similarly-structured WM tasks, compared to WM training without strategies instruction (Forsberg et al., 2020; Malinovitch et al., 2021). Yet, sufficient WM training without strategies instruction ultimately lead to the discovery of efficient strategy taught in the WM strategy training group (Fellman et al., 2020). All these evidence, taken tougher, suggest the transfer of WM training to structurally similar non-trained tasks may be due in part to the same strategy use in both training and non-trained tasks and due to the efficiency of such strategy use in a WM-demanding condition (e.g., Bailey et al., 2008; Fellman et al., 2020; von Bastian & Oberauer, 2013).

Review of studies

Recently, there are some WM training studies in line with the domain-specific approach of WM training. Next, we reviewed these articles to illustrate how WM-academic tasks and task-linking strategies may work in an educational context. Specifically, articles for review were identified in two ways. First, a computer search of published articles and dissertations in Education Resources Information Center (ERIC), MEDLINE, and PsychInfo for literature was conducted. We used the earliest possible start date till January 2021. Abstracts were searched for the following terms: ("working memory" OR "executive function") AND ("training" OR "intervention") AND ("academic" OR "reading" OR math* OR "strategy"). * can help include different forms of search terms (e.g., math* can include mathematics). Second, we searched in previous relevant WM intervention reviews/*meta*-analysis. The initial search yielded 1,917 articles. The first author and a graduate student in education then reviewed all studies by titles and abstracts. After excluding 391 duplicate articles and 664 irrelevant articles, the remaining 862 articles were closely reviewed using two specific criteria: First, studies must have samples as school-aged students (before grade 12) and must have academic tasks as part of the outcomes. Based on this criterion, we excluded 187 articles. Second, studies have to be quasi-experimental or randomized control trials that included the training components tapping WM-academic tasks (performing academic tasks in a complex span paradigm or performing academic tasks with the central executive task paradigm such as inhibition and switching) or strategy (task-specific strategies such as rehearsal or *meta*-cognitive strategies) instruction/training in WM tasks. Based on this criterion, we excluded another 666 articles.

From the literature, we have identified and reviewed 9 randomized controlled trials of interventions tapping WM-academic tasks or task-linking strategy during WM training and including academic outcomes among school-aged children. Instead of doing a systematic review or *meta*-analysis, we have reviewed these studies here to serve as a proof-of-concept for the domain-specific approach of WM training, in order to provide examples of WM-academic task training and task-linking strategy training across various academic areas. Effect sizes and p values (if significant) were provided to indicate the magnitude of effects of interventions. The summary of each reviewed study is also presented in Table 1.

WM-academic task training

Several reading and mathematics intervention studies adopted complex span paradigms to teach students how to efficiently use WM and relevant executive functions (e.g., updating) with the retrieval of relevant information about the academic tasks from long-term memory (Fuchs et al., 2018; García-Madruga et al., 2013; Goodrich et al., 2020; Kroesbergen et al., 2014; Swanson, 2016). For example, in their randomized controlled trial with 31 students in the intermediate grades (8- to 9-year-olds; 50 min per session for 12 sessions over 4 weeks), García-Madruga et al. (2013) embedded WM within reading comprehension instruction. In one complex reading span task, students silently read anaphora and analogy problems presented on a screen, actively recalled the word solution for each anaphora or analogy problem, and then wrote the words down in order. In a WM task with reading and updating, students read different texts with a stream of information in which the relevant facts constantly changed. Students kept track of the changing information during reading, and at several points in the story, they evaluated different aspects of the story at that time (e.g., the order of the horses in a race, the score on a scoreboard in a football match). García-Madruga et al. found that such WM embedded reading comprehension instruction produced effects on reading comprehension (standardized measures) in comparison to the active controls (Cohen's $d = 0.79$).

In a randomized controlled trial of a reading intervention with 116 struggling readers in grades 3 and 5 over 14 weeks (42 sessions, 20–45 min per session), Fuchs et al. (2018) tested WM-reading task training using a three-group design: a WM-reading training group, a reading instruction-only group, and a business-as-usual control group. The WM-reading training group received the same reading instructions and instruction time as did the reading instruction-only group. However, the WM-reading training group received about 5 $\sim$ 10 min (based on the authors' intervention log) devoted to WM-reading task training in each session. Three WM-reading tasks were used. In a "What If?" task, students read a story, retold it, and retold it again with a changed ending, which required updating story information in WM. In a cloze span task, students first completed the cloze task and then recalled in order words that they selected to fill in the blanks in the cloze task. In a WM-main idea task, after summarizing the main idea of each paragraph of a story, students recalled in order all main ideas from previous paragraphs together with the one that they had just created. Because of the limited time devoted to these WM-reading tasks in each session, the training on these tasks was not adaptive.

Fuchs et al. (2018) did not find significant treatment effects on standardized reading measures, but most effect sizes for treatment groups versus controls were educationally meaningful (Hedges' $g \geq 0.25$; Lipsey et al., 2012). Moreover, among 3rd graders, both of Fuchs et al.'s (2018) treatment groups significantly outperformed controls on researcher-developed vocabulary and on reading comprehension (Hedges' $g = 0.31$–1.46). Effect sizes were larger for the WM-reading training group versus controls on researcher-developed vocabulary and reading comprehension (except on reading comprehension among 5th graders) than for the reading instruction-only group versus controls on those measures. Note too that students in the reading instruction-only group practiced on more reading materials than did those in the WM-reading training group, who needed to spend time working on memory activities (e.g., recalling words or main ideas). Thus, the relatively larger (although not statistically significant) effects favoring the WM-reading training group on some reading measures provide promising evidence that for at-risk readers, WM-reading task training may be a more efficient way to improve reading skills than traditional reading instruction or drill-and-practice.

