SCHOOL AS A CONTEXT FOR DEVELOPMENT: EXAMINING FIRST-GRADERS' STRATEGY USE ACROSS DOMAINS

Kesha N. Hudson

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology in Neuroscience.

Chapel Hill
2019

Approved by:

Peter A. Ornstein

Jennifer L. Coffman

Patrick J. Curran

Jeff A. Greene

Beth E. Kurtz-Costes

ABSTRACT

Kesha N. Hudson: School as a Context for Development: Examining First-Graders’ Strategy Use across Domains
(Under the Direction of Peter A. Ornstein)

Data from four cohorts of first-grade teachers and students were used to (1) describe students’ strategy use on memory and mathematics tasks across the first-grade, (2) consider differences in teachers’ use of *Cognitive Processing Language* (CPL), or instruction that encourages deep levels of processing and metacognition over time and across subject matter, and (3) examine the effects of teachers’ CPL on students’ strategy use on memory and mathematics tasks. Minimal change was observed in students’ recall, accuracy, and strategy use on memory and mathematics tasks over the course of the year, which contributed to the failure to establish linkages between teachers’ use of CPL and students’ performance. The results highlight the need for additional research that carefully considers whether the tasks being used are developmentally appropriate and sensitive enough to capture change over the course of an academic year. In addition, efforts to further understand the association between strategy use and performance across domains should also include measures of metacognition and self-regulated learning.

ACKNOWLEDGEMNTS

Throughout the writing of this dissertation I have a received a great deal of support and guidance. I would first like to express my deepest appreciation to my advisors Drs. Peter Ornstein and Jennifer Coffman who provided invaluable mentorship throughout graduate school. Their unparalleled support was instrumental in the completion of this dissertation. I am also appreciative of my committee members, Patrick Curran, Jeff Greene, and Beth Kurtz-Costes for being engaged throughout this process and providing thoughtful and critical feedback on my work.

I must also recognize Dr. Steve Reznick, for inspiring my interest in developmental psychology. The freshman seminar I took with him as an undergraduate student introduced me to the field of psychology and his continued guidance helped me find my place at Carolina.

A special thanks to my lab mate Taylor Thomas. In Taylor, I found a friend whose shared love of concerts and iced coffee kept me sane. This journey would not have been possible with the unwavering support she continues to show me and I am grateful to call her a friend.

I would also like to acknowledge the funding agencies that made this work possible. The research reported here was supported in part by funding from National Science Foundation through Grant BCS-0217206 and Grant DRL-0843961. This work was also supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A120402. In addition to the support of NSF and IES, a predoctoral fellowship provided by the National Institute of Child Health and Human Development (T32-HD07376) through the

Center for Development Science provided additional interdisciplinary training opportunities that enriched my graduate education.

I am grateful to the school administrators, teachers, families, and students who participated in this research. Finally, I would like to recognize the time and effort of the team of research assistants and undergraduate students who were instrumental in carrying out this work over the last several years.

TABLE OF CONTENTS

LIST OF TABLES……..……………………………………………………………………...xi

LIST OF FIGURES…………………………………….……………………………….xiv

LIST OF ABBREVIATIONS………………………………………………………...xvi

INTRODUCTION……………………………………………………………..…1

Age-related Changes in Strategy Use…………………………..……………….…3

Strategy Use from a Developmental Perspective………..……………………………...7

The Role of Behavioral Self-Regulation……………………………………………..…11

School as a Context for Strategy Development………..………………………………13

Evidence from Mathematics Education Literature…………………………...14

Evidence from the Memory Development Literature……………………..…17

The Present Study……………………………………………………………….…23

Aim 1…………………………………………………………………....24

Aim 2……………………………………………………………………25

Aim 3………….…………………………………………………………….26

Aim 4.……...……………………………………………………………26

Aim 5……………………………………………………………………26

Aim 6……………….…………………………………………………….…27

METHOD………….……………………………………………………………..…..28

Sample and Demographics…………………………………………...……….28

Procedures and Measures……………………………………………………….31

Taxonomy of Teacher Behaviors………..…………………………………..…31

Instructional Activities……………………………………………….…32

Cognitive Structuring Activities……………………………..……32

Memory Requests……………..…………………………………..…33

Metacognitive Information……………………………………...…34

Cognitive Processing Language…………………………………...…34

Child Assessments…………… ….…………………………………….…36

Object Memory Task…………..……….………………………………36

Free Recall with Organizational Training Task……………..….………….....37

Math Problem Solving Task…………………………………………….…39

Head Toes Knees Shoulders………...……………………………………..40

Digit Span…………………………………………………………….……39

RESULTS…….……………………………………………………………………41

Aim 1: Describing Children's Memory and Math Performance……………………..41

Aim 2: Examining Associations in Performance within and across Tasks……......…43

Correlations between Strategy Use and Performance………………………..43

Correlations between Recall in Memory and Accuracy in Mathematics within Domain…….………………………………………………………43

Correlations between Recall in Memory and Accuracy in Mathematics across Domain………………………………………………………….…44

Correlations between Strategy Use within Domain…………………………45

Correlations between Strategy Use across Domains………..………………46

Aim 3: Characterizing the Classroom Context……………………………………46

Cognitive Processing Language…………………………………...……47

Aim 4: Examining CPL by Cohort……………………………………………...…49
CPL during Mathematics Instruction across Cohorts……………..…..……..50
CPL during Language Arts Instruction across Cohorts……...………………50
Aim 5: Comparing CPL in Mathematics and Language Arts…………………..……51
Aim 6: Linking CPL to Children's Memory and Mathematics Performance……..….52
Recall during the Object Memory Task…………………………...………53
Strategy Use during Object Memory Task……………………………...…55
Recall during the Free Recall with Organizational Training Task………...…56
Strategy Use during the Free Recall with Organizational Training Task…….58
Accuracy during the Mathematics Task…………………………………...…59
Strategy Use during the Mathematics Task…………………………..……60
Additional Analyses……………………………………………………..……62
DISCUSSION……..…………………………………………………………………..64
Children's Performance in Memory and Mathematics………………………………65
The Classroom Context and Teachers' Provision of Cognitive Processing Language……..68
CPL across Cohorts and Subject Matter……………..………………………………69
Linking CPL to Children's Performance……………………..………………………71
Limitations………...…………………………………………………………………73
Child-level Considerations……………………………………………..………73
Classroom-level Considerations……………………...…………………..……75
Future Directions………………………………………………………..………77
The "New" First Grade……………………………………………………..……78
The Role of Metacognition………………………………………...……79

The Role of Self-Regulated Learning……………………………………..….82

Appendix A: Exploratory Analyses…………….…………………………………………..85

Appendix B: Recall in Object Memory Task Unconditional Model…………………...………..94

Appendix C: Recall in Object Memory Task Random Intercept Regression
Model with Level 1 Predictors……………………....………………………..95

Appendix D: Recall in Object Memory Task Random Intercept Regression
Model with Level 1 and Level 2 Predictors...………..…………..…………..……96

Appendix E: Recall in Object Memory Task Slopes-as-Outcome Model
with Cross-level Interaction……...…………………………...…………………97

Appendix F: Strategy Use in Object Memory Task Unconditional Model……………..…….98

Appendix G: Strategy Use in Object Memory Task Random Intercept
Regression Model with Level 1 Predictors.……….……………………………99

Appendix H: Strategy Use in Object Memory Task Random Intercept
Regression Model with Level 1 and Level 2 Predictors.……………………...….100

Appendix I: Strategy Use in Object Memory Task Slopes-as-Outcomes
Model with Cross-level Interaction……….………….……………………...…….101

