DOCTORAL DISSERTATION

Thesis booklet

Selma Boz

WORKING MEMORY TRAINING: COGNITIVE AND MATHEMATICAL IMPLICATIONS IN SCHOOL-AGE CHILDREN

Doctoral School of Education

Special Needs Education PhD Program

Supervisor: Dr. Klara Marton, PhD, habil

Budapest, 2024

Table of Content

List of Tables and Figures

1 Introduction

Mathematical tasks involve various processes, such as understanding numerical concepts, logical reasoning, counting, and problem-solving. Working memory (WM) functions including storage, monitoring, and manipulation of information are strongly associated with these processes in mathematics (De Stefano & LeFevre, 2004; Raghubar et al., 2010). From an educational perspective, enhancing WM can provide a promising approach to improve mathematics education outcomes. Specifically, an interference framework of WM (Oberauer, 2001; Cowan, 2000-2001) provides a distinctive structure to the understanding of individual differences in WM performance. Efficient interference control can potentially lead to enhanced WM capacity. WM training programs can help reduce the likelihood of interference and perform cognitive-based tasks more efficiently (Klingberg, 2010; Jaeggi et al., 2008; Salminen et al., 2012). A considerable amount of research has been conducted to interpret practical gains of WM training (Dahlin et al., 2008; von Bastian et al., 2013; Shipstead et al., 2012; Conway et al., 2011). Many of them have demonstrated significant improvement in participants' skills (Melby-Lervåg & Hulme, 2013; Lövdén et al., 2012; Brehmer et al., 2012). From this perspective, the rationale of the present research is noteworthy to investigate the impacts of targeted WM training throughout an interference framework.

The present experimental study was conducted to discover how WM training contributes to cognitive-based and mathematical improvements in school-age learners from various schools in Istanbul. The effects of the training were evaluated with the pre- and post-training performance measures in a set of cognitive and mathematics tasks. The differences between training and control groups were compared with pre-test performance for each task and through the analysis of changes following the training; pre- and post- test performances were compared for each task within the training groups. Cognitive (e.g., WM capacity, attentional control) and non-cognitive (e.g., age, motivation, SES) factors were used to understand whether individual differences in these factors are related to training performance.

The outcomes of this study provide substantial evidence to endorse the use of WM training for the improvement of children's cognitive and mathematics skills. With additional investigation, practical applications can be established for learners. Educators, especially curriculum developers, can utilize these findings to devise mathematics programs and integrate principles about cognitive skills with mathematics programs. Such interventions might be promising for mathematics learning, as they can render the curriculum more effective and responsive to individual needs (Sternberg, 2003).

2 Background

Interference Framework of Working Memory

WM is acknowledged as a more processing-oriented construct and provides active processing and temporary storage of task-relevant information dynamically (Baddeley, 1992). One of the most important characteristics of WM is its limited capacity, which restricts cognitive performance.

Interference theory has been considered a distinctive and broader perspective of WM and its functionality across various domains. This theory attributes capacity limitation primarily to interference among memory representations and processes (Oberauer, 2009). When an individual faces a flow of information, his/her limit of capacity is exceeded. This exceeding in capacity might result in limitations to hold information in memory and to update those items during the processing of new information. It eventually becomes difficult to differentiate previously learned information from the subsequently learned one. This phenomenon is called proactive interference (Jonides & Nee, 2006), where previous or current information in memory is distracting subsequent information while performing a task. The present study focused exclusively on the phenomenon of proactive interference.

Interference theory points out that the ability to resist interfering information is a key element in updating WM contents and is a source of individual differences in WM performance (Cowan, 1995). Accordingly, the terms, interference control and resistance to interference were cornerstones of this research and these terms are described as the capability to resist irrelevant information and distractors in a given task (Nigg, 2000).

Binding and Updating: Mechanisms to Resist Interference in WM

Oberauer and his colleagues (2007) note that capacity is not only determined by the number of items that can be maintained in WM separately, but also by the number of composite items that can be bound together simultaneously. The process of combining different pieces of information into meaningful representations is called binding. It helps to hold multiple representations as a single unit within WM. Content and context of information are integrated to create structural or relational representations in WM (Wheeler & Treisman, 2002). The system where activated representations form a new relational representation corresponds to context for binding content representations to spatial or temporal positions (e.g., words linked with list positions), or a *"schema"* for binding content items to slot (e.g., *"words bound to a syntactic schema or numbers bound to roles in an equation"*) (Oberauer & Lange, 2009, p.104).

While holding and manipulating multiple bindings, interference may occur between competing representations and processes. Dynamic binding is required for the mechanism where the new construction is set up and maintained in WM by integrating it with its representations. In this view, bindings must be quickly built and dissolved again when the representations are updated or discarded (Oberauer & Lange, 2009).

Context in WM is maintained but content bound to it is updated for each trial throughout a task (Kessler et al., 2023). Updating needs to be specific to keeping items in WM independently and make decisions about which one needs to be kept and which one needs to be removed or replaced to adapt to the new information (Vockenberg, 2006; Kessler & Meiran, 2008). The balance between shielding existing information and updating serves effective WM functioning.

Besides the WM updating functions mentioned above, another important factor of building bindings and updating information is to support activation and recognition of target information (Oberauer, 2009). Recognition refers to the decision process used to determine when an item or event occurred in the past. Familiarity and Recollection

Familiarity is based on the identification and activation of items in long-term memory during the recognition process. Recognition in memory derives from the assessment of familiarity and from the retrieval of a set of structural information that also involves its associated items (McElree et al., 1999). During this process, the accurate binding of content to a context is essential to hold current information in WM, since the retrieval of the content depends on its context. Familiarity is not sensitive to the context; therefore, it is impossible to keep the current information active only through familiarity when updating is rapid, leading to the possible retrieval of no-longer-relevant items (Kessler & Meiran, 2008). However, familiarity-based decision may create errors, such as interference errors. Recognition is not based on an automatic assessment of familiarity only, but also on recollection that is controlled consciously, not automatically (Yonelinas & Jacoby, 1994).

Recollection is considered as a systematic search process involving the context of an item that was previously encountered (Szmalec et al., 2011). According to Oberauer and Lange (2009), *"familiarity arises from activated representations in long-term memory, ignoring their relations; recollection retrieves bindings in the capacity-limited component of working memory."* (p.102).

The conflict between familiarity and recollection contributes to proactive interference in WM (Oberauer, 2005). For example, high-level familiarity can create a conflict among representations from irrelevant information. When these irrelevant representations are rejected during recollection, interference among them can be avoided (Oberauer, 2005). The *n*-back task is a typical example to demonstrate the role of binding and updating of WM representations in a conflict paradigm (Gray et al., 2003).

n-back Paradigm

In an *n*-back task, a participant is rapidly exposed to stimuli such as letters or shapes presented one at a time. The goal is to judge whether the current item matches the one that was presented "*n*" items prior. The "*n*" can be manipulated to increase or decrease the load in the WM system. In this task, stimuli can be either target, new distractor or interference items (lures). A target item is a stimulus that matches an item presented "*n*" steps prior in the sequence, requiring a correct acceptance from the participant. On the other hand, a new distractor is a stimulus that is different from any of the preceding stimuli and does not match the *n*-back stimulus presented before.