Goodrich et al. (2020) conducted a randomized controlled trial among low-performing preschool English language learners. Sixty-nine children were randomly assigned to three conditions: WM-reading training, reading instruction-only, and business-as-usual control. Students in the treatment groups completed instructional sessions lasting approximately 25 min per day (4 days each week) for 7 weeks. Every week, the first day and the third day were devoted to reading instruction only. The second day and the fourth day were devoted to WM-reading training using the same materials from the prior reading session (e.g., the WM-reading training materials on the second day were from the reading materials on the first day). With this design, Goodrich et al. (2020) purposefully exercised students' ability to retrieve relevant long-term memory information during the WM-reading tasks. The reading instruction and instruction time were the same for the two treatment groups, except that some instructional time was devoted to WM-reading task training in the WM-reading training group. WM-reading tasks used the complex span paradigm with adaptive training and instruction in rehearsal strategy. The WM-reading training tasks required children to complete a "trial" of a designated early literacy activity (e.g., identifying the initial sound/letter of a word; matching lower and upper case letters), rehearse the relevant word/letter/sound used during that trial, complete the next trial of the early literacy activity, and then recall the relevant words/letters/sounds in order from each trial. If students succeeded in recalling two words/letters/sounds, the rehearsal procedure continued with a third word/letter/sound being introduced, followed by an attempted recall of all three words/letters/sounds. This process continued until children did not successfully recall all relevant words/letters/sounds.

The treatment groups in Goodrich et al. (2020) significantly outperformed controls on standardized vocabulary and syntax measures (Hedges' $g = 0.67$ - 1.30), and the WM-reading training group also showed higher performance on syntax (a WM demanding task) than the instruction-only group (Hedges' $g = 0.78$). An aptitude-by-treatment interaction was reported, such that WM-reading training was most effective on reading outcomes in comparison with controls among students with relatively high pretreatment WM abilities. This study systematically tapped both WM-reading task training and task-linking strategy training in an adaptive way, and as in Fuchs et al. (2018), the WM-reading training group spent less time on reading instruction (less exposure to reading materials) due to practice on memory activities in comparison with the reading instruction-only group. Thus, Goodrich et al. (2020) have provided promising evidence in support of the domain-specific approach of WM training, given a small sample.

Besides reading intervention, we have found two studies in mathematics intervention that tapped the domain-specific approach of WM training. Kroesbergen et al. (2014) randomly assigned 51 five-year-old children with below-average mathematics skills to three study groups: domain-general WM training (with non-numerical materials), domain-specific WM training (with numerical materials), and a business as usual control. The training included eight 30-min sessions across 4 weeks. The two WM training activities shared similar training paradigms and training dosage, except that the domain-specific WM focused on exercising numerical skills and retrieving numerical information from long-term memory (e.g., on simultaneously processing and remembering numbers). For example, in a "Counting Recall" activity, children saw a set of triangles and circles on a screen, counted the circles, remembered the count for each set, and recalled the (count) number from each set in order. In a "linear board game," children took turns to throw a die with the numbers one, two, and three and remembered the number they threw. When all the children had thrown the die, they recalled the number they threw and attempted to take the right number of steps on the board. In a "Memory" activity, on one side of a table, there were cards with different amounts of dots; on the other side of the table, there were cards with numbers. Children took turns and turned one card with dots and one card with a number each time to find matched pairs. When matched, the child could take the cards, but the cards had to be turned back if not matched. The domain-specific WM group significantly outperformed the control group on the standardized numeracy measure (Hedges' $g = 0.93$), and outperformed (although not statistically significant) the domain-general WM group on both researcher-developed and standardized numeracy measures (quantity discrimination and various complicated counting skills) that tap numerical WM heavily among young children (Kroesbergen et al., 2009) (Hedges' $g = 0.49$ - 0.53).

Swanson (2016) studied the effects of systematically increasing WM load embedded in word problem solving instruction that practices the retrieval of problem-solving schema from the long-term memory. This randomized control trial with 162 elementary school children with mathematics difficulties included 5 conditions: (1) practice with increasing WM load only (practice in the incremental presentation of irrelevant propositions during word problem solving without explicit strategy instruction), (2) practice with verbal emphasis (explicit verbal strategies for word problems) and increasing WM load, (3) practice with visual emphasis (explicit diagrams representing strategies) and increasing WM load, (4) practice with verbal + visual emphasis and increasing WM load, and (5) business-as-usual control. In comparison to the business-as-usual control, the practice with increasing WM load only group for those with relatively higher WM capacity showed the strongest effect size on word problem performance (Hedges' $g = 0.47$, $p <.05$) and calculations (Hedges' $g = 0.80$, $p <.05$), whereas those with lower WM capacity performed better in the control group. There was an aptitude-by-treatment effect. In those with relatively lower pretreatment WM capacity, simply practicing verbal and/or visual strategies or practicing word problems with increasing WM load may overload WM capacity and compromise performance on complex/WM-demanding word problems. In contrast, in children with relatively higher pretreatment WM capacity, simply practicing verbal and/or visual strategies also seemed to overload WM capacity, but practicing world problems with increasing WM load (WM-word problem tasks) helped them allocate WM capacity more efficiently to achieve better word-problem solving.

To sum, the above-mentioned studies adopted the WM-academic tasks. The common pattern across these WM-academic tasks is using WM complex span as the paradigm, in which students practiced academic skills (e.g., summarizing main ideas; counting numbers) and tried to remember the outcome of those academic practice by retrieving this information from long-term memory (e.g., the main ideas; the number from counting). Yet, different studies placed a different emphasis on those WM-academic tasks. While some studies involved a lot of practice on those WM-academic tasks with adaptive rules (García-Madruga et al., 2013; Goodrich et al., 2020; Swanson, 2016), other studies only devoted a small amount of time to those WM-academic tasks without adaptive rules (Fuchs et al., 2018).

Task-linking strategy training

Research on WM training tends to focus on either of two major sets of strategies: training on metacognitive strategies such as planning, organizing, monitoring, and evaluation, and training on task-specific strategies.