Appendix J: Recall in Free Recall with Organizational Training Task
Unconditional Model……,…………………………………………………..102

Appendix K: Recall in Free Recall with Organizational Training Task
Random Intercept Regression Model with Level 1 Predictors………..………….103

Appendix L: Recall in Free Recall with Organizational Training Task
Random Intercept Regression Model with Level 1 and Level 2 Predictors.......…104

Appendix M: Recall in Free Recall with Organizational Training Task
Slopes-as-Outcomes Model with Cross-level Interaction...……………………...105

Appendix N: Strategy Use in Free Recall with Organizational Training
Task Unconditional Mode………...……………………………...……………106

Appendix O: Strategy Use in Free Recall with Organizational Training
Task Random Intercept Regression Model with Level 1Predictors…………..……..107

Appendix P: Strategy Use in Free Recall with Organizational Training
Task Random Intercept Regression Model with Level 1 and

Level 2 Predictors………………………………………………………………...108

Appendix Q: Strategy Use in Free Recall with Organizational Training Task Slopes-as-Outcomes Model with Cross-level Interaction.……..…………..109

Appendix R: Accuracy Unconditional Model……………………….…………………………110

Appendix S: Accuracy Random Intercept Regression Model with Level 1 Predictors…………….……………………………………………………………....111

Appendix T: Accuracy Random Intercept Regression Model with Level 1 and Level 2 Predictors…………………………………………………………………112

Appendix U: Accuracy Slopes-as-Outcomes Model with Cross-level Interaction……………………………………………………………..……….113

Appendix V: Strategy Use in Mathematics Unconditional Model…………………….……….114

Appendix W: Strategy Use in Mathematics Random Intercept Regression Model with Level 1 Predictors..…………..……………………………………..115

Appendix X: Strategy Use in Mathematics Random Intercept Regression Model with Level 1 and Level Predictor…………………….…………………..…..116

Appendix Y: Strategy Use in Mathematics Slopes-as-Outcomes Model with Cross-level Interaction…………...……………………………….…117

REFERENCES………………...………………………………………………………………181

LIST OF TABLES

Table 1 – Task-Oriented Strategic Behaviors Coded in the Object Memory Task……..……..118

Table 2 – Code Definitions for the Taxonomy of Teacher Behaviors………………………….119

Table 3 – Component Codes from the Taxonomy Used to
Index Cognitive Processing Language……………………..…………..…………121

Table 4 – Sample Instruction from High CPL Classrooms
in Mathematics and Language Arts…….…..……………………………………..122

Table 5 – Children's Strategies (adapted from Siegler & Jenkins, 1989)…………………..….124

Table 6 – Child-level Tasks and Measures………………………………………………..….125

Table 7 – Summary of Overall Child-level Performance in the Fall and Spring………………126

Table 8 – Summary of Correlations between Recall and Strategy Use
during the Object Memory Task in the Fall and Spring………...……...……………133

Table 9 – Summary of Correlations between Recall and Strategy Use
in the Free Recall with Organizational Training Task in the Fall and Spring……....134

Table 10 – Summary of Correlations between Accuracy and Strategy Use
in Mathematics in the Fall and Spring…...……………………...………………..135

Table 11 – Summary of Correlations between Recall in the Object Memory Task,
the Free Recall with Organizational Training Task, and Accuracy in Math……..…136

Table 12 – Summary of Correlations between Strategy Use in the Object Memory Task,
the Free Recall with Organizational Training Task, and the Math Task……………137

Table 13 – Mean and Range of Percent Occurrence of Teacher Behaviors
for Whole Sample Across Subjects……………………………..……….………138

Table 14 – Mean and Range of Cognitive Processing Codes
for Whole Sample across Subjects…………..…………………...……………...141

Table 15 – Means and ANOVA Summary Statistics for CPL during
Mathematics and Language Arts across Cohorts…………..………………………144

Table 16 – Means and ANOVA Summary Statistics for CPL during
Mathematics across Cohorts ………....……………………..………………..…146

Table 17 – Means and ANOVA Summary Statistics for CPL during
Language Arts across Cohorts……..……...……………………………………148

Table 18 – Summary of Unconditional HLM Analysis for Recall in
the Object Memory Task……………………….…………………………..………155

Table 19 – Summary of Condition HLM Analysis for Recall in

the Object Memory Task with Level-1 Predictors…..…………………………..…..156

Table 20 – Summary of Conditional HLM Analysis for Recall in
the Object Memory Task with Level-1 and Level-2 Predictors…………………….157

Table 21 – Summary of Conditional HLM Analysis for Recall in
the Object Memory Task with Cross-level Interaction……………………..……158

Table 22 – Summary of Unconditional HLM Analysis for Strategy Use in
the Object Memory Task ………………………………………………....……..159

Table 23 – Summary of Conditional HLM Analysis for Strategy Use in
the Object Memory Task with Level-1 Predictors………………..………………...160

Table 24 – Summary of Conditional HLM Analysis for Strategy Use in
the Object Memory Task with Level-1 and Level-2 Predictors……...……………..161

Table 25 – Summary of Conditional HLM Analysis for Strategy Use in
the Object Memory Task with Cross-level Interaction……………..………………162

Table 26 – Summary of Unconditional HLM Analysis for Recall in the
Free Recall with Organizational Training Task……………..……………….……..163

Table 27 – Summary of Condition HLM Analysis for Recall in the
Free Recall with Organizational Training Task with Level-1 Predictors...……...…164

Table 28 – Summary of Conditional HLM Analysis for Recall in
the Free Recall with Organizational Training Task with Level-1
and Level-2 Predictors….……………………………………………………….… .165

Table 29 – Summary of Conditional HLM Analysis for Recall in
the Free Recall with Organizational Training Task with
Cross-level Interaction...……………………………………………………....…166

Table 30 – Summary of Unconditional HLM Analysis for Strategy Use
in the Free Recall with Organizational Training Task ………………..…………167

Table 31 – Summary of Condition HLM Analysis for Strategy Use in
the Free Recall with Organizational Training Task with Level-1 Predictors.......…168

Table 32 – Summary of Conditional HLM Analysis for Strategy Use in
the Free Recall with Organizational Training Task with Level-1
and Level-2 Predictors……………………………………………………....…169

Table 33 – Summary of Conditional HLM Analysis for Strategy Use in
the Free Recall with Organizational Training Task with
Cross-level Interaction……...……………………………………………...….170

Table 34 – Summary of Unconditional HLM Analysis for Accuracy in Mathematics……..…171

Table 35 – Summary of Condition HLM Analysis for Accuracy
in Mathematics with Level-1 Predictors……………………...………..……………172

Table 36 – Summary of Conditional HLM Analysis for Accuracy
in Mathematics with Level-1 and Level-2 Predictors………………………………173

Table 37 – Summary of Conditional HLM Analysis for Accuracy
in Mathematics with Cross-level Interaction……………...………..………………174

Table 38 – Summary of Unconditional HLM Analysis for Strategy Use
in Mathematics…………………………………..…………………….....……..175

Table 39 – Summary of Condition HLM Analysis for Strategy Use
in Mathematics with Level-1 Predictors…...……………...……………..…………176

Table 40 – Summary of Conditional HLM Analysis for Strategy Use
in Mathematics with Level-1 and Level-2 Predictors…………..…………………..177

Table 41 – Summary of Conditional HLM Analysis for Strategy Use
in Mathematics with Cross-level Interaction….……..…..………………………178

LIST OF FIGURES

Figure 1 – Number of Items Recalled during the Object Memory Task in the Fall and Spring……………...…………………………………………...127