During the task, the participant is required to make a recognition decision on each item by accepting targets and rejecting distractors in accordance with the *n*-back rule. Successful performance on this task requires the binding of each letter (content) to the appropriate temporal position (context) and the updating of these content-context bindings as they change with incoming new information (Oberauer et al., 2007).

Interference Framework in Mathematics

Interference in WM can be also a source of performance limitations in measures of mathematics learning processes. There is a common consent that arithmetic facts are constructed in interrelating structures in long-term memory (e.g. Campbell, 1995) and that when one encounters an arithmetic problem, pertinent incorrect answers might be activated. As a result, incorrect answers create competition with correct answers which interfere with the process of retrieving correct answers (Campbell & Tarling, 1996). This interference results in failures and slow processing of information (Noël & De Visscher, 2018). Tasks with increased complexity typically have more steps and thereby are more susceptible to inaccurate results. As an example, when an individual is performing arithmetic operations, holding intermediate result is required while carrying and borrowing numbers. During this process, recalling and using procedures of arithmetic could potentially be disrupted by proactive interference.

De Visscher and Noël (2014) proposed that similarity-based interference in arithmetic problems determines the performance on arithmetic facts. When two items have a definite amount of overlap with respect to their features, they share considerable amount of feature of their representations regarding similarity of items and these features interact with each other, resulting in interference (Oberauer & Kliegl, 2006).

Solving word problems is a multi-step process where primarily understanding the narratives in the problem and then using relevant information while rejecting irrelevant ones before building up a mathematical sentence is indispensable to solve the problem (Peng et al., 2016). The influence of numerically and literally irrelevant information in a word problem-solving task may affect differently the degree of interference in WM (Ng et al., 2017). Whereas numerically irrelevant words are perceived as values that must be used in the operation or in any mathematical calculations (i.e., equation), literally irrelevant information may be detected as unnecessary information for solving the problem.

Larger problems are more prone to be connected with incorrect answers due to increased likelihood of errors and problem size effect. For example, two-digit or complex problems are assumed to be large problems (Thevenot et al., 2010), due to their higher probability of incorrect answers. The common clarification for problem size effect is that smaller problems (e.g., simple additions, single-digit multiplication) are more frequently solved using direct retrieval strategy than larger problems (e.g., complex subtraction, multi-digit problems) (Thevenot et al., 2010; Zbrodoff & Logan, 2005).

The exploration of the relationship between mathematics and interference displays a complex interplay between WM mechanisms and information processing in mathematics. Studies have highlighted how complexity and similarity in mathematics problems amplify the challenge by increasing interference in WM. The findings provide justifications and insights into WM functionality while solving mathematics problems, especially arithmetic problems, and the possible factors that create interferences in this process were considered as a measure for their mathematics proficiency in the current study.

Working Memory Training and Transfer Effects

It is proposed that despite the limit in WM capacity, the efficiency of WM processes can be improved with WM training (Klingberg et al., 2005; Verhaeghen et al., 2004; Westerberg et al., 2007; Jaeggi et al., 2014; Melby-Lervag & Hulme, 2013). The key role that WM plays in many processes and the individual differences in WM performance have inspired research questions about the potential to train the WM system and to transfer this training effect to performance on complex tasks that are known to recruit the WM system, such as language and mathematics (Sternberg, 2020).

The transfer effect refers to how training improvements promote other skills or performance in various cognitive tasks. It can be categorized as near or far transfer effect. As proposed by some studies (e.g., Klingberg, 2010), the improvements which can be observed behaviorally result from increasing performance on tasks similar to the trained tasks, defined as near transfer effects. Near transfer effects reflect direct acquisition from the training. For instance, near transfer effect can be observed after receiving n-back training if an individual exhibits better performance on a digit span task which refers to recalling numbers in order. On the other hand, a broader cognitive improvement is required for far transfer effect that occurs when training improves performance on specific tasks which do not share the same cognitive processes with the trained task. For example, performance improvement in language or in mathematical tasks as a result of completing an n-back training (Dahlin et al., 2008; Jaeggi et al., 2008). The efficiency of training would be predicted from its transfer effects to untrained tasks (Shipstead et al., 2012). WM training is considered to provide both near and far transfer effects (Barnett & Ceci, 2002).

Participants' personality traits can contribute to variations in training and training effects (von Bastian & Oberauer, 2014). The findings illustrate that motivation and individual differences can influence the performance of participants during training and may also impact the training effects.

The studies on WM training are engaged in continual WM tasks which are carried out experimentally in a controlled manner and this experimental context impacts cognitive functions (Jaeggi et al., 2014; Melby-Lervag & Hulme, 2013). The n-back training serves as a proper tool where performance on the n-back task requires both strong and flexible bindings that can promote recollection and support resistance to interference related to familiarity. Strong binding and updating of items may enhance recollection, leading to accurate retrieval in this recognition task. Therefore, binding and updating mechanisms are the cornerstone of interference control in the WM system.

3 Rationale and Purpose of the Research

A considerable number of children struggle with solving basic mathematics problems in elementary schools, resulting in difficulty in their future mathematical abilities. Understanding the cognitive processes is crucial to develop more effective tools to support children's mathematical skills. Recent research (Ji & Guo, 2023) has also demonstrated that children who have strong WM skills can perform better in mathematics, underlying the relationship between mathematical skills and cognitive processes. WM skills are required to manage multiple steps in complex calculations and solve problems effectively and more accurately. These findings assure that supporting WM development can enhance school-age children's mathematics proficiency (Sala & Gobet, 2017; Shipstead et al., 2012).

Since cognitive skills are essential for success in mathematics, interventions designed to enhance these abilities can yield long-term efficacy not only in mathematics but also in other academic disciplines (Holmes & Adams, 2006). This study can provide evidence to enhance children's cognitive abilities to maintain and process information in cognitive related tasks and may contribute to the identification of special educational strategies for children with learning difficulties or cognitive impairment.

4 Aims and Hypotheses

The hypotheses of the present study were constructed within four main aims. First, to explore how WM training affects school-age children's WM systems within the interference framework, with a focus on binding and updating functions and limitations. Second, to show near and far transfer effects in performance. Third, to find evidence for the relationship between the mechanisms used for resisting interference in WM and mathematics performance in conditions where WM load and interference are manipulated. Fourth, to understand if individual differences in both cognitive (e.g., WM capacity, attentional control) and non-cognitive (e.g., motivation, SES) factors are related to *n*-back training performance. Overall, these aims highlight the necessity of the enhancement of WM capacity to perform better in cognitive-based tasks within an interference control framework of WM.

The effects of the training were evaluated with the pre- and post-training performance measures in a set of cognitive and mathematics tasks. The differences between training and control groups were compared with pre-test performance for each task and through the analysis of changes following the training; pre- and post- test performances were compared for each task within the training groups. There were two experimental and two control groups: One experimental group received training and completed pre- and post-tests, whereas the other experimental group received training and completed only post-tests; regarding control groups, one control group completed pre- and post-tests, while the other only completed post-tests (see Figure 1 below).