Jones et al. (2020) and Partanen et al. (2015) trained school-aged children to use metacognitive strategies during WM training, using a randomized controlled design. Jones et al. randomly assigned 95 typically developing children aged 9–14 years to a WM training group, WM + metacognitive strategy training group, or a control group. In the WM + metacognitive strategy group, children were explicitly taught to think about their thinking as they completed WM training tasks or academic tasks, and to plan, monitor, and evaluate specific metacognitive strategies that served to self-motivate and refocus. For example, during training, questions prompted children to plan at the beginning, reminded them to monitor thoughts during training and required them to self-evaluate at the end. The questions tapped the goal of the task, the generation of task-specific strategies (although not specifically mentioned), steps needed to complete the task, strategies to stay focus, and evaluation of strategies. As children progressed in the training, the questions were replaced with prompting reminders of how to plan, monitor and evaluate. Jones et al. (2020) found that the WM training group and the WM plus *meta*-cognitive strategy training group significantly outperformed the control group (visual search training) on non-trained WM tasks and mathematics in post-intervention and follow-up points (Hedges' $g = 0.24$ - 0.65), with larger effect sizes associated with WM plus *meta*-cognitive strategy training group. In particular, the authors argued that such *meta*-cognitive strategies prompted children to be aware and develop task-specific strategies that can help improve the efficiency use of their WM resources, which is in line with the strategy mediation hypothesis.

Partanen et al. (2015) adopted the same randomized controlled design with the same three groups in 64 children with special needs (average mean age is 8 years old). Partanen et al. (2015) used questions as prompts to facilitate students' reflections on cognitive and metacognitive strategies during WM training. The questions guided the students to pay attention to WM training task characteristics, formulate the goal for performance on the task, and identify and formulate strategies for both success and failure on the task. Partanen et al. (2015) found WM plus *meta*-cognitive strategy training significantly outperformed the WM training group on non-trained WM tasks in post-intervention (Hedges' $g = 0.77$, $p < .05$) but no significant training effects were found on academic skills. Similar to Jones et al. (2020), Partanen et al. (2015) acknowledged that such *meta*-cognitive strategies inevitably lead to "task-oriented" strategy use that can add to the training effects.

In contrast to metacognitive strategies, task-specific strategies are mentally effortful goal-directed processes that have also been found to enhance WM performance (Oberauer, 2019; Peng & Fuchs, 2017). One of the most common WM-specific strategies is rehearsal, which involves the repetition of to-be-remembered information (Baddeley, 1999; Oberauer, 2019; Sadoski et al., 2000; Turley-Ames & Whitfield, 2003). Although rehearsal is traditionally considered a mnemonic tool for memory or WM tasks, it is also an important task-specific strategy commonly applied in academic tasks.

Rehearsal, an important component of the phonological loop, plays an important role in early phonological processing and decoding development (Baddeley et al., 1986). Children often use rehearsal to keep sounds in the phonological loop so that they can manipulate and blend sounds in phonological processing and decoding tasks, and such use of rehearsal plays a crucial role in learning the novel phonological forms of new words (Baddeley et al., 1998; Cunningham et al., 2021). Rehearsal is also important for early reading comprehension development. That is, rehearsal directly supports the rereading strategy widely adopted by young and old readers. Through reading a sentence or text over and over, students not only practice reading fluency but are more likely to remember more of the sentence or text's content, which facilitates the use of WM to integrate information (from the sentence or text and from long-term memory) and to achieve better reading comprehension (Marley et al., 2010).

Moreover, rehearsal during WM tasks can greatly facilitate the central executive of WM. Research has demonstrated that the central executive is heavily involved in short-term storage of information when more information is being rehearsed (Unsworth & Engle, 2007). Under a WM training condition, rehearsing a long list of to-be-remembered items is an effortful information processing task that requires the central executive (Peng & Fuchs, 2017; Unsworth & Engle, 2007). Salmi et al. (2018) conducted a neuroimaging *meta*-analysis on WM training and found that only dorso- and ventrolateral prefrontal cortex showed consistent training-related modulations specific to WM training, possibly due to their well-established role in short-term storage and rehearsal. This suggests that rehearsal may be heavily involved during (verbal) WM training and maybe a critical factor associated with WM improvement from WM training.

We found two WM intervention studies with school-aged children that specifically trained rehearsal strategies during WM training. One is that of Goodrich et al. (2020), who explicitly had children practice rehearsal strategies on WM-demanding reading tasks in their embedded WM reading instruction group. Goodrich et al. found that such training significantly improved several reading outcomes even in contrast to an active control group (i.e., reading instruction only). However, without including an embedded WM reading instruction group that did not have rehearsal instruction, Goodrich et al. did not demonstrate the unique effects of rehearsal instruction added to the WM training. In contrast, Peng and Fuchs (2017) used a randomized controlled trial with 58 at-risk first-grade readers and examined whether rehearsal strategy instruction would add extra benefits to WM training. Children were randomly assigned to three groups: WM training with rehearsal strategy instruction, WM training without strategy training, and business-as-usual control. The rehearsal strategy was taught and practiced in an adaptive WM complex span paradigm for 10 sessions (30 min per session on 10 consecutive days). The authors found that in comparison with the WM training without strategy training and the business-as-usual control, the WM training with strategy instruction group showed better performance on listening comprehension and retell outcomes (non-standardized measures, Hedges' $g = 0.43$ - 0.65, $ps < 0.05$). The authors argued that improved WM capacity together with fluent use of rehearsal in high WM-demanding conditions may have enabled children to retain information on listening comprehension and retell tasks.

Taken together, for studies on *meta*-cognitive strategies during WM training, the effects on non-trained WM tasks seemed to be limited only to structurally similar WM tasks (e.g., near-transfer effects on visuospatial WM tasks, as a result of training on visuospatial WM tasks), which suggests that *meta*-cognitive strategies instruction during WM training likely facilitated children's awareness on task-specific strategies and improve the efficiency of applying these task-specific strategies to structurally similar non-trained WM tasks. For task-specific strategy training, strategies afforded by both WM tasks and academic tasks seemed to produce stronger effects on academic outcomes. These findings are in line with the strategy affordance hypothesis and strategy mediation hypothesis, suggesting the transfer of WM training to structurally similar non-trained tasks may be due in part to the same strategy use in both training and non-trained tasks and also due to the efficiency of such strategy use in a WM-demanding condition (e.g., Bailey et al., 2008; Fellman et al., 2020; von Bastian & Oberauer, 2013). That said, all those task-linking strategies in WM training aimed to reduce the cognitive load of WM/WM-academic tasks, as to save cognitive resources for the knowledge retrieval from long-term memory and maintaining information to facilitate WM/WM-academic tasks performance.