Figure 2 – Amount of Time Spent Engaged in Helpful Behavior during the Object Memory Task in the Fall and Spring…...……………………..…128

Figure 3 – Number of Items Recalled during the Free Recall with Organizational Training Task in the Fall and Spring….……...…………..……129

Figure 4 – Sorting ARC Scores during the Free Recall with Organizational Training Task in the Fall and Spring………..…………….…..………130

Figure 5 – Number of Correct Answers Provided during the Mathematics Task……………...131

Figure 6 – Percentage of Effective Strategy Use during the Mathematics Task………..…...…132

Figure 7 – Provision of CPL during Mathematics and Language Arts Lessons Combined across Cohorts…………..……………………………..………142

Figure 8 – Provision of Instructional Activities and Memory Requests during Mathematics and Language Arts Lessons by Cohort...………...……………143

Figure 9 – Provision of CPL during Mathematics Instruction by Cohort………..…………..145

Figure 10 – Provision of Instructional Activities and Memory Requests during Mathematics Instruction by Cohort………….……………………………..…147

Figure 11 – Provision of Cognitive Structuring Activities and Memory Requests during Mathematics Instruction by Cohort….…………………149

Figure 12 – Provision of CPL during Language Arts Instruction by Cohort…….………………..150

Figure 13 – Provision of Instructional Activities and Memory Requests during Language Arts Instruction by Cohort……………….....………….……….151

Figure 14 – Provision of Metacognitive Questions during Language Arts Instruction by Cohort…………….……………...………152

Figure 15 – Provision of Metacognitive Information and Memory Requests during Language Arts Instruction by Cohort…………...…………………....…153

Figure 16 – Provision of CPL during Mathematics and Language Arts Instruction…………...154

Figure 17 – Percentage of Intervals Containing CPL among “Low” and “High” Teachers across Cohorts………...……………………..…………………………179

Figure 18 – Children’s Sorting ARC Scores across Cohorts…………………….……………..180

LIST OF ABBREVIATIONS

CPL	Cognitive Processing Language
FRT	Free Recall Task
OBJ	Object Memory Task

INTRODUCTION

Throughout the preschool and elementary school years children's performance improves across a range of cognitive domains. In the areas of memory and mathematics, in particular, these gains in performance are paralleled by corresponding changes in children's strategy use. To illustrate, consider the systematic transition from using passive or relatively simple techniques to employing more sophisticated and efficient strategies to meet task demands (Ornstein, Haden, & Elischberger, 2006; Schneider & Pressley, 1997). On deliberate memory tasks, for example, this transition is marked by a shift away from the use of simple single-item rehearsal techniques toward the adoption of more sophisticated multi-item rehearsal procedures to aid in performance (Lehmann & Haselhorn, 2010; Lehmann & Hasselhorn, 2012; Ornstein & Naus, 1978). To illustrate, when presented with a list of to-be-remembered words, younger children tend to rehearse each item alone as it is introduced, (e.g., *table, table, table*) whereas older children engage in cumulative rehearsal by repeating several previously presented stimuli together (e.g., *table, car, flower*) (Ornstein, Naus, & Liberty, 1975).

A similar pattern of development has been documented in the mathematics literature, whereby preschool-aged children evidence a shift away from concrete counting strategies, such as *sum*, in which both addends are first represented and then counted together, toward the more frequent use of abstract counting strategies, such as *min*, in which the sum is determined by counting-on from the cardinal value of the larger addend (Fuson, 1982; Siegler & Jenkins, 1989). Furthermore, across both domains, children's strategy use becomes more effective over time, such that strategy use is increasingly associated with greater recall in the context of memory tasks and greater accuracy in the context of mathematics tasks (Baker-Ward, Ornstein, & Holden, 1984; Bjorklund, 1987; Hudson, Coffman, Ornstein, 2018; Folds, Footo, Guttentag, & Ornstein, 1990; Lehmann & Hasselhorn, 2007; Siegler, 1987; Siegler & Robinson, 1982).

Given that improvements in performance and strategy use occur with age in both domains, it is important to consider which factors contribute to strategy acquisition and refinement. A number of researchers have demonstrated that individual differences in behavioral self-regulation (Duckworth & Seligman, 2005; McClelland, Acock, & Morrison, 2006; Blair & Razza, 2007; Duncan et al., 2007) as well as environmental factors, such as early educational experiences (Coffman, Ornstein, McCall, & Curran, 2008; Grammer, Coffman, Sydney, & Ornstein, 2016; Hudson et al., 2018), are associated with the development of children's cognitive and academic skills. In order to understand more fully the origins and subsequent refinement of children's strategy use it is imperative for future research to (1) examine strategy use across multiple domains over time and (2) consider interactions between child-level characteristics and the classroom context. As illustrated above, distinct bodies of work document comparable patterns of improvement in children's strategy use on memory and mathematics task over time. However, children's performance on memory and mathematics task has yet to be examined simultaneously. Exploring the association between children's domain specific strategy use in memory and mathematics during first-grade provides a unique opportunity to consider whether strategic knowledge and skill also become more generalized or integrated over time. In addition, examining the influence of both child- and environmental-level characteristics is important because it allows for an examination of specific subgroups of children who may be more or less prepared to succeed in the classroom.

In this dissertation, I take advantage of a unique opportunity to integrate data from four cohorts of first-grade teachers and their students to examine questions about the association among teachers' instructional language, children's self-regulation, and children's strategy use that remain unaddressed in the existing body of research. This study represents the first attempt to examine the development of children's strategy use within and across the domains of memory and mathematics during the first-grade school year. In addition, the relatively large sample size that is achieved by combining multiple cohorts will provide a unique opportunity to describe the classroom context with greater specificity than has previously been possible.

In the sections that follow, I begin by presenting an overview of the foundational research on age-related changes in children's strategy use in the domains of memory and mathematics. I then introduce findings from longitudinal and microgenetic studies that provide key insights into the developmental pattern of individual children's strategy use over time. Following a review of the literature on strategy development, I establish the importance of children's self-regulation in the context of the classroom. I then review two distinct lines of research that have examined the influence of formal schooling experiences and more specifically, teachers' instructional practices during the early elementary school years, on children's developing cognitive skills. Finally, I introduce the specific aims and hypotheses of the current project.

Age-related Changes in Strategy Use

During the elementary school years, children become increasingly facile at using strategies across a range of problem-solving contexts. A strategy can be defined as any procedure that is (1) nonobligatory such that it does not represent the only way to solve a problem, and (2) goal directed, in that it is intended to accomplish a specified purpose (Flavell, 1979; Harnishfeger & Bjorklund, 1990; Siegler & Jenkins, 1989). In the context of children's memory skills, a rich literature documents differences in children's use of rehearsal (Ornstein & Naus, 1978), organization (Lange, 1978), and elaboration (Rohwer, 1973) strategies. In a separate literature, researchers have also identified a number of strategies that children acquire to solve simple arithmetic problems, such as fact-based retrieval, counting on from one, and decomposition (Siegler & Jenkins, 1989).

Early depictions of strategy development are based on cross-sectional research designs that enable a characterization of the contrasting skills of children of different ages. Indeed, a considerable body of research demonstrates that as early as preschool children are quite strategic in the context of deliberate memory (Baker-Ward et al., 1984) and mathematics tasks (Siegler & Shrager, 1984). Moreover, early research on strategy use indicated that with age, children use increasingly sophisticated strategies more effectively. To illustrate, in the context of deliberate memory tasks, children systematically transition from employing relatively simple single-item techniques to more complex multi-item rehearsal procedures (Ornstein & Naus, 1978). When presented with a list of to-be-remembered words, younger children tend to rehearse each item alone as it is introduced, (e.g., *table, table, table*) whereas older children engage in cumulative rehearsal by repeating several previously presented stimuli together (e.g., *table, car, flower*) (Ornstein, Naus, & Liberty, 1975).