Figure 1. Design of the study

Training group 1 (T1):	Pre-test →	Training →	Post-test
Training group 2 (T1):		Training →	Post-test
Control group 1 (C1):	Pre-test →	→	Post-test
Control group 2 (C2):			Post-test

The research was guided by five parts of hypothesis:

1) Training performance and progress (T1 & T2)

We hypothesized that

a) Participants in the four groups would not differ in age, nonverbal IQ, and basic memory and language tasks (i.e. digit span and semantic verbal fluency).

b) *n*-back performance of participants in the experimental training groups is negatively affected by increased WM load (higher *n*-back levels) and by the presence of interference lures.

c) Individuals who more strongly believe that they can improve their abilities through training would show higher engagement in the training than individuals who believe that their abilities are given, and individuals who have higher nonverbal IQ scores would improve more rapidly in the n-back tasks over the training program.

d) Individual differences in Semantic Verbal Fluency Test would predict training progress.

e) The factors such as "age", "socioeconomical status (SES)", "participants' motivation", "performance change over the sessions" (The completed n-level from each session was considered.) and "baseline cognitive resources" – specifically assessed by the highest reached n-level from first two sessions – can predict the training progression.

2) Pre-test performance (T1 & C1)

We anticipated that

a) *n*-back task (1-back and 2-back): Participants would show better performance on new distractor items than on target items at each set size. Their performance on target items would be better than on interference items in the proactive interference tasks.

b) Modified digit span task (MDS): Performance on solving arithmetic operations was expected to be significantly better than memorizing the last items. Additionally, participants would perform better in transformation-based tasks than retrieval and substitution-based tasks.

c) Arithmetic operations: The pattern of performance would be different for the two conditions of multiplication and division. Participants would perform better in small-size problems than long-size problems.

d) Word problems: Participants would perform significantly better on the neutral condition than on the biased condition. Additionally, they would perform significantly better on the problems which have literal irrelevant information than on the problems with numerical irrelevant information.

3) Training effects, pre- to post-tests (T1)

We expected that

a) *n*-back task (1-back and 2-back): Participants would perform significantly better on both *n*-back levels after the training. Performance on target items would predominantly determine changes and participants would show no significant difference in pre- to post improvement on rejection of new distractors in both *n*-back tests.

b) MDS task: Performance on the retrieval and substitution-based tasks would improve significantly, whereas performance on the transformation-based task would not.

c) Arithmetic operations: Performance on multiplication and division would improve significantly, while performance on addition and subtraction would not because the effect of training is more likely to be observed in complex tasks.

d) Word problem: Performance on problems involving literal and numerical information would significantly improve, whereas performance on problems which have only relevant information would not.

4) Training effect post-test performance (T2 & C2)

It was anticipated that

a) *n*-back task (1-back and 2-back): Group T2 would show better performance on target items than Group T2. The performance on new distractor items would not differ significantly between the groups.

b) MDS task: The groups would not differ in transformation-based tasks while the training group would show better performance in retrieval and substitution-based tasks. Additionally, the training group would perform significantly better at recalling items than the control group.

c) Arithmetic operations: Groups would differ significantly either in small-size problems in all types of operations nor in addition and subtraction problems.

d) Word problems: The training group would perform significantly better on problems which contain irrelevant information. However, the difference between training and control groups' performances were not expected to be significant in problems that involved only relevant information.

5) Test-retest effect post-test performance (only C1) / (C1 & C2)

a) It was hypothesized that changes in performance for the C1 group would not be significant on any task.

b) There was expected no significant difference between C1 and C2 control groups in post-test performance. C1 and C2 groups would be compared in each task to confirm that there is no testing effect in performance of participants who complete the pre-tests.

5 Research Design and Methods

The study implemented a training paradigm and interpreted its effects on groups of elementary level children in mathematics and cognitive skills. Participants between the ages of 9 and 12 included 44 children in elementary level classes from various schools in Istanbul, Türkiye. The Solomon four-group design was applied to account for testing effects in the study.

The training groups completed sixteen 20-minute sessions of adaptive *n*-back task over four weeks. Participants performed pre- and post-tasks or only post-tasks depending on their group assignment, before and after the training period. The experimental groups completed the adaptive *n*-back training with lures. Within these groups, the effects of individual differences were tested with the Theories of Cognitive Abilities Scale (Dweck et al, 2000) to understand whether participants' intrinsic motivation influenced the results of the study. Between the experimental and control groups, three categories of pre-/post-tests were administered to determine transfer effects:

n-back tasks (1-back and 2-back), modified digit span (MDS) task, and mathematics proficiency tasks (see Table 1, below).

Table 1. Detailed information on pre and post-tests

Tasks	**Conditions**	**Item Types**
1-back	Neutral	New Distractor Target
	Possible Interference	New Distractor Target Target (Possible error)
2-back	Neutral	New Distractor Target
	Interference	New Distractor Target Proactive interference
Modified Digit Span	Transformation	New Distractor Target
	Substitution	
	Retrieval	
	Retrieval and Substitution	
Arithmetic Operations	Baseline	Target
	Interference-based	
Word Problems	Baseline	Target
	Literal irrelevant information	
	Numerical irrelevant information	

n-back 1-back and 2-back tasks were administered to examine participants' working memory updating skills. Participants completed neutral condition of *n*-back tasks before possible proactive interference condition for the 1-back task and proactive interference condition for the 2-back task in the order which they received the 1-back and 2-back tasks respectively.

Modified Digit Span (MDS) Four tasks were included as different components of WM system. Arithmetical operations were presented sequentially in two boxes for each task on participants' computer screens. In some tasks, participants were responsible to retrieve the information in the box where they used it as an operand to apply the operation and then substitute the result in the relevant box. In the t task, only transformation was required, and participants simply did the calculations and typed the results in the corresponding boxes without retrieving or substituting any information. Due to no presence of any information to memorize, this condition did not include initial and recall items. In the tS task, participants had to remember first presented initial items and apply the operations in each box. Since this task involved substitution, they were required to hold the result of each operation in mind to memorize recall items at the end of each list. In the Rt task, which included retrieval, two initial numbers presented in the beginning of each list were required to be memorized, associating with the box. Then, each number was retrieved to use it in incomplete operations (e.g. ? + 2) depending on its box. After participants performed the operations and entered the results, they typed these initial numbers in their associated boxes, at the end of the list. In the RtS task, which included retrieval and substitution, participants were responsible to remember the first presented initial item for each box and use it to perform the first incomplete operation (e.g. + 3) of the associated box. The result of each box had to be remembered to use it as an operand for the following operations.

Arithmetic operations Baseline and interference condition of arithmetical operations were administered to measure far transfer effect of the training to performance to resist interference while solving arithmetical operations within large size. Participants were presented questions on their computer screen consecutively without limiting time until they answered them. This task was separated into four tasks: Addition, Subtraction, Multiplication and Division and each task included 10 questions within two different conditions.