Summary of the review

Across all the reviewed studies, researchers tried to 1) train students how to efficiently retrieve and utilize relevant information from long-term memory during the WM-academic tasks under a WM-demanding situation, and 2) train students with strategies such as rehearsal and *meta*-cognitive strategies to reduce the WM load during the WM-academic tasks as to aid the information retrieval, combining the retrieved information with information extracted from WM-academic tasks, and maintaining the new information (the integration of information from the long-term memory and information from the current WM-academic tasks). Based on the (adaptive) training on these processes, students are expected to efficiently use strategies to maintain information and facilitate the role of long-term memory in WM-demanding academic tasks as to acquire new academic information in a more general learning setting.

Recommendations for future work

In this theory paper, we have discussed the domain-specific approach of WM training to emphasize explicit instruction and practice on academic tasks with a WM training paradigm and task-linking strategies that can be applied to academic tasks and WM tasks. In comparison to the domain-general WM training, domain-specific WM training is developmental in nature and more educationally relevant. More importantly, using WM along with the domain-specific materials can make the material learning and the retrieval of long-term memory knowledge more effortful, which can create a helpful "generation effect" that is known to boost the long-term learning (Blair, 2002; Diamond, 2013; Liew, 2012). Although our review suggested promising effects for this approach, we noted several points in the following about how future studies can further refine and validate this approach.

Study Design. Future research should include randomized control trials, with a focus on improving academic outcomes among school-aged student samples (at-risk or typically developing children in K–12). Ideally, there could be five study groups: academic instruction-only, the domain-general WM training, WM-academic training (based on the WM-academic task training and task-linking strategy training), a business-as-usual control, and domain-general WM training + academic instruction. The academic content and instructional time should be equated among treatment groups, and information on relevant academic skills instruction in classrooms should be collected for all groups to control for counterfactual effects (Lemons, et al. 2014). Treatment fidelity should be closely documented. For the WM-academic training group, children's performance on each WM-academic task and strategy use in each session should be documented to closely monitor whether children are mastering knowledge of the WM-academic tasks and task-linking strategies (Peng & Fuchs, 2017). WM-academic task performance can also be collected in other study groups to investigate WM-academic task performance changes across study groups explain near-or far-transfer effects (as mentioned in the mediator section below).

Further, fade-out effects are quite common in academic interventions and often occurs among those with learning difficulties (Bailey et al., 2016), likely due to many factors including un-sustained learning environment after interventions and that the trained skills not strongly transferring to later academic skills learning (Bailey et al., 2020). Moreover, the "vicious circle" theory between learning disabilities and cognitive deficits further highlights the gap in WM between children with learning difficulties and their typically developing peers widens across grades. It is possible that the deficits in retrieving academic information in long-term memory, together with domain-general attention control in WM, and possible insufficient instruction contribute to the consistent low performance among individuals with learning disabilities (Peng et al., 2022). Thus, our proposed domain-specific WM training may provide students with more cognitive strategies and skills to retrieve and utilize relevant background knowledge in WM-demanding academic tasks in general learning setting after interventions (Doebel, 2020; Duke & Cartwright, 2021; Peng & Goodrich, 2020). Unfortunately, among those studies we reviewed here, none measured long-term follow-up effects. Follow-up testing is needed to examine the possible fade-out effects in the future.

Academic outcomes. The academic outcomes and training skills for WM-academic tasks should be WM-demanding academic skills across reading or mathematics. That said, any type of academic skill can be WM demanding if the skill is novel to children (e.g., even simple calculation problems or word-reading tasks can be quite WM demanding for young children who are just starting to learn those skills; Peng et al., 2018; Peng et al., 2016). Thus, the choice of academic outcomes and training skills should be appropriate for the sample's level of academic development, with materials that can activate and facilitate the retrieval of knowledge from long-term memory. Ideally, the trained academic skills should be the academic outcomes, with structurally similar tasks as potential near-transfer outcomes (e.g., reading comprehension measures that tap summarization and retell can be near-transfer measures for training on WM-summarization or WM-retell). Through long-lasting training for various WM-demanding subskills in a specific academic domain (e.g., word recognition, language comprehension, reading comprehension in the literacy domain), far-transfer effects may be more likely on standardized measures that require children to efficiently apply WM to retrieving, monitoring, and manipulating various WM-demanding academic skills simultaneously.

Unlike reading, with its two major instructional components (decoding and reading comprehension; Hoover & Gough, 1990), mathematics implicates many more instructional components (e.g., numeracy skills, calculation, fraction, algebra, word problem solving, and geometry; Common Core State Standards Initiative, n.d.). These mathematic skills are usually taught sequentially at different stages through K–12 and become more integrated in later grades. Further research should explore the feasibility and validity of designing instruction for various mathematics skills within the domain-specific approach of WM training. One way is to facilitate the efficient use of WM for mathematics skills that require language comprehension such as word problem instruction (Fuchs, 2020). Another approach is to focus on the use of WM in mathematics information retrieval from long-term memory such as combining complex span paradigm with on mental calculations with multiple addends, subtrahends, and multipliers.