In one training study, Ornstein and colleagues (Ornstein, Naus, & Stone, 1977) demonstrated that this age-related change in rehearsal style was causally linked to corresponding differences in recall. Second and sixth grade students were assigned to one of three training conditions in which they were instructed to rehearse the currently presented item (1) alone (e.g., one-item group), in conjunction with the item immediately preceding it (e.g., two-item group), or to practice each word with as many of the preceding words as possible (e.g., multi-item group).

A control group, which served as a comparison of children's spontaneous use of the rehearsal strategy was told to practice the items aloud. Age-related differences in recall were observed, as evidenced by greater recall by sixth-graders compared to second-graders in all three training conditions. Importantly, the recall of non-instructed second-graders was similar to the performance of children who were instructed to rehearse only one or two items together. In contrast, the recall performance of second-grade students who were instructed to rehearse many items together was enhanced, especially in terms of the initial list items, and more comparable in form to that of the non-instructed sixth graders.

Younger children are also less likely than older children to (1) organize related items into groups on the basis of meaning and to (2) make use of this organized grouping during recall in an attempt to maximize memory performance. Just as researchers were able to train students to use a verbal rehearsal strategy, a consistent body of literature demonstrates that children can also be trained to use organizational techniques to facilitate recall of low-associated items. In a series of three experiments, Bjorklund and colleagues demonstrated that in spite of age related differences in third- fifth- and seventh-grade students' spontaneous use of organizational techniques, specific instructions to group items on the basis of meaning was sufficient to produce gains in recall (Bjorklund, Ornstein, & Haig, 1977).

Similar age-related changes in strategy sophistication have been observed for children's arithmetic strategies. Initial research on children's use of addition strategies suggested that children progress from using overt (e.g., counting on fingers) to covert (e.g., mental math) techniques and eventually to fact retrieval (Ashcraft, 1982; Groen & Parkman, 1972). Ashcraft (1987) further proposed that when children enter school they often add by counting from one. Sometime during first grade, however, they adopt a more advanced approach by counting from the larger addend, and by third or fourth grade, they are able to recall or retrieve the answer from memory – without using a specific strategy – for problems for which they know the answer. Empirical support for this trend was observed by Siegler (1987) who presented kindergarten, first grade, and second grade students with 45 addition problems. Clear age-related trends towards increasing use of retrieval and decreasing use of counting-all (e.g., counting each addend individually and then counting them together) were observed. Kindergartners used retrieval on 16% of problems whereas second-graders used retrieval on 45% of problems. In contrast, kindergartners used the counting all strategy on 22% of trials, whereas second-graders never used that strategy.

Across both domains older children are more likely to benefit from strategy use as evidenced by improved recall and accuracy. Consider the findings reported by Baker-Ward et al. (1984) in which 4-, 5-, and 6-year-old children were given instructions to remember a set of items and then during a two minute study period were observed to be naming and visually examining the to-be-remembered objects. Despite using the same strategy for remembering, only the 6-year-olds demonstrated superior recall. Comparable findings have been reported in studies of children's use of organizational strategies (see, e.g., Bjorklund, 1987; Folds, Footo, Guttentag, & Ornstein, 1990). To illustrate, consider a study by Harnishfeger, Bjorklund, and Halleck (1986) in which third- and seventh-grade students received training in the use of an organizational strategy in the context of a free recall task. Children were trained to identify categories among a list of items and to try to remember the items according to their category. Both third- and seventh-graders used the organizational strategy, as evidenced by increased clustering scores on the trained trial in comparison to the uninstructed trial. However, significant improvement in recall was only observed for seventh-graders.

Similar findings have been documented in studies of children's strategy use and accuracy on addition tasks. Although accuracy improved on problems in which a strategy was used in general, Siegler and Robinson (1982) also reported that age-related differences in accuracy were present such that 3-year-olds solved 20% of problems correctly, in comparison with 66% of 4-year-olds and 79% of 5-year-olds. In a similar study of elementary school students' use of arithmetic strategies, Siegler (1987) reported that the percentage of errors decreased from 51% for kindergartners to 5% for second graders.

Taken together, this body of cross-sectional work led researchers to conclude that strategy development proceeds in a gradual fashion, whereby older children use more advanced strategies and demonstrate more effective strategy use as evidenced by improvements in recall and accuracy. However, cross-sectional designs fail to consider variability in strategy use among different children of the same age and do not account for changes in individual's strategy use over time. This led researchers to question whether the developmental course of strategy use had been mischaracterized (Schneider & Weinert, 1995; Schneider, Kron, Hunnerkopf, & Krajewski, 2004). As will be seen in the following section, the emergence of longitudinal and microgenetic evidence highlighted the inadequacy of traditional "staircase models" of development that depicted development as a unidirectional progression toward increasingly sophisticated and consistently effective strategy use.

Strategy Use from a Developmental Perspective

In one of the first longitudinal studies of children's cognitive competencies, the Munich Longitudinal Study on Individual Development (LOGIC), researchers assessed children's use of organizational strategies as well as their recall on a sort-recall task when participants were 4, 6, 8, 10 and 12 years old. An analysis of group means for recall and strategy use revealed a pattern of development that is consistent with the findings reported in the cross-sectional literature, in terms of the extent to which children made use of organizational strategies and correlations between strategy use and recall that appeared to gradually increase over time.

However, additional analyses of the performance of individual children revealed that nearly 80% of children went from chance-level performance to near perfect organization between successive time points. In addition, strategy loss, or the failure to maintain the use of a strategy over time was also prevalent. Indeed, 70% of children who used an organizational strategy at the age 4 or 6 did not continue to use it at one or more later time points. In addition, considerable variation in the onset of children’s strategy use was also observed. Forty-percent of children engaged in conceptual organization by age 6, 24% by age 8, 21% by age 10, and 5% as late as age 12 (Schneider & Sodian, 1997; Schneider & Weinert, 1995; Schneider et al., 2004; Sodian & Schneider, 1999; Weinert & Schneider, 1999).

Evidence from microgenetic research carried out by Siegler and his colleagues provides additional insight into the variability and flexibility of children’s strategy use in mathematics. As exemplified by Siegler’s (1996) overlapping waves theory, at any given age children have available to them, and make use of, a variety of strategies. Indeed, as early as first grade the same child, given the same problem, will sometimes count from one, sometimes count from the larger addend, and sometimes retrieve the answer. Similarly, children will select different strategies to solve different problems on a trial-by-trial basis. Even when they have acquired more sophisticated strategies, such as decomposition, children nonetheless continue to use older strategies that are slower and less accurate, such as shortcut sum, as well (Carpenter & Moser, 1982; Fuson, 1982; Ginsburg, 1982; Siegler, 2003). To illustrate, in a number of studies researchers have reported that the majority of children use at least three distinct strategies when presented with a series of addition problems, including retrieval, decomposition, and counting from one (Siegler, 1987; Siegler & Robinson, 1982). Although children increasingly use more advanced strategies over time, it is not the case that simple and inefficient strategies are replaced completely with complex and efficient ones. Rather, at any given time children's strategic repertoire can be characterized by a wavelike pattern that consists of more- and less-advanced strategies that coexist and seem to compete for use (Siegler, 1989, 1996).