Word Problems A total of six-word problems (Ng et al., 2017) were administered to measure far transfer effect of the training to performance on mathematics problems which included literal and numerical irrelevant information. The task consisted of two conditions: The baseline condition had two problems which did not include any irrelevant information, and the biased condition contained four problems within irrelevant information either literally or numerically.

All participants completed a battery of tests (Table 1) depending on their group assignment to understand the effect of the training within-groups and to compare their performance between groups. All tasks and training sessions were administered online using E-Prime Go which was obtained from E-Prime 3.0 software (Psychology Software Tools, 2020) to present stimuli and record responses remotely. Accuracy and reaction time data for each task were collected from E-

Prime Go result sheets. R studio (2022) was used for both data processing and analysis. The analysis goals were addressed using a mixed-effects regression analysis to investigate both within- and between-subject effects in hierarchical data.

6 Results

In what follows, I summarized the results of five parts of the hypothesis.

Participants in training and control groups did not significantly differ in age $F(3, 40) = 0.53$, p= 0.66, scaled TONI scores $F(3, 40) = 1.00$, p= 0.40, SVFT scores $F(3, 40) = 0.41$, p= 0.75. Since DST scores were not normally distributed, Kruskal Wallis test was conducted to check the difference between groups. Participants in four groups did not significantly differ in DST scores $H(3) = 1.39$, p = 0.71.

Table 2. Participant characteristics by group

Group	N	Age	TONI	SVFT	DST
T1	11	10.5 (1.0)	110.8 (12.5)	12.3 (3.1)	6.9 (1.4)
T2	11	10.4 (0.8)	114.8 (11.2)	12.5 (2.9)	7.7 (2.8)
C1	11	10.8 (1.0)	112.0 (10.8)	13.5 (3.0)	7.3 (2.0)
C2	11	10.5 (0.8)	106.9 (8.6)	13.0 (2.1)	6.5 (1.3)

Preliminary assessment was conducted to test skewness, kurtosis and normality of scaled TONI scores and SVF scores for all participants. According to the Shapiro-Wilk test (W), TONI score, SVFT score and achieved training level data ($W \sim 1$, $p > 0.05$) were normally distributed (see Table 2, below). However, only training groups were predicated on the achieved training level.

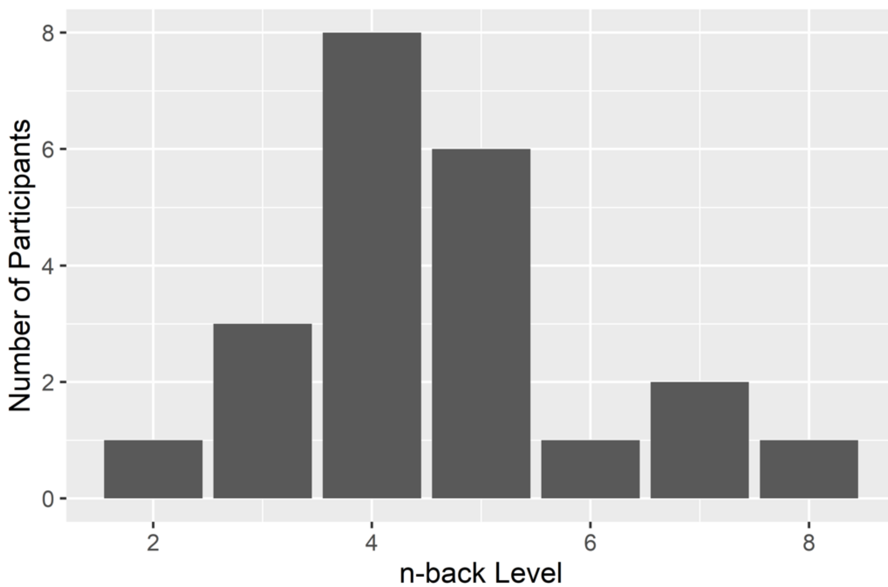
Table 3. Normality of participants' characteristics

Characteristic	*Skewness*	*Kurtosis*	*Normality (W)*	*Normality (p)*
TONI score	0.332	2.373	0.971	0.330
SVFT score	0.707	3.185	0.953	0.073
Achieved training level	0.653	3.187	0.918	0.070

Training performance and progress

According to descriptive analyses, participants in training and control groups did not significantly differ in age, nonverbal IQ, semantic verbal and digit span test. The participants did not differ in self-reports of motivation which was measured by the Theories of Cognitive Abilities questionnaire (Dweck, 2000). The minimum level which was achieved by only one participant was 2-back level with proactive interference lures and the maximum level which was also achieved by only one participant was 8-back level with proactive interference lures (see Figure 2).

Figure 2. *n*-back level achieved by participants in training groups



As predicted by hypothesis 1b, participants would demonstrate improvement throughout the training period. The analysis of the distribution of maximum *n*-level achieved in each session provided insight into the progression of participants' training over time. The training progress of participants in the training groups was not uniform or consistent. Among the participants, four individuals showed slow progress and only reached the 3-back level, while other six individuals initially progressed slowly but eventually completed the entire training up to the 4-back level. Two participants who were able to reach 7- and 8-back levels showed faster improvement in their performance.

The results obtained from the Theories of Cognitive Abilities questionnaire indicated a relatively high level of motivation among the participants. However, it is not evident that the scores from the questionnaire exhibit a significant association with individual differences in training progress. The analysis examined two variables, nonverbal IQ, and semantic verbal fluency, to explore their relationship with training progress. The findings did not support the hypotheses, revealing that nonverbal IQ and semantic verbal fluency were not significantly associated with individual progress in the training program.

The observations provide evidence that individuals whose performance in baseline cognitive tasks was higher achieved higher n-levels with the training. On the other hand, age, socioeconomic status (SES), and motivation did not predict their training progress. Participants' capabilities in the training might be relatively resilient to individual differences in these factors.

Pre-and post-test performance

Different models were compared for each task, including the null model (no predictors), a model with random intercept, a model with random intercept and random slope, and a model with random intercept, random slope, and interactions. The null model was particularly important in determining whether the change score for each task significantly differed from zero. After the initial assessment, predictors of change were added to each model to determine the best fit.

n-back task

Hypothesis 2a proposed that participants would demonstrate better performance on new distractor items compared to target items in both the 1-back and 2-back conditions of the task. Additionally, it was expected that their performance on target items would be better than on interference items in the proactive interference tasks.

In the 1-back task, participants exhibited significantly lower accuracy in the tasks, where possible proactive interference (PI) errors were available, compared to the neutral condition. Moreover, their reaction times were significantly higher in the PI condition compared to the neutral condition.