Mediators. We propose four possible mediator variables to explain advantageous effects (if any) of the domain-specific approach of WM training: (1) possible changes on the domain-general WM measures (e.g., changes on WM measures that are not structurally similar to WM-academic tasks or academic outcomes), (2) possible changes on the domain-specific WM measures (e.g., changes on WM measures that are structurally similar to WM-academic tasks or academic outcomes), (3) possible changes on task-linking strategy use efficiency (e.g., changes on rehearsal efficiency on trained or untrained tasks), and (4) changes on trained WM-academic tasks, which can be used as progress monitoring measures for near-transfer training effect. Evidence supporting the domain-specific approach of WM training would be even stronger if these variables (single or combined) explain the treatment effects (WM-academic + Task-linking strategy training vs. academic instruction/domain-general WM training/domain-general WM training + academic instruction/business-as-usual controls) on trained academic skills or structurally similar academic outcomes.

More diverse strategies support. Further research should explore the effects of diverse strategy training within the domain-specific approach of WM training (Brod, 2020). Studies with both children and adults suggest that during WM training, individuals in general are prone to use metacognition to find and apply the most efficient task-specific strategies (e.g., Fellman et al., 2020; Jones et al., 2020; Peng & Fuchs, 2017). Thus, combining metacognitive and task-specific strategy training might facilitate WM-academic tasks performance (e.g., students would be directly taught and guided to use metacognitive strategies to monitor task-specific strategies during training). In addition, the strategy-linking model of WM suggests strategies be used in both academic tasks and WM tasks. Given the academic instructional feature of the domain-specific approach of WM training, it is important to explore evidence-based strategies used in reading or mathematics instruction and how those strategies can be used in a WM training paradigm. For example, there is also a growing body of research that suggested individuals who spend time using elaborative strategies on processing/encoding maintained items in a meaningful way are more likely to retrieve relevant information from long-term memory during WM tasks (Jarjat et al., 2018; Loaiza & Lavilla, 2021; Souza & Oberauer, 2017). This evidence suggests elaborative strategies such as semantic linking (linking to-be-remembered items in a meaningful way) can best activate and link information from long-term memory with information extracted from the WM-academic tasks.

Meanwhile, it is important to consider the strategy use from a developmental perspective. Strategy mediation hypothesis categorized strategies used in WM tasks into effective ones and ineffective ones (Gonthier & Roulin, 2020). However, the effectiveness of strategies vary across individuals and development (Alexander et al., 1998). For example, for young children who are still developing cognitive abilities (with relatively lower WM capacity), simple strategies such as rehearsal are often adopted to reduce WM load as to retain more information during learning (Ornstein et al., 1988). For adults, there are various strategies with different levels of complexity and the complexity of strategies is partly determined by long-term memory of the task performance. That is, if individuals are more familiar with the tasks, they will have more WM available to use more complicated and potentially more effective strategies. If individuals are not familiar with the tasks, their WM devoted to complex strategy use is limited and they are more likely to use simpler but less effective strategies (Gonthier & Roulin, 2020). Thus, the choice of the task-linking strategies is developmental and should take into account individual's general WM capacity, knowledge/fluency of strategies, and knowledge on the academic skills in the WM-academic training tasks simultaneously.

Aptitude-by-treatment effects. Prior research on WM training has often reported aptitude-by-treatment effects. For example, several studies, together with the ones reviewed here, indicated that both WM training without explicit academic content instruction and WM training embedded within explicit skill instruction seem to benefit children with stronger general cognitive abilities (e.g., higher nonverbal intelligence or higher WM capacity) on both trained WM and untrained cognitive or academic tasks (e.g., Gathercole et al., 2019; Swanson, 2016; Swanson & McMurran, 2018; Goodrich et al., 2020). However, the cognitive–academic bidirectional model suggests that at-risk children, in comparison with typically developing peers, are less likely to trigger cognitive–academic bidirectionality to improve both cognition and academic skills even with academic instruction only (Peng & Kievit, 2020). Thus, future research should further investigate how to make the domain-specific approach of WM training more beneficial for children with learning disabilities or/and lower general cognitive abilities.

Declaration of Competing Interest

This research was supported by R324A200165 from the Institute of Education Sciences; We would like to thank Dr. Michael J. Kane, Dr. Robert Logie, Dr. David C. Geary, and anonymous reviewers for their cosntructive feedback on the previous versions.

Data availability

Data is included in the manuscript

References

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*(2), 129–154.

Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading, 18*(5), 325–346. https://doi.org/10.1080/10888438.2014.902461

Baddeley, A. [D.] (1986). *Working memory*. Clarendon Press/Oxford University Press.

Baddeley, A., Logie, R., Bressi, S., Sala, S. D., & Spinnler, H. (1986). Dementia and working memory. *The Quarterly Journal of Experimental Psychology Section A, 38*(4), 603–618.

Baddeley, A. [D.] (1992). Working memory. *Science, 255*(5044), 556-559. 10.1126/science.1736359.

Baddeley, A. D. (1999). *Essentials of human memory* (Cognitive Psychology: A Modular Course). Taylor & Francis.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158–173. https://doi.org/10.1037/0033-295X.105.1.158

Baddeley, A., Allen, R. J., & Hitch, G. J. (2010). Investigating the episodic buffer. *Psychologica Belgica, 50*(3–4), 223–243. https://doi.org/10.5334/pb-50-3-4-223

Bailey, D. H., Nguyen, T., Jenkins, J. M., Domina, T., Clements, D. H., & Sarama, J. S. (2016). Fadeout in an early mathematics intervention: Constraining content or preexisting differences? *Developmental Psychology, 52*, 1457–1469. https://doi.org/10.1037/dev0000188

Bailey, D. H., Duncan, G. J., Cunha, F., Foorman, B. R., & Yeager, D. S. (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest, 21*(2), 55–97. https://doi.org/10.1177/1529100620915848

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition, 36*(8), 1383–1390. https://doi.org/10.3758/MC.36.8.1383

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111–127. https://doi.org/10.1037//0003-066X.57.2.111

Brod, G. (2020). Generative learning: Which strategies for what age? *Educational Psychology Review*. Advance online publication. 10.1007/s10648-020-09571-9.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Chooi, W. T., & Logie, R. H. (2020). Changes in error patterns during N-back training indicate reliance on sub-vocal rehearsal. *Memory and Cognition, 48*, 1484–1503. https://doi.org/10.3758/s13421-020-01066-w

Common Core State Standards Initiative. (n.d.). *Common core state standards for mathematics*. Retrieved from http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf.