In order to accurately capture this variability, Siegler (1987) warned against the dangers of averaging data across strategies as well as averaging data across individuals. To illustrate, in previous work Groen and Parkman (1972) observed that the size of the smaller addend was the best predictor of first-graders' solution times on simple addition problems, which led them to propose that children of this age consistently use the min strategy (e.g., counting up from the larger addend) to solve such problems. To determine whether 5- to 7-year-olds consistently use the min strategy Siegler (1987) examined solution time as well as children's verbal reports of their strategy use on each trial. Importantly, when all problem-solving trials were analyzed together, the results were consistent with those reported by Groen and Parkman (1972) and indicated that the min strategy accounted for 76% of the variance in solution time. However, according to children's verbal reports, a total of five different strategies were applied across different trials, including shortcut sum, min, retrieval, decomposition, and guessing. To illustrate, children reported using min on 36% of trials, retrieval on 35% of trials, shortcut sum on 8% of trials, decomposition on 7% and guessing on 14% of trials. Importantly, the results of multiple regression analyses conducted separately for each strategy revealed a different pattern of performance. Indeed, on the trials in which children reported using the min strategy, the size of the smaller addend was still the best predictor of solution time, accounting for 86% of the variance. However, on the remaining 64% of trials in which children reported using a strategy other than min, the size of the smaller addend was no long a predictor of solution times.

Siegler and Jenkins (1989) also demonstrate that generalizations about children's use of addition strategies that are based on averages of group data may not actually reflect how any single child solved a problem. Whereas all children evidence multiple strategy use, individuals vary considerably in the relative frequency with which different strategies are used. For example, some children used a single strategy on a majority of trials but the specific strategy that was predominantly used varied, whereas others did not use any single strategy on a majority of problems.

As evidenced by the findings reviewed above, individual trajectories of development may be obscured by cross-sectional data. Therefore, generalizations about the developmental course of strategy acquisition that are based on comparison across different age groups can be unreliable. Longitudinal data, which reveal individual patterns of change, indicate that developmental changes in strategy use may be more accurately characterized by nonlinear and dramatic leaps in performance, rather than gradual and unidirectional increases (Schneider & Sodian, 1997; Schneider & Weinert, 1995; Schneider et al., 2004; Sodian & Schneider, 1999; Weinert & Schneider, 1999). Evidence from microgenetic studies provide additional support for this nonlinear progression of skill and also demonstrate that children continue to use multiple strategies over an extended period of time (Siegler & Jenkins, 1989).

Despite the rise of longitudinal and microgenetic approaches to study the developmental progression of strategy use, a critical gap remains with respect to understanding the mechanisms that underlie the emergence and subsequent refinement of strategic skill. Specifically, how do children become aware of strategies and how to execute them effectively? Children's ability to regulate their behavior to meet the learning- and task-oriented demands of the classroom is one factor that may influence strategy acquisition.

The Role of Behavioral Self-Regulation

Behavioral self-regulation involves the deliberate and flexible management of attentional resources, working memory, and inhibitory control (McClelland et al., 2007; Morrison, Ponitz & McClelland, 2010; McClelland & Cameron, 2011). These three elements are important for meeting school- and task-related demands, such as learning to following directions and persisting on difficult tasks. To illustrate, attentional flexibility involves the ability to ignore environmental distractions in order to focus on a particular task, such as listening to the teacher in a busy classroom (Barkley, 1997; Rothbart & Posner, 2005). Working memory involves cognitively maintaining and manipulating information, such as remembering and then implementing a teacher's instructions (Gathercole & Pickering, 2000; Kail 2003). Finally, inhibitory control aids children in preventing or modifying an incorrect behavior in favor of a more adaptive response, such as remembering to raise his or her hand in class instead of shouting out (Diamond, Kirkham, & Amso, 2002). Together, these components contribute to children's ability to successfully navigate learning environments by focusing their attention, remembering instructions, and controlling their behavior.

A large body of research indicates that children's ability to effectively regulate their own behavior by controlling these three aspects of executive function is associated with their academic success across multiple domains. To illustrate, in one longitudinal study researchers observed that higher teacher ratings of children's self-regulatory skills in kindergarten predicted greater growth in literacy and mathematics between kindergarten and second grade and also significantly higher reading and mathematics achievement through sixth grade (McClelland et al., 2006). In addition to teacher ratings, children's performance on observational measures of self-regulation has also been linked to academic achievement. In one study, researchers administered the Head-to-Toes task in which preschool-aged children were instructed to do the opposite of what the experimenter said. For example, when the experimenter said "touch your head" instead of following the command, the child was instructed to touch his or her toes, and vice versa. To be successful, children must apply the three cognitive skills associated with behavioral regulation to their overt motor movements by (1) focusing on instructions and commands, (2) using working memory to remember and execute new rules while processing commands, and (3) inhibiting the prepotent response in order to respond correctly (Ponitz, McClelland, Matthews, & Morrison, 2009).

Higher scores on the Head-to-Toes task, were correlated with greater achievement in literacy, vocabulary, and mathematics in the fall and spring of the prekindergarten year. Moreover, children who demonstrated greater gains on the Head-to-Toes task between the fall and spring also evidenced significantly greater gains on all three measures of academic achievement compared to their peers who demonstrated less growth in self-regulation over the prekindergarten year (McClelland et al., 2007). Comparable findings have been reported in a kindergarten sample using the Head-Toes-Knees-Shoulders (HTKS) task, a more complex version that includes instructions corresponding to four body parts instead of two. Children's behavioral regulation skills, as evidenced by performance on the HTKS task in the fall predicted end-of-year achievement on measures of literacy, vocabulary, and mathematics. Additionally, after controlling for prior achievement, children who scored higher in self-regulation at kindergarten entry, exhibited greater gains in mathematics over the course of the year than their less-regulated peer (Ponitz et al., 2009).

This literature consistently demonstrates that behavioral self-regulation skills are required for successfully navigating the demands of the classroom and are consequently predictive of academic success. If variations in children's self-regulatory skill contribute to their ability to learn in formal classroom settings then it is important to consider which aspects of the classroom environment may provide supportive opportunities for less-regulated students to remain actively engaged and benefit and from instruction. For example, it is possible that children who are less well-regulated may benefit from being in a classroom setting in which the teacher provides external supports through modeling or direct instruction to promote children's attentional skills and facilitate their awareness of their own thinking and problem-solving skills.

School as a Context for Strategy Development

A growing body of research documents the influence of the classroom context on the development of children's cognitive skills. Among mathematics education researchers, there has been an increasing interest in teachers' instructional practices, in response to standards-based reform efforts that emphasize the importance teacher quality and effectiveness. A distinct but complementary body of research has emerged from a series of cross-cultural investigations of children's memory development and emphasizes the role of formal schooling, and teachers' instructional language more specifically, on the development of children's strategy use in the context of deliberate memory tasks.

In the sections that follow, I review evidence of the association between instructional practices and corresponding gains in student performance from the perspective of researchers in mathematics education and memory development. In doing so, I point out that underlying both of these research traditions is an emphasis on social-constructivist theory (Wood, Bruner, & Ross, 1976; Vygotsky, 1978; Brown & Reeve, 1987) and the importance of scaffolded, social interactions for student learning.