Table 4. Pre-test 1-back accuracy and reaction time predicted by condition

Accuracy

Variable	Estimate (SE)	z	*p*
Fixed effects			
Intercept	0.840 (0.310)	2.713	< **.05**
Proactive	1.653 (0.175)	9.465	< **.001**
Random effects	Variance	*sd*	

Intercept	1.492	1.221	
Reaction time			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	485.65 (34.040)	14.267	< **.001**
Proactive	0.218 (0.022)	9.982	< **.001**
Random effects	Variance	*sd*	
Intercept	19399	139.3	
Residual	96797	311.1	

Note: Sample n = 22. The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

In the 2-back task, participants demonstrated significantly higher accuracy and higher reaction times in the PI condition compared to the neutral condition. These findings support the assumptions about the impact of interference items on n-back performance, leading to longer reaction time in the PI condition compared to the neutral condition, but the assumptions about performance in accuracy contradict with the finding where the accuracy was higher in PI condition. The reason might be related to the nature of the task, which is low level of n-back task and after practicing the 1-back tasks and neutral condition of the 2-back task, the participants might have become better in performance. This can also be explained by the trade-off theory (Heitz, 2014) which suggests that decisions made with more accuracy require more time.

Table 5. 2-back pre-test accuracy and reaction time predicted by condition

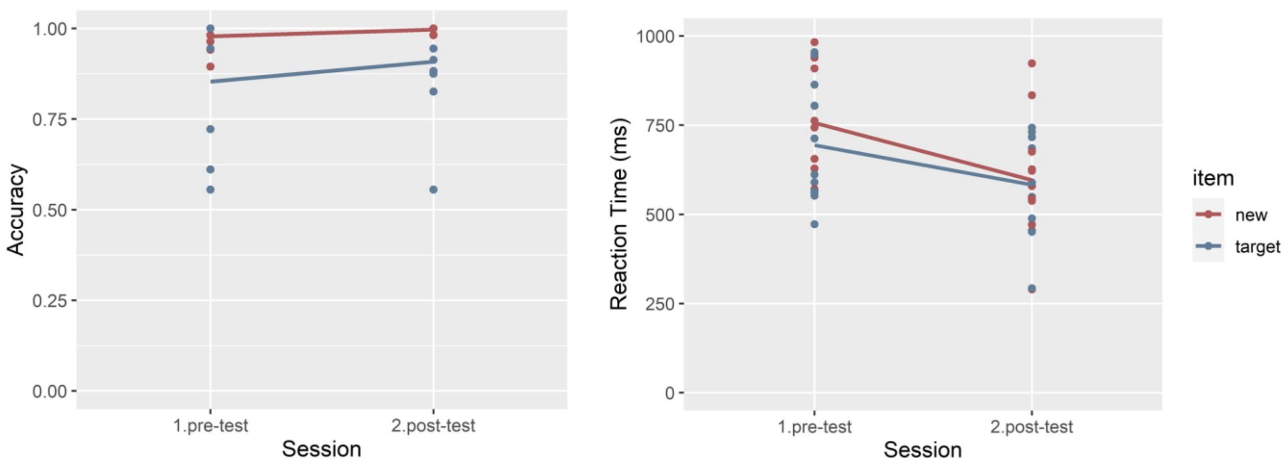
Accuracy			
Variable	Estimate (SE)	z	*p*
Fixed effects			
Intercept	1.228 (0.262)	4.683	< **.001**
Proactive	0.311 (0.144)	2.155	< **.05**
Random effects	Variance	*sd*	
Intercept	1.143	1.069	
Reaction time			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	671.7 (53.98)	12.442	< **.001**
Proactive	0.072 (0.021)	3.357	< **.001**
Random effects	Variance	*sd*	
Intercept	55691	236.0	
Residual	86351	293.9	

Note: Sample n = 22. The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

The results of the assumption which suggested a significant change in performance of the pre-test training group, proposed by hypothesis 3a, revealed that participants showed varying degree of performance change in terms of accuracy and reaction time across different *n*-back levels after they completed all sessions of the training. The intercepts for null accuracy and reaction time change were not significant. Therefore, the changes on condition and item type of 1-back or 2-back tasks were not analyzed.

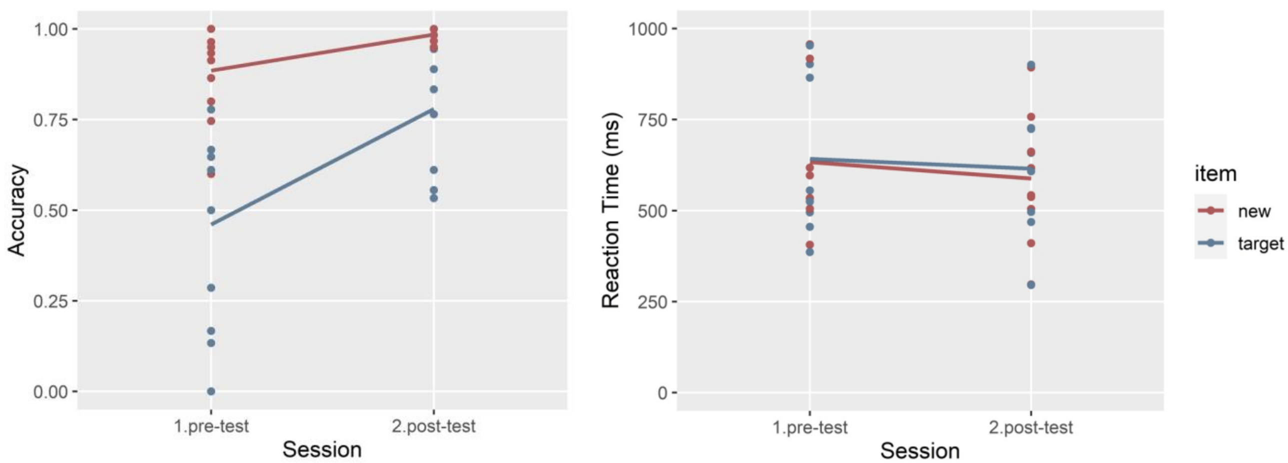
Based on the mean change in accuracy and reaction time, an insignificant mean change in accuracy and a significant decrease in reaction time was observed in 1-back task from pre to post-test. The ceiling effect in this task indicated that participants reached a maximum of correct responses; therefore, further improvement could not be observed in accuracy, but they became faster after the training.

Figure 3. 1-back training group accuracy and reaction time change



Furthermore, a significant increase in accuracy for target items and insignificant mean change in reaction time was noticed in the 2-back task. This suggests that the performance of the participants was similar in accuracy for the 1-back task at both pre and post-tests, while they improved their response time at the post-test. The shorter temporal span between the maintained and updated items, and the reduced cognitive load of the 1-back task likely resulted in quicker response times (Jaeggi et al., 2010).

Figure 4. 2-back training group accuracy and reaction time change



The additional analysis of the training effect supports Hypothesis 4a, which postulates n-back task performance difference between post-test training group and post-test control group, indicating an improvement in the training group. These findings are consistent with the hypothesis and show that the training enabled the training group to have a positive effect to perform the n-back task more accurately. The participants in this training group showed higher performance in accuracy not only on the new items but also with the target items compared to the participants in the control group, however, the groups did not significantly differ in reaction time performance. The results suggest that the training might have underscored skills that were either acquired or improved, leading to better accuracy performance in the n-back tasks through enhancement of WM capacity.