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (Eds.). (2007). *Variation in working memory*. Oxford University Press.

Cowan, N., Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

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

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433. https://doi.org/10.3758/BF03214546

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology, 103*(2), 186–201. https://doi.org/10.1016/j.jecp.2009.01.004

Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science, 15*(4), 942–956. https://doi.org/10.1177/1745691620904771

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly. Advance online publication.*. https://doi.org/10.1002/rrq.411

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science, 13*(2), 190–193. https://doi.org/10.1177/1745691617720478

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145–199). Elsevier Science.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Ericsson, K. A., Hoffman, R. R., Kozbelt, A., & Williams, A. M. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.

Fellman, D., Jylkkä, J., Waris, O., Soveri, A., Ritakallio, L., Haga, S., … Laine, M. (2020). The role of strategy use in working memory training outcomes. *Journal of Memory and Language, 110*, Article 104064. https://doi.org/10.1016/j.jml.2019.104064

Forsberg, A., Fellman, D., Laine, M., Johnson, W., & Logie, R. H. (2020). Strategy mediation in working memory training in younger and older adults. *Quarterly Journal of Experimental Psychology, 73*(8), 1206–1226. https://doi.org/10.1177/1747021820915107

Friso-Van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General, 129*(1), 61–83. https://doi.org/10.1037/0096-3445.129.1.61

Fuchs, D., Hendricks, E., Walsh, M. E., Fuchs, L. S., Gilbert, J. K., Tracy, W. Z., … Peng, P. (2018). Evaluating a multidimensional reading comprehension program and reconsidering the lowly reputation of tests of near-transfer. *Learning Disabilities Research & Practice, 33*(1), 11–23. https://doi.org/10.1111/ldrp.12162

Fuchs, L. S. (2020, February). *Contrasting the effects of working memory training to direct skills intervention on word-problem solving* [Paper presentation]. 28th Annual Pacific Coast Research Conference, Coronado, CA.

García-Madruga, J. A., Elosúa, M. R., Gil, L., Gómez-Veiga, I., Vila, J.Ó., Orjales, I., … Duque, G. (2013). Reading comprehension and working memory's executive processes: An intervention study in primary school students. *Reading Research Quarterly, 48*(2), 155–174. https://doi.org/10.1002/rrq.44

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A.-M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*(3), 265–281. https://doi.org/10.1016/j.jecp.2005.08.003

Gathercole, S. E., Dunning, D. L., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language, 105*, 19–42. https://doi.org/10.1016/j.jml.2018.10.003

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American psychologist, 50*(1), 24–37. https://doi.org/10.1037//0003-066x.50.1.24

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*(2), 121–151. https://doi.org/10.1016/j.jecp.2004.03.002

Gonthier, C., & Roulin, J.-L. (2020). Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition. *Journal of Experimental Psychology: General, 149*(3), 564–579. https://doi.org/10.1037/xge0000660

Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *Journal of Experimental Psychology: General, 144*(5), 916–924. https://doi.org/10.1037/xge0000101

Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young school-age children and its relation to intelligence. *Journal of Memory and Language, 92*, 183–201. https://doi.org/10.1016/j.jml.2016.06.004

Goodrich, J. M., Peng, P., Bohaty, J., Leiva, S., & Thayer, L. (2020). *Embedding executive function training into early literacy instruction for dual* language *learners: A pilot study*. Manuscript submitted for publication.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*(4), 339–387. https://doi.org/10.1006/cogp.2001.0769

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12*(4), F9–F15. https://doi.org/10.1111/j.1467-7687.2009.00848.x

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology, 24*(6), 827–836. https://doi.org/10.1002/acp.1589

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2), 127–160. https://doi.org/10.1007/BF00401799

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Jarjat, G., Hoareau, V., Plancher, G., Hot, P., Lemaire, B., & Portrat, S. (2018). What makes working memory traces stable over time? *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.13668

Jones, J. S., Milton, F., Mostazir, M., & Adlam, A. R. (2020). The academic outcomes of working memory and metacognitive strategy training in children: A double-blind randomized controlled trial. *Developmental Science, 23*(4). https://doi.org/10.1111/desc.12870

Kail, R. V., Lervåg, A., & Hulme, C. (2016). Longitudinal evidence linking processing speed to the development of reasoning. *Developmental Science, 19*(6), 1067–1074. https://doi.org/10.1111/desc.12352

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*(1), 66–71. https://doi.org/10.1037/0033-2909.131.1.66

Kendeou, P., Van Den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice, 29*(1), 10–16. https://doi.org/10.1111/ldrp.12025

Kievit, R. A., Lindenberger, U., Goodyer, I. M., Jones, P. B., Fonagy, P., Bullmore, E. T., the Neuroscience in Psychiatry Network, & Dolan, R. J. (2017). Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychological Science*, *28*(10), 1419–1431. 10.1177/0956797617710785.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kroesbergen, E. H., Van Luit, J. E. H., Van Lieshout, E. C. D. M., Van Loosbroek, E., & Van de Rijt, B. A. M. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment, 27*, 226–236. https://doi.org/10.1177/0734282908330586

Kroesbergen, E. H., van't Noordende, J. E., & Kolkman, M. E. (2014). Training working memory in kindergarten children: Effects on working memory and early numeracy. *Child Neuropsychology*, *20*(1), 23–37. 10.1080/09297049.2012.736483.

Laine, M., Fellman, D., Waris, O., & Nyman, T. J. (2018). The early effects of external and internal strategies on working memory updating training. *Scientific Reports, 8*(1), 1–12. https://doi.org/10.1038/s41598-018-22396-5

Liew, J. (2012). Effortful control, executive functions, and education: Bringing self-regulatory and social-emotional competencies to the table. *Child Development Perspectives, 6*(2), 105–111. https://doi.org/10.1111/j.1750-8606.2011.00196.x

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSER 2013-3000). National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf.