Evidence from mathematics education literature. Over the last two decades, researchers interested in the development of children's mathematical skills have emphasized the role that teachers' instructional language plays in the development of children's mathematical competencies. One feature of instruction that has been considered is teachers' use of mathematically relevant language or "math talk" (Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006). Klibanoff and colleagues observed preschool teachers and coded the number of times that math-relevant language, such as references to counting, equivalence, or calculation, was included in classroom activities. Dramatic differences in the amount and diversity of math-related talk were observed across classrooms. During the hour-long observation interval, the amount of math input teachers provided ranged from 1 to 104 instances, and the type of input varied between 1 and 9 different types. Importantly, teachers' math-talk was related to the growth of children's mathematical knowledge over the school year, such that preschoolers who experienced higher amounts of math talk evidenced greater growth in mathematical knowledge than their peers who experienced lower levels of math talk.

Beyond the amount of math talk to which children are exposed, findings from an observational study of preschool classrooms suggest that there is also variability in the type of mathematical content that teachers provide. Rudd et al. (2008) reported that lower-level mathematical concepts such as a counting and labeling amount, describing location with spatial words (e.g., in, on, between), and using comparative words for size, length, or weight, are discussed much more frequently than higher-level mathematical concepts such as, demonstrating counting-based strategies and providing opportunities to identify patterns or to sort or compare items in groups. In a similar study, Engel, Classens, and Finch (2013) reported that kindergarten teachers teach basic counting and shapes nearly 13 days per month, whereas only 4 days per month are spent teaching addition and subtraction. Importantly, an analysis of the relation between mathematics content and student achievement revealed a negative association between teachers' emphasis on lower-level concepts, such as counting and shapes, and end-of-grade mathematics test scores. Devoting additional days per month to teaching more advanced concepts, such as place value, currency, addition, and subtraction, was positively associated with students' end-of-grade mathematics test scores.

There is also evidence to suggest that in addition to the content of mathematical instruction, the nature of instruction is also important for student learning. In this regard, researchers have examined (1) the extent to which teachers and students co-construct a mathematical dialogue that serves to promote students' understanding of and ability to explain their own thinking, as well as (2) the influence of teachers' understanding of children's mathematical thinking on their own instructional practices (Carpenter, Fennema, Franke, Levi, & Empson, 1999; Cobb, Boufi, McClain, & Whitenack, 1998). Productive mathematical discourse is fostered when teachers elaborate on student ideas, encourage students to generate, explain, and compare multiple problem-solving strategies, question student reasoning, and create opportunities for justification and argumentation. Engaging in rich mathematics discourse has been linked to students' ability to correctly answer and explain their problem-solving strategies on a range of mathematical reasoning tasks (Webb et al., 2008; Webb, et al. 2013).

Teachers' conceptual understanding of the development of children's mathematical thinking has also been related to children's outcomes in a series of studies. Cognitively Guided Instruction (CGI) is a professional development program designed to enhance teachers' understanding of students' mathematical thinking in order to improve instructional practices. In the first CGI study, Carpenter and her colleagues (1989) examined the effect of CGI professional development on teachers' knowledge, beliefs, and instruction. Not only did teachers who participated in the professional development program demonstrate greater knowledge about their students' thinking than did control teachers, but they also (1) placed greater emphasis on problem solving skills and less on computational skills, (2) expected students to provide multiple strategies for problem-solving, and (3) listened to their students more. As a result, students in CGI classrooms evidenced significantly higher levels of achievement on measures of problem-solving skills than their peers in control classrooms.

Additional longitudinal research (Fennema, Carpenter, Franke, Levi, Jacobs, & Empson, 1996) provided confirmatory evidence of the effects of CGI on student outcomes. First, second, and third grade teachers who participated in a CGI workshop were followed for four years. Changes in the instructional practices of individual teachers were related to changes in their students' achievement on measures of mathematical concepts, such as place value and solving addition, subtraction, multiplication, and division problems. Moreover, the influence of CGI appeared to be cumulative, such that repeated exposure to CGI across multiple years of school led to greater gains in student achievement in the second and third years of the study (Carpenter, Fennema, Franke, Levi, & Empson, 2000; Fennema, et al., 1996).

The quality of classroom instruction has also been examined from a more global perspective by researchers using the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, Hamre, 2008). The CLASS was developed to characterize teaching across three broad domains – emotional, organizational, and instructional support. The instructional support domain in particular captures the extent to which teachers (1) foster the development of students' metacognitive knowledge, (2) promote higher-order thinking skills, (3) use feedback to extend understanding, and (4) stimulate language development. In a series of studies, higher scores on the instructional support domain were shown to be positively associated with preschool children's performance on standardized math measures, such as the Applied Problems subscale of the Woodcock Johnson Tests of Achievement (Keys et al. 2013; Burchinal et al., 2008; Crosnoe et al., 2010).

Additional findings reported by Mashburn and colleagues (2008) indicate that preschool children in classrooms that were rated as higher in the domain of instructional support also evidenced greater gains in mathematics achievement over the course of the year than their peers in classrooms that were rated lower on this domain. Similar findings were reported by Burchinal and colleagues (2010) who observed that the instructional support domain of the CLASS was positively associated with children's math performance, such that children from classrooms that scored higher on this domain tended to score higher on standardized measures of math achievement. In addition to documenting an overall effect of instructional quality on mathematics performance, Burchinal et al. (2010) also observed that the magnitude of the association varied in relation to the quality of teacher-child interactions. A series of spline regression models indicated that instructional quality, as measured by the CLASS, predicted children's mathematics performance more strongly in higher quality classrooms than classrooms in the low-to-average range.

Evidence from the memory development literature. Memory researchers began exploring the influence of formal schooling on children's strategic development in a series of cross-cultural investigations. In a review of the cross-cultural memory literature, Rogoff (1981) noted that the use of deliberate techniques for remembering is not present among non-schooled children and concluded that school seemed necessary for the acquisition of these skills. To illustrate, in Morocco, researchers contrasted the performance of children matched in chronological age but who differed in terms of whether or not they experienced Western-style schooling. Wagner (1978) observed that although memory capacity did not vary as a function of schooling experience, as evidenced by comparable rates of recall across both groups, only those students who attended school demonstrated the primacy effect in free recall, which suggests the use of a deliberate rehearsal strategy. Consistent findings were observed by Cole and his colleagues in Liberia, such that only those children who attended government-sponsored secondary school made use of meaning-based organizational strategies during a free recall task (Sharp, Cole, & Lave, 1979). These findings suggest that there is something specific about the formal school context that facilitates children's strategic behavior in the context of deliberate memory tasks.

These initial findings of the impact of formal schooling were extended by researchers in the 1980s who began examining cross-national differences in educational practices between Germany and the United States (Carr, Kurtz, Schneider, Turner, & Borkowski, 1989; Kurtz, Schneider, Carr, Borkowski, & Rellinger, 1990; Schneider, Borkowski, Kurtz, & Kerwin, 1986). German teachers reported more direct instruction in the use of task-specific strategies, such as relating new information to old material, organizing, and elaboration, than did American teachers (Carr et al., 1989, Kurtz et al., 1990). Moreover, these differences in teachers' strategy instruction were paralleled by systematic differences in American' and German children's strategy use in a sort-recall task. During the study period, German children engaged in organized sorting at earlier ages and more often than their American peers. In addition, German children also demonstrate higher rates of clustering, or organized grouping during recall, at younger ages than their American peers. Taken together, the results from cross-cultural and cross-national research suggest that formal schooling, in general, and teachers' instructional language more specifically, are associated with the development of students' strategic behaviors.

During the 1990's additional efforts were made to examine the classroom context in greater detail as researchers began conducting observations in elementary school classrooms that captured the linguistic nature of the social interactions that take place during the course of instruction. Mercer (1996) surveyed first grade teachers and found that although teachers consider memory to be important in their classrooms, they do not explicitly teach mnemonic skills. Consistent with these reports, observational evidence also suggested that direct strategy instruction is rare. Moely and her colleagues (1992) observed mathematics and language arts lessons in kindergarten through sixth grade classrooms and reported that teachers frequently seek correct answers (seen in an average of 32.3% of coded intervals), acknowledge student responses (27.8%), describe procedures (27.1%), and provide specific task-related information (26.1%). In contrast, teachers spent much less time teaching children about cognitive processes and strategies they could use to approach learning tasks. Teachers' referenced potential strategies in 2.28% of observed intervals and provided rationales or explanations of strategies in less than 1% (.47%) of intervals. Ten percent of the teachers in the study did not provide any strategy instruction during the observationf period. These findings demonstrate that although strategy instruction is rare, there is nonetheless considerable variability across teachers.

Coffman and Ornstein and their colleagues replicated and extended the initial cross-sectional work of Moely et al. (1992) in a series of longitudinal studies. An observational coding system, the *Taxonomy of Teacher Behaviors*, was developed to identify important variability in the extent to which teachers use metacognitively-rich language during mathematics and language arts instruction that promotes the deep processing of information by encouraging students to actively monitor and assess their own thinking and understanding. Differences across teachers in the use of these language forms were captured in Coffman et al.'s measure of *Cognitive Processing Language* (CPL)[1]. To illustrate, using a median split, Coffman and her colleagues observed that in classrooms taught by high CPL teachers an average of 7.32% of coded intervals contained a strategy suggestion, such as "*You can use a tens frame to help you solve the problem*," and that 6.90% of intervals contained metacognitive language, such as "*Who can think of another way we could solve this problem?"* Low CPL teachers, in constrast, provided strategy suggestions (2.50% of intervals) and engaged in metacognitive questioning (2.86% of intervals) comparatively less often than their high CPL colleagues.

In addition to examining the frequency with which teachers employ various types of instructional language, Coffman et al. (2008) also explored associations between teachers' use of CPL and their students' strategy use in the context of deliberate memory tasks over the course of the first-grade year. Consider students' performance on a Free Recall with Organizational Training task (Moely et al., 1992) in which the children were presented with 16 pictures drawn from four semantically related categories (e.g., spoon, fork, cup, bowl) and instructed "to work to remember them." Following a baseline trial, each child was trained to organize the cards into meaning-based groups during the study period and during recall. After a 15-minute delay, children completed a generalization trial using a new set of pictures. At the first assessment point in the fall of first grade, children's performance on the baseline and generalization trials did not vary as a function of teachers' use of CPL. However, differences in children's response to training were evident on generalization trials at the second and third time points, such that

[1] Coffman et al. (2008) referred to the measure as *Mnemonic Style* but it has since been relabeled as *Cognitive Processing Language* (Grammer, Coffman, & Ornstein, 2013).

children whose teachers used more CPL during instruction were more likely to sustain the use of an organizational strategy following a training trial than their peers taught by teachers who used less CPL.

A comparable pattern of results was reported for children's performance on an Object Memory task (Baker-Ward et al., 1984) that was designed to assess the types of strategies that children employ as they attempt to remember a set of low-associated stimulus objects. The amount of time that children spent engaged in strategic behaviors such as visual examination, naming, or making verbal associations was recorded (See Table 1 for a complete list of strategic behaviors). As expected, there were no differences between children's performance as a function of teachers' language at the beginning of first-grade, but by the end of the year, teachers' use of CPL was associated with children's recall and strategy use. Children exposed to greater amounts of CPL remained engaged in strategic behaviors for a longer period of time during the 2-minute study period and had higher recall scores, than their peers who experienced lower levels of CPL.

In the most recent set of studies, Coffman, Ornstein, and their colleagues have extended their examination of the association between teachers' instructional language and children's developing memory skills to include measures of children's academic achievement and indicators of strategy use on mathematics and study skills tasks. Grammer, Coffman, Sydney, and Ornstein (2016) demonstrated that the extent to which teachers incorporate CPL into mathematics lessons is related to growth in students' math achievement over the course of second grade. Children whose teachers used greater amounts of CPL evidenced greater growth on the Math Fluency and Calculation subscales of the Woodcock Johnson Tests of Achievement than did their peers who were exposed to lower levels of CPL (Grammer et al., 2016). Similar findings have been reported in studies of the association between teachers' use of CPL and students' strategy use on mathematics and study skills tasks. Kindergarten students who were exposed to greater amounts of CPL were more effective at using addition strategies to correctly solve single-digit addition problems than their peers who experienced lower levels of CPL (Hudson et al., 2018). Moreover, data from a longitudinal investigation revealed that exposure to CPL during the first grade has a sustained impact of on children's strategic skill later in elementary school. Indeed, first grade teachers' use of CPL predicted children's use of study strategies, such as highlighting or underlining, notetaking, and self-testing during fourth grade (Coffman et al., 2018).

In addition to the observational work carried out in elementary school classrooms, Grammer et al. (2013) carried out an experimental study in after-school programs in order to determine if there was a causal linkage between teachers' instructional language and children's performance. In this study, first and second graders participated in a two-week LEGO-based science and engineering unit involving *Things that Move*. Although the activities and material covered in the lessons were the same for all participants, the language used by the teachers to present the content was manipulated in a manner that captured the higher versus lower levels of CPL used by teachers in the Coffman et al., 2008 study. The results of this experiment demonstrate that all students acquired new factual knowledge as a result of instruction, but also that those students assigned to receive higher CPL exhibited greater levels of strategic knowledge and were better able to apply this knowledge in the service of a memory goal than their peers who has been exposed to lower levels of CPL. Indeed, when presented with a set of 15 cards containing pictures of Lego pieces (e.g., plates, beams, gears) and instructed to work to remember the cards, children exposed to more CPL engaged in more strategic sorting behaviors, as evidenced by the extent to which the cards were sorted into meaning-based groups (e.g., all gears placed together, regardless or color or shape).

The Present Study

As illustrated above, similar patterns of strategy development have been consistently documented in memory and mathematics research. Moreover, during the last two decades, observational and experimental findings from the Classroom Memory Study consistently indicate that the instructional context of the early elementary school classroom is associated with the emergence and refinement of children's strategic skill in both domains. However, due to the time intensive nature of collecting and coding observational data at both the student- and teacher-level, investigations of the classroom context have been limited to relatively small sample sizes. To illustrate, in the original examination of the first-grade classroom Coffman et al. (2008) recruited 107 students from 14 classrooms. Similar samples were recruited in later replication studies, ranging from 87 students recruited from 14 classrooms to 130 students recruited from 17 classrooms. As a result of the relatively small sample of classrooms in which observations have been carried out in each of these studies, researchers have relied on the use of a median split to form groups of teachers based on contrasting levels of CPL. In the current project, data collected from four cohorts of first-grade teachers and students between the 2002-2003 and 2015-2016 academic year are combined. By making use of multiple samples, it should be possible to capitalize on a larger sample size and to examine teachers' provision of CPL in a continuous fashion.

In addition to characterizing the context of the first-grade classroom, several indicators of children's performance are also being examined. More specifically, children's performance is characterized in terms of recall on memory tasks, accuracy on a mathematics tasks, and strategy use in both domains. In the existing literature, little is known about how children's strategic competence develops across domains over time. This question is being addressed in a preliminary fashion in this study by examining associations in performance within and across domains at the beginning and end of the year. Finally, observational measures of children's self-regulation are included to determine whether child-levels factors moderate the effect of teachers' CPL on student outcomes. As outlined in the following section Aims 1-4 are descriptive in nature and are included to provide relevant background information about (1) the patterns of association between children's performance and strategy use across the first-grade year and (2)to characterize the general instructional context of first-grade classrooms over time and across subject matter domain. Aims 5 and 6 address more substantive analytic questions regarding the effect of child-level and teacher-level effects on student performance. Specific aims are outlined below.

Aim 1: To describe children's developing competencies in memory and mathematics across first grade.

Hypothesis 1: Children will demonstrate greater recall on the Object Memory Task and the Free Recall Task with Organization Training at the end of the first grade in comparison to their initial performance at the beginning of the year.

Hypothesis 2: Children's strategy use on OBJ will improve over the course of first grade as evidenced by increases in the amount of time that children spend engaged in strategic behaviors during the two-minute study period.

Hypothesis 3: Children's strategy use on the Free Recall with Organizational Training task will improve over the course of the first grade as evidenced by (1) higher sorting ARC scores at the end of first-grade in comparison to the beginning of the year.

Hypothesis 4: Children will answer more problems correctly on the mathematics task at the end of the first grade in comparison to their initial performance at the beginning of the year.

Hypothesis 5: Children will use addition strategies to solve problems correctly more frequently at the end of first-grade than they do at the beginning of the year.

Aim 2: To examine the association among recall, accuracy, and strategy use within and across domains at the beginning and end of first grade.

Hypothesis 1: A significant positive association between strategy use and recall during both memory tasks and strategy use and accuracy during the mathematics task is expected to be observed in both the fall and spring. The strength of this association is expected to increase over the course of the year as strategy use becomes less cognitively demanding over time.

Hypothesis 2: At both time points, correlations between strategy use within domain are expected to be higher than correlations between strategy use across domains. The amount of time that children spend engaged in "helpful" strategic behaviors FRT task are expected to be more highly correlated with each other than with children's effective strategy use on the addition task. Similarly, children's performance is expected to be more strongly correlated within task, such that recall during OBJ will be more strongly correlated with recall during the FRT task than with children's accuracy in math.

Hypothesis 3: Correlations between strategy use, both within and across domains, are expected to increase across the year.

Aim 3: To describe the types of language that teachers use during instruction, including their use of CPL.

Hypothesis 1: Teachers' instructional language will predominantly consist of providing general information as well as task specific information. Teachers will also spend a considerable amount of time making connections with previously learned material and asking students to remember information. Teachers will provide comparatively less metacognitive language during the course of instruction.

Aim 4: To identify potential cohort effects between the four cohorts of teachers in terms of whether the provision of CPL changed between 2002-2003 and 2015-2016 school years?

Hypothesis 1: Overall CPL scores are expected to increase over time as a result of the introduction of the Common Core (NGACBP, 2010) which emphasizes higher-order thinking. More specifically, rates of cognitive structuring activities, strategy suggestions, metacognitive questions, and memory requests paired with metacognitive information are expected to increase over time.

Aim 5: To compare the provision of CPL during mathematic and language arts instruction.

Hypothesis 1: It is expected that individual teacher's CPL scores will be correlated across subject matter. However, CPL is expected to occur more frequently in mathematics lessons compared to language arts lessons due to the more explicit emphasis on strategy use during mathematics identified by the Common Core (NGACBP, 2010)

Aim 6: To examine linkages between students' self-regulation, teachers' use of CPL and students' strategic behaviors and performance in memory and mathematics at the beginning and end of first grade.

Hypothesis 1: Children's strategy use and performance on each task is expected to vary as a function of self-regulation at the end of first grade, such that more highly regulated students will demonstrate better performance compared to their less well-regulated peers.

Hypothesis 2: Children taught by teachers who use greater amounts of CPL will evidence greater strategy use and performance on memory and mathematics tasks than their peers taught by teachers who use less CPL at the end of the year.

Hypothesis 3: A significant interaction is expected between students' self-regulation skills and teachers' use CPL, such that CPL will be more strongly associated with strategy use and performance in memory and mathematics for those students who are low in self-regulation. Children who are low in self-regulation may benefit more from external supports provided by teachers that orient them to learning goals and relevant strategies.

METHOD

Sample and Demographics

Data were collected across four cohorts of first-grade teachers in the Durham Public Schools and the Chapel Hill-Carrboro City Schools. With the assistance of the administrations of these two public school systems, first-grade teachers were recruited to participate in the study. The teachers were told that the purpose of the study was to examine the classroom as a context for children's cognitive development and agreed to be observed as they taught regularly-scheduled lessons in mathematics and language arts. Letters were sent to families of all children in the participating classrooms, and any student who returned a consent form was enrolled, with no criteria for exclusion.

During the 2002-2003 school year, the first cohort of teachers was recruited from four elementary schools, two schools in each district. A sample of 15 teachers was recruited from 14 classrooms (two teachers co-taught). All of the teachers were female – 12 Caucasian and 3 African-American – and their average age was 36 years (range 23-51 years). They had a mean of 10.6 years of teaching experience (range 1-30 years), with an average of 7.5 years of teaching first grade (range 1-30). Ten of the teachers received Bachelor's degrees, whereas four had Masters degrees, and one had received a Ph.D. A sample of 107 children, 49 boys and 58 girls, was also recruited. At the initial time point, the children's mean age was 79 months (range 71-90 months) or 6 years, 7 months. The diversity of the sample was representative of the school systems from which the participants were drawn, with 45% of families describing their ethnicity as Caucasian, 26% as African-American, 15% as Asian, 10% as Mixed ethnicity, and 4% as Hispanic. Of the total sample, 26 children (24%) received free or reduced lunch at school[2].

The second cohort of teachers was recruited during the 2010-2011 school year from four elementary schools, two schools in each district. A sample of 17 first-grade teachers agreed to participate. All of the teachers were female – 14 Caucasian and 3 African-American – and their average age was 38 years (range 22-62). They had a mean of 11 years of teaching experience (range 1-26), with an average of 6.2 years of teaching first grade (1-18). Twelve of the teachers' received Bachelor's degrees and five had Masters degrees. A second cohort of 130 first-graders, 69 boys and 61 girls, was recruited during the 2010-2011 school year. At the initial time point the children's mean age was 81 months (range 72-95) or 6 years, 9 months. The diversity of the sample was representative of two school systems from which the participants were drawn, with 43% of families describing their ethnicity as Caucasian, 21% as African-American, 16% as Hispanic, 12% as Mixed ethnicity, and 8% as Asian. Of the total sample, 29 children (22%) received free or reduced lunch at school.

During the following academic year, 2011-2012, a third cohort of first-grade teachers was recruited from four elementary schools, all in the Chapel Hill-Carrboro school district. A sample of 15 teachers was recruited. All of the teachers were female – 11 Caucasian, 2 African-American, 1 Hispanic, and 1 Native American – and their average age was 36 years (range 24-60). They had a mean of 12.6 years of teaching experience (range 1-30) with an average of 6.9 years of teaching first grade. Five of the teachers held Bachelor's degrees, eight of the teachers had Masters degrees, and two teachers declined to report their level of education. A sample of 100 children, 50 boys and 50 girls, participated. At the initial time point the children's mean age

[2] Data from Cohort 1 were collected during the 2002-2003 school year and findings have been previously reported in Coffman et al., 2008.

was 78 months (range 60-84) or 6 years, 6 months. The diversity of the sample was representative of the Chapel-Hill Carrboro City Schools from which the participants were drawn, with 62% of families describing their ethnicity as Caucasian, 15% as Hispanic, 10% as Asian, 7% as Mixed Ethnicity, and 6% as African