MDS task

The investigation was based on hypothesis 2b that participants would exhibit various performances in the conditions that required retrieval or substitution or both while performing basic arithmetic operations in all conditions of the MDS task. The participants were expected to perform better in the conditions where only transformation was employed as compared to the conditions in which they were engaged in retrieving and substituting items during the task. The performance in accuracy for the retrieval and substitution-based condition (RtS) with both short and long listed block items was lower among other conditions, while performance in reaction time was significantly higher for the transformation and substitution-based condition (tS) and the RtS condition with long-listed block items for both. The challenge in the retrieval and substitution-based task highlights the role of WM systems in retrieval process because information requires to be executed and manipulated while other information is being simultaneously retrieved from WM resources (Unsworth & Engle, 2007).

Table 6. Pre-test MDS task accuracy and reaction time predicted by condition of interference condition

Accuracy			
Variable	Estimate (SE)	z	*p*
Fixed effects			
Intercept	3.147 (1.169)	2.692	**< .01**
tS (long list)	-1.426 (1.094)	-1.303	0.193
Rt (short)	-0.015 (0.423)	-0.036	0.972
Rt (long)	0.454 (0.408)	1.112	0.266
RtS (short)	1.783 (0.437)	4.077	**< .001**
RtS (long)	-0.981 (0.438)	-2.239	**< .05**
Random effects	Variance	*Sd*	
Intercept	0.535	0.732	
Reaction time			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	1987.5 (165.4)	12.01	**< .001**
tS (long list)	229.1 (67.29)	3.406	**< .001**
Rt (short)	-53.82 (59.20)	-0.909	0.364
Rt (long)	152.0 (100.8)	1.508	0.132
RtS (short)	89.41 (67.95)	1.316	0.189
RtS (long)	173.1 (68.72)	2.519	**< .05**
Random effects	Variance	*sd*	
Intercept	526589	725.7	
Residual	1514693	1230.7	

Note: Sample n = 22. The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

The results of performance change of the training group provide insight into the near effects of the WM training. In this aspect, hypothesis 3b predicted that n-back training would contribute to significant improvement in performance on retrieval and substitution-based tasks compared to transformation tasks, reflecting near transfer effect to the updating processes of the WM system. The findings showed that the change in performance varied across the conditions of MDS task. The training group did not improve significantly on retrieval (Rt) and retrieval and substitution based (RtS) tasks. This outcome contradicts the predictions and supports the idea that trained cognitive skills may not directly adjust to the improvement in retrieval and substitution processes. Transfer effects of WM training can be more task-specific, bound to specific cognitive processes employed in the trained tasks (Dahlin et al., 2008).

Table 7. MDS task accuracy and reaction time change predicted by condition

Accuracy Change			
Variable	Estimate (SE)	t	*p*
Fixed effects			
Intercept	0.088 (0.100)	0.880	0.392
tS (long)	0.682 (0.263)	2.590	**< .05**
Rt (short)	0.087 (0.437)	0.199	0.845
Rt (long)	-0.239 (0.450)	-0.532	0.602
RtS (short)	0.013 (0.381)	0.035	0.973
RtS (long)	-0.011 (0.377)	-0.030	0.977
Random effects	Variance	*sd*	
Intercept	0.000	0.000	
Residual	0.104	0.322	
Reaction Time Change			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	-76.88 (102.9)	-0.747	0.469
tS (long)	0.706 (0.113)	6.222	**< .001**
Rt (short)	-0.417 (0.228)	-1.828	0.088
Rt (long)	0.289 (0.151)	1.915	0.079
RtS (short)	0.097 (0.195)	0.497	0.626
RtS (long)	0.182 (0.108)	1.686	0.115
Random effects	Variance	*sd*	
Intercept	48229	219.6	
Residual	36912	192.1	

Note: Group presented is the T1 group (n = 11). The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

As stated in hypothesis 4b, the post-test training group would perform better than the post-test control group in retrieval and substitution-based tasks as well as in operation and recall items in each condition, except baseline condition. The training group did not improve performance significantly in accuracy, since groups did not differ at post-test in terms of accuracy. On the other hand, the training group showed faster response times in all interference conditions than the control group, but their performance in baseline condition was similar. This improvement only in response time might suggest that the training was related to certain cognitive processes more than others, such as complex retrieval and substitution.

Arithmetic operations

While performance on accuracy was lower, it was higher on reaction time in the interference condition compared to baseline condition. These findings are in alignment with the assumptions postulated in hypothesis 2c. Participants exhibited similar accuracy performance in the baseline condition across the operation types, the reaction time performance was higher in subtraction and multiplication. In the interference condition, the accuracy performance was lower across all operation types, but reaction time performance was higher in multiplication and division. Multiplication and division are more complex operations and involve larger numbers. Unlike addition and subtraction, retrieval of the multiplication table is required for those operations. Binding and updating processes have also a significant role in explaining how large-size problems create proactive interference. In large-size problems, the complexity and quantity of information contribute to the increase in binding numerical values to their corresponding operation types (Oberauer et al., 2012).

Table 8. Pre-test arithmetic operations accuracy and reaction time predicted by operation type of interference condition

Accuracy			
Variable	Estimate (SE)	z	*p*
Fixed effects			
Intercept	-1.068 (1.408)	-0.758	0.448
Subtraction	0.824 (0.861)	0.956	0.339
Multiplication	1.512 (1.181)	1.280	0.200
Division	0.911 (1.144)	0.796	0.426
Random effects	Variance	*sd*	
Intercept	0	0	
Reaction time			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	35547.1 (2512.2)	14.15	**< .001**
Subtraction	4140.4 (2485.3)	1.666	0.103
Multiplication	5881.8 (2863.1)	2.054	**< .05**
Division	7720.4 (1623.7)	4.755	**< .001**
Random effects	Variance	*sd*	
Intercept	77714019	8816	
Residual	99851109	9993	

Note: Sample n = 22. The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

Hypothesis 3c predicted a pattern change in both conditions of this task and enhancement in multiplication and division performance in the interference condition, leading to far transfer effect of the *n*-back training. Performance changes of the pre-test training group were observed only in reaction time in both baseline and interference conditions. Performance change in reaction time was significantly larger in the interference condition than in the baseline condition, suggesting that participants showed more improvement in the interference condition after the training. Further analyses showed that participants were faster at post-test than pre-test in all operations in both conditions, but the improvement from pre- to post-test was significantly greater for subtraction tasks in the baseline condition and for division tasks in the interference condition. Far transfer effect is not compatible with some cognitive tasks. In this study, no significant improvement was found in addition and subtraction under the interference condition, probably due to less complexity of the tasks than in multiplication and division.

Table 9. Arithmetic operations interference condition reaction time change by operation type

Reaction time change			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	-11951.8 (4488.7)	-2.663	**< .05**
Subtraction	-1138.9 (3520.4)	-0.324	0.749
Multiplication	-1837.5 (3233.9)	-0.568	0.576
Division	8254.7 (2502.1)	3.299	**< .01**
Random effects	Variance	*sd*	
Intercept	140549787	11855	
Residual	135986663	11661	

Note: Group presented is the T1 group (n = 11). The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

Another hypothesis on training effect proposed that the post-test training group would outperform the post-test control group in large-size arithmetic operations. Contrary to expectations, according to the analyses, no significant difference was observed in either accuracy or reaction time performance between these two groups.

Word Problems

The primary purpose of implementing this test was to investigate the far transfer effect of the n-back training. This segment of the study aimed to furnish further supportive evidence for theories about the interference mechanisms within the context of solving word problems. The performance between group T1 (the pre-test training group) and group C1 (the pre-test control group) on word problems was similar. However, the findings revealed that the presence of irrelevant information, either literal or numerical, did not significantly influence the performance of participants on word problems.

Table 10. Pre-test word problems biased condition accuracy and reaction time predicted by information type

Accuracy			
Variable	Estimate (SE)	z	*p*
Fixed effects			
Intercept	70.77 (31.39)	2.255	< **.05**
Literal distractor	-33.42 (31.39)	-1.065	0.287
Numerical distractor	-24.04 (31.39)	-0.766	0.444
Random effects	Variance	*Sd*	
Intercept	2908	53.93	
Reaction time			
Variable	Estimate (SE)	*t*	*p*
Fixed effects			
Intercept	89490 (18048)	4.958	< **.001**
Literal distractor	-31426 (15190)	-2.069	0.052
Numerical distractor	66702 (32986)	2.022	0.058
Random effects	Variance	*sd*	
Intercept	2.084e+09	45649	
Residual	2.115e+09	45994	

Note: Sample n = 22. The reported data for fixed effects consists of unstandardized coefficients, while for random effects, it includes variance.

Group T1 did not exhibit performance change in terms of accuracy and reaction time as a result of training, as proposed by hypothesis 3d.

Groups T2 (the post-test training group) and C2 (the post-test control group) differed in only reaction time performance but interestingly group C2 performed faster than T2 in the problems which included numerical irrelevant information. The faster response time of group C2 might be attributed to their reliance on more familiar and less complex approaches while performing the task (Karpicke & Roediger, 2008; Dehaene, 2011). The contradiction between the assumptions and the findings in this part of the study might be caused by high individual differences and inadequacy of problem-solving skills due to the presence of younger children. Only type and amount of irrelevant information cannot be indicative to determine performance in word problems with interferences. Participants are also required to be able to use cognitive strategies and to have reading comprehension and numeracy skills (Holmes & Adams, 2006). The training groups probably did not sufficiently address these multifaceted demands required for proficiency in word problems with irrelevant information.

Practice effect

To evaluate the practice effect of repeated testing, two different sets of analyses were conducted within the control groups. It was anticipated that the pre-test control group would not show any improvement in tasks from pre-test to post-test, and no significant difference would be observed between the pre-test control group and the post-test control group at post-test.

The performance of the pre-test control group in terms of either accuracy or reaction time did not change in the n-back, arithmetic operations and word problems tasks from pre to post-test, but the reaction time performance in the MDST was lower at post-test. As expected, without training, the participants did not exhibit significant improvement in the n-back task (Redick & Lindsey, 2013), nor in the arithmetic operations and word problems tasks (Karbach & Verhaeghen, 2014). Conversely, a decrease in reaction time in the MDST at post-test, despite no training, could indicate a test-retest effect. The faster response might be a consequence of increased familiarity with the task.

The findings focusing on the differences between the pre-test and post-test only control groups were partially consistent with the assumptions. At post-test, while these two groups did not show significant difference in either the n-back or arithmetic operations tasks in terms of accuracy and reaction time performance, the pre-test control group performed faster in the MDST and demonstrated higher accuracy in the word problems task. This supports the notion that repeated exposure to cognitive tests can contribute to improvement over time even if training was not received (Collie et al., 2003).

7 Overall Conclusions

The primary aim of cognitive training studies is to discover whether improvements in performance can be transferred to areas that were not directly targeted within the training (Jaeggi et al., 2008; Melby-Lervåg & Hulme, 2013; Simons et al., 2016; Sala & Gobet, 2017). Studies on cognitive training exhibit multifaceted and complex findings which were gathered from different methodological approaches and designs. The degree of overlap between the underlying mechanisms that are targeted by the training and the outcome measures have a key role in obtaining a transfer effect to untrained tasks.

In line with this, the current study provides significant insights to demonstrate training effects from a methodological perspective, and across the tasks designed to observe transfer effect of the training. The Solomon 4-group design differentiated this study from previous research and provided advanced and extensive analyses of data gathered from distinct combinations of tasks to examine transfer effects of the training. Detailed analysis revealed that transfer effects of the training were observed in the cognitive tasks (1-back, 2-back and MDS) and arithmetic operations, particularly in the condition including proactive interference lures or possible errors. Participants demonstrated increased speed in those tasks despite the absence of significant improvement in accuracy. However, transfer effects in word problem-solving tasks were not observed in this study. These specific points provide an important contribution to exploration of the mechanisms which underlie transfer effects of n-back training.

This study underscores the importance of the interference framework of the WM system for consistent explanations of why individuals differ in performance across different WM measures and which specific WM functions are crucial or task-oriented skills. Performance differences observed across tasks consisting of different conditions and item types can be attributed to interference control skills. Completing these task goals successfully depends on the ability to use binding and updating mechanisms to maintain and manipulate information amidst interference.

The findings of this study have implications for research in cognitive skills and education, as well as practical applications in educational settings. The observations that predict mathematics performance based on cognitive abilities provide a steppingstone for developing educational strategies and interventions aimed at improving mathematics performance (Bull & Lee, 2014). These findings also highlight the necessity for early intervention in children with low working memory (WM) capacity or poor interference control. Since these cognitive skills are essential for success in mathematics, interventions designed to enhance these abilities can yield long-term efficacy not only in mathematics but also in other academic disciplines (Holmes & Adams, 2006). This study can provide evidence to enhance children's cognitive abilities to maintain and process information in cognitive related tasks and may contribute to the identification of special educational strategies for children with learning difficulties or cognitive impairment. To achieve this, further research is required for better understanding of which specific tasks can be tailored to diverse cognitive abilities.

8 Publications Connected to the Topic of the Research

Boz, S., & Erden, M. (2021). The Effect of different components of working memory on multiplication skills of 3rd grade children. *Hacettepe University Journal of Education, 36*(1), 177-185. doi: 10.16986/HUJE.2020058880

Boz, S. (2021). Transfer effect of n-back training: Mathematical implications in school-age children. In *International Conference on Education and New Developments*. https://doi.org/10.36315/2021end121

Boz, S. (2024). Enhancing arithmetic skills through working memory: A review of interventions and strategies. *Journal of Mathematics and Science Teacher, 4*(3), em067. https://doi.org/10.29333/mathsciteacher/14632

Boz, S. (2024). The effect of problem size on children's arithmetic performance: Interference control in working memory. *European Journal of Psychology and Educational Research*, *7*(2), 83-92. https://doi.org/10.12973/ejper.7.2.83

References

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556–559.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637. https://doi.org/10.1037/0033-2909.128.4.612

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Frontiers in human neuroscience*, *6*, 63. https://doi.org/10.3389/fnhum.2012.00063

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8*(1), 36-41.

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition*, *1*(2), 121–164.

Campbell, J. I. D., & Tarling, D. P. M. (1996). Retrieval processes in arithmetic production and verification. *Memory and Cognition, 24*(2), 156–172. doi:3758/BF03200878

Collie A., Maruff P., Darby D. G., McStephen M. (2003). The Effects of Practice on the Cognitive Test Performance of Neurologically Normal Individuals Assessed at Brief Test–Retest Intervals. *Journal of the International Neuropsychological Society, 9*(3), 419–428.

Conway, A. R. A., Getz, S. J., Macnamara, B., & Engel de Abreu, P. M. J. (2011). Working Memory and Intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 394–418). chapter, Cambridge: Cambridge University Press.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (2000-2001). Processing limits of selective attention and working memory: potential implications for interpreting. *Interpreting, 5*(2), 117 146. https://doi.org/10.1075/intp.5.2.05cow

Dahlin, E., Nyberg, L., Bäckman, L., & Neely, A. S. (2008). Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23*(4), 720-730.

De Stefano, D., & LeFevre, J-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353-386. doi:10.1080/09541440244000328

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. Oxford University Press.

Dweck, C. S. (2000). *Self-theories: Their Role in Motivation, Personality, and Development*. Psychology Press. Retrieved from https://books.google.com.tr/books?id=P0Mccblm6eUC

Gray, J.R., Chabris, C.F., Braver, T.S. (2003). Neural mechanisms of general fluid intelligence. *Nat Neurosci. 6*(3), 316-22. doi: 10.1038/nn1014. PMID: 12592404.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*, 150.

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology, 26*(3), 339-366.

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.

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & cognition*, *42*(3), 464–480. https://doi.org/10.3758/s13421-013-0364-z

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence, 38*(6), 625-635.

Ji, Z., & Guo, K. (2023). The association between working memory and mathematical problem solving: A three-level meta-analysis. *Frontiers in psychology*, *14*, 1091126. https://doi.org/10.3389/fpsyg.2023.1091126

Jonides, J., & Nee, E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139,* 181-193.

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: a meta-analysis of executive-control and working memory training in older adults. *Psychological Science, 25*(11), 2027-2037.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319* (5865), 966-968. https://doi.org/10.1126/science.1152408

Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Joumal of Experimental Psychology: Leaming, Memory, and Cognition, 34,* 1339-1348.

Kessler, Y., Zilberman, N., & Kvitelashvili, S. (2023). Updating, fast and slow: Items, but not item-context bindings, are quickly updated into working memory as part of response selection. *Journal of Cognition, 6*(1), Article 11, 1–11. https://doi.org/10.5334/joc.257

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317-324.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(2), 177–186. https://doi.org/10.1097/00004583-200502000-00010

Lövdén, M., Schaefer, S., Noack, H., Bodammer, N. C., Kühn, S., Heinze, H. J., Düzel, E., Bäckman, L., & Lindenberger, U. (2012). Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. *Neurobiology of aging*, *33*(3), 620.e9–620.e22. https://doi.org/10.1016/j.neurobiolaging.2011.02.013

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 563–582.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental psychology, 49*(2), 270.

Ng, J., Lee, K., & Khng, K. H. (2017). Irrelevant information in math problems need not be inhibited: Students might just need to spot them. *Learning and Individual Differences, 60,* 46–55. https://doi.org/10.1016/j.lindif.2017.09.008

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220-246. http://dx.doi.org/10.1037/0033-2909.126.2.220

Noël, M.-P., & De Visscher, A. (2018). Hypersensitivity-to-interference in memory as a possible cause of difficulty in arithmetic facts storing. In *Heterogeneity of function in numerical cognition* (pp. 387–408). Academic Press. https://doi.org/10.1016/b978-0-12-811529-9.00018-2

Oberauer, K. (2001). Removing irrelevant information from working memory. A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 948–957.

Oberauer, K. (2005). Control of the contents of working memory—A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 714–728.

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

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, *55*(4), 601–626.

Oberauer, K., & Lange, E. B. (2009). Activation and binding in verbal working memory: A dual-process model for the recognition of nonwords. *Cognitive Psychology, 58(1), 102–136.* doi:10.1016/j.cogpsych.2008.05.003

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19,* 779-819.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2007). *Individual differences in working memory capacity and reasoning ability.* In A. R. A. Conway, C. Jarrold, M. J. Kane (Eds.) & A. Miyake & J. N. Towse (Ed.), *Variation in working memory* (p. 49–75). Oxford University Press.

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

Psychology Software Tools. (2020). *E-Prime 3.0.* Psychology Software Tools, Inc. https://www.pstnet.com/eprime/

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.

RStudio Team. (2022). *RStudio: Integrated development environment for R* (Version 2022.02.0+443). RStudio, PBC. https://www.rstudio.com/

Sala, G., & Gobet, F. (2017). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*(4), 671-685.

Salminen, T., Strobach, T., & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience, 6*, Article 166. https://doi.org/10.3389/fnhum.2012.00166

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

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. Cambridge University Press.

Sternberg, R. J. (2020). The concept of intelligence. In R. J. Sternberg (Ed.), *The Cambridge handbook of intelligence* (2nd ed., pp. 3–17). Cambridge University Press.

Szmalec, A,Verbruggen, F.,Vandierendonck, A., & Kemps, E. (2011) *Control of interference during working memory updating. Journal of Experimental Psychology: Human Perception and Performance, 37*(1),137-151

Thevenot, C., Castel, C., Fanget, M., Fayol, M., (2010) Mental subtraction in high- and lower skilled arithmetic problem solvers: verbal report versus operand-recognition paradigms. *J. Exp. Psychol. Learn. Mem. Cogn. 36* (5), 1242–1255. http://dx.doi.org/10.1037/a0020447.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104−132.

Verhaeghen P., Cerella J., Basak C. (2004). A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*, 1322–1337.10.1037/0278-7393.30.6.1322

Vockenberg, K. (2006). *Updating of representations in working memory* [Unpublished doctoral dissertation]. University of Potsdam.

von Bastian, C. C., Langer, N., Jäncke, L., & Oberauer, K. (2013). Effects of working memory training in young and old adults. *Memory & Cognition, 41*(4), 611-624. https://doi.org/10.3758/s13421-012-0280-7

Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Ostensson, M. L., Bartfai, A., & Klingberg, T. (2007). Computerized working memory training after stroke--a pilot study. *Brain injury*, *21*(1), 21–29. https://doi.org/10.1080/02699050601148726

Wheeler, M. E., & Treisman, A. M. (2002). Binding in Short-Term Memory. *Journal of Experimental Psychology: General, 131*, 48-64. https://doi.org/10.1037/0096-3445.131.1.48

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 48*(4), 516–535. https://doi.org/10.1037/1196-1961.48.4.516

Zbrodoff, N.J. & Logan, G.D. (2005). *What everyone finds: the problem-size effect*. In: Campbell, J.I.D. (Ed.), Handbook of Mathematical Cognition.