Loaiza, V. M., & Lavilla, E. T. (2021). Elaborative strategies contribute to the long-term benefits of time in working memory. *Journal of Memory and Language, 117*, Article 104205.

Logie, R. H. (2012). Cognitive training: Strategies and the multicomponent cognitive system. *Journal of Applied Research in Memory and Cognition, 1*(3), 206–207. https://doi.org/10.1016/j.jarmac.2012.07.006

Logie, R., Camos, V., & Cowan, N. (2020). *Working memory: The state of the science*. Oxford University Press.

Malinovitch, T., Jakoby, H., & Ahissar, M. (2021). Training-induced improvement in working memory tasks results from switching to efficient strategies. *Psychonomic Bulletin & Review, 28*(2), 526–536.

Mammarella, I. C., Giofrè, D., & Caviola, S. (2017). Learning geometry: The development of geometrical concepts and the role of cognitive processes. In *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 221–246). Academic Press.

Marley, S. C., Levin, J. R., & Glenberg, A. M. (2010). What cognitive benefits does an activity-based reading strategy afford young Native American readers? *The Journal of Experimental Education, 78*(3), 395–417. https://doi.org/10.1080/00220970903548061

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? *A meta-analytic review. Developmental Psychology, 49*(2), 270–291. https://doi.org/10.1037/a0028228

Melby-Lervåg, M., Lervåg, A., Lyster, S.-A.-H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-Repetition Ability Does Not Appear to Be a Causal Influence on Children's Vocabulary Development. *Psychological Science, 23*(10), 1092–1098. https://doi.org/10.1177/0956797612443833

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*(4), 512–534. https://doi.org/10.1177/1745691616635612

Miller-Cotto, D., & Byrnes, J. P. (2020). What's the best way to characterize the relationship between working memory and achievement? An initial examination of competing theories. *Journal of Educational Psychology, 112*(5), 1074–1084. https://doi.org/10.1037/edu0000395

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*(2), 139–158. https://doi.org/10.1006/jecp.1999.2498

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 411–421. https://doi.org/10.1037/0278-7393.28.3.411

Oberauer, K. (2019). Is rehearsal an effective maintenance strategy for working memory? *Trends in Cognitive Sciences, 23*(9), 798–809.

Ornstein, P. A., Baker-Ward, L., & Naus, M. J. (1988). The development of mnemonic skill. *Memory development: Universal changes and individual differences* (pp. 31–50).

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32*(1), 1–8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0

Partanen, P., Jansson, B., Lisspers, J., & Sundin, Ö. (2015). Metacognitive strategy training adds to the effects of working memory training in children with special educational needs. *International Journal of Psychological Studies, 7*(3), 130–140. https://doi.org/10.5539/ijps.v7n3p130

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., … Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*(1), 48–76. https://doi.org/10.1037/bul0000124

Peng, P., & 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. *Journal of Learning Disabilities, 50*(1), 62–80. https://doi.org/10.1177/0022219415594609

Peng, P., & Goodrich, J. M. (2020). The cognitive element model of reading instruction. *Reading Research Quarterly, 55*(S1), 577–588. https://doi.org/10.1002/RRQ.336

Peng, P., & Kievit, R. (2020). The development of academic achievement and cognitive abilities: A bidirectional perspective. *Child Development Perspectives, 14*(1), 15–20. https://doi.org/10.1111/cdep.12352

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108*(4), 455–473. https://doi.org/10.1037/edu0000079

Peng, P., Zhang, Z., Wang, W., Lee, K. J., Wang, T. F., Wang, C. C., … Lin, J. Z. (2022). A meta-analytic review of cognition and reading difficulties: Individual differences, moderation, and language mediation mechanisms. *Psychological Bulletin. Advance online publication.*. https://doi.org/10.1037/bul0000361

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and *n*-back measures of working memory: A meta-analysis. *Psychonomic Bulletin & Review, 20*(6), 1102–1113. https://doi.org/10.3758/s13423-013-0453-9

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology, 51*(5), 573–582. https://doi.org/10.1037/a0038981

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? *A Meta-Analysis. Psychological Science, 29*(8), 1358–1369. https://doi.org/10.1177/0956797618774253

Robison, M. K., & Unsworth, N. (2017). Working memory capacity, strategic allocation of study time, and value-directed remembering. *Journal of Memory and Language, 93*, 231–244. https://doi.org/10.1016/j.jml.2016.10.007

Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology, 92*(1), 85–95. https://doi.org/10.1037/0022-0663.92.1.85

Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science, 26*(6), 515–520. https://doi.org/10.1177/0963721417712760

Sala, G., & Gobet, F. (2020). Working memory training in typically developing children: A multilevel meta-analysis. *Psychonomic Bulletin & Review, 27*(3), 423–424. https://doi.org/10.3758/s13423-019-01681-y

Salmi, J., Nyberg, L., & Laine, M. (2018). Working memory training mostly engages general-purpose large-scale networks for learning. *Neuroscience & Biobehavioral Reviews, 93*, 108–122. https://doi.org/10.1016/j.neubiorev.2018.03.019

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span *n*-back, and other working memory indicators in psychometric context. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01475

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*(2), 138–166. https://doi.org/10.1080/00461520.2015.1036274

Seigneuric, A., Ehrlich, M.-F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing, 13*(1–2), 81–103. https://doi.org/10.1023/A:1008088230941

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*(1), 4–27. https://doi.org/10.1037/0096-3445.125.1.4

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138*(4), 628–654. https://doi.org/10.1037/a0027473

Souza, A. S., & Oberauer, K. (2017). Time to process information in working memory improves episodic memory. *Journal of Memory and Language, 96*, 155–167. https://doi.org/10.1016/j.jml.2017.07.002

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*(4), 1077–1096. https://doi.org/10.3758/s13423-016-1217-0

Stockard, J., Wood, T. W., Coughlin, C., & Khoury, C. R. (2018). The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research, 88*(4), 479–507. https://doi.org/10.3102/0034654317751919

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30*(3), 261–288. https://doi.org/10.1016/S0160-2896(01)00100-3

Swanson, H. L. (2016). Word problem solving, working memory and serious math difficulties: Do cognitive strategies really make a difference? *Journal of Applied Research in Memory and Cognition, 5*(4), 368–383. https://doi.org/10.1016/j.jarmac.2016.04.012

Swanson, H. L., Martin, A. J., Sperling, R. A., & Newton, K. J. (Eds.). (2020). *Handbook of educational psychology and students with special needs* (pp. 19–51). Routledge.

Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook: Vol 1. Theories, constructs, and critical issues* (pp. 327–366). American Psychological Association. 10.1037/13273-012.

Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research, 76*(2), 249–274. https://doi.org/10.3102/00346543076002249

Swanson, H. L., & McMurran, M. (2018). The impact of working memory training on near and far transfer measures: Is it all about fluid intelligence? *Child Neuropsychology, 24*(3), 370–395. https://doi.org/10.1080/09297049.2017.1280142

Sweller, J. (2021). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review*. https://doi.org/10.1007/s10648-021-09647-0

Sweller, J., van Merriënboer, J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292. https://doi.org/10.1007/s10648-019-09465-5

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120*(3), 439–471. https://doi.org/10.1037/a0033138

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49*(4), 446–468. https://doi.org/10.1016/S0749-596X(03)00095-0

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition, 41*(2), 242–254. https://doi.org/10.3758/s13421-012-0261-x

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*(6), 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language, 62*(4), 392–406. https://doi.org/10.1016/j.jml.2010.02.001

Unterrainer, J. M., Kaller, C. P., Halsband, U., & Rahm, B. (2006). Planning abilities and chess: A comparison of chess and non-chess players on the Tower of London task. *British Journal of Psychology, 97*(3), 299–311. https://doi.org/10.1348/000712605X71407

Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Verbal and visual-spatial working memory and mathematical ability in different domains throughout primary school. *Memory & Cognition, 43*(3), 367–378. https://doi.org/10.3758/s13421-014-0480-4

von Bastian, C. C., & Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *Journal of Memory and Language, 69*(1), 36–58. https://doi.org/10.1016/j.jml.2013.02.002

Wang, S., & Gathercole, S. E. (2013). Working memory deficits in children with reading difficulties: Memory span and dual task coordination. *Journal of Experimental Child Psychology, 115*(1), 188–197. https://doi.org/10.1016/j.jecp.2012.11.015

Waters, A. J., Gobet, F., & Leyden, G. (2002). Visuospatial abilities of chess players. *British Journal of Psychology, 93*(4), 557–565. https://doi.org/10.1348/000712602761381402

Willoughby, M. T., Wylie, A. C., & Little, M. H. (2019). Testing longitudinal associations between executive function and academic achievement. *Developmental Psychology, 55*(4), 767. https://doi.org/10.1037/dev0000664

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. (I). *Psychological Review, 8* (3), 247–261. https://doi.org/10.1037/h0074898

Zhang Z., & Peng, P. (2022). *The investigation of co-development among reading, executive function, and social emotional skills.* https://osf.io/mhdpx/?view_only=8963582fd5584ad28e78f6398c48e956.

Zsambok, C. E., & Klein, G. (Eds.). (2014). *Naturalistic decision making* (eBook ed.). Taylor & Francis/Psychology Press. 10.4324/9781315806129.

Further reading

Donchin, E., Fabiani, M., & Sanders, A. (1989). The learning Strategies Program: An examination of the strategies in skill acquisition. *Acta Psychologica, 71*, 1–312.

Feldon, D. F., & Litson, K. (2021). Modeling Theories and Theorizing Models: An Attempted Replication of Miller-Cotto & Byrnes'(2019) Comparison of Working Memory Models Using ECLS-K Data. *Educational Psychology Review, 1–28*. https://doi.org/10.1007/s10648-021-09596-8

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*(4), 216–227. https://doi.org/10.2307/1593674

Geary, D., Berch, D., Geary, D., & Berch, D. (Eds.). (2016). *Evolutionary perspectives on child development and education* (pp. 217–249). Basel: Springer.

Kovacs, K., Molenaar, D., & Conway, A. R. A. (2019). The domain specificity of working memory is a matter of ability. *Journal of Memory and Language, 109*, Article 104048. https://doi.org/10.1016/j.jml.2019.104048

Lervåg, A. (2020). Some roads less travelled—different routes to understanding the causes of child psychopathology. *Journal of Child Psychology and Psychiatry, 61*(6), 625–627. https://doi.org/10.1111/jcpp.13274

Mammarella, I. C., Caviola, S., Giofrè, D., & Szűcs, D. (2018). The underlying structure of visuospatial working memory in children with mathematical learning disability. *British Journal of Developmental Psychology, 36*(2), 220–235. https://doi.org/10.1111/bjdp.12202

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition, 29*(1), 10–17. https://doi.org/10.3758/BF03195736

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press. 10.1017/CBO9781139174909.

Morrison, A. B., Rosenbaum, G. M., Fair, D., & Chein, J. M. (2016). Variation in strategy use across measures of verbal working memory. *Memory & Cognition, 44*(6), 922–936. https://doi.org/10.3758/s13421-016-0608-9

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 51, pp. 45–100). Elsevier Academic Press. 10.1016/S0079-7421(09)51002-X.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7*(1), 75–79. https://doi.org/10.1038/nn1165

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C)*. Psychological Corporation.

Rodríguez, M., & Sadowki, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning, 50*(2), 385–412. https://doi.org/10.1111/0023-8333.00121

Sweller, J. (2020). Cognitive load theory and instructional technology. *Educational Technology Research and Development, 68*, 1–16. https://doi.org/10.1007/s11423-019-09701-3

Vaughn, S., Cirino, P. T., Wanzek, J., Wexler, J., Fletcher, J. M., Denton, C. D., … Francis, D. J. (2010). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. *School Psychology Review, 39*(1), 3–21. https://doi.org/10.1080/02796015.2010.12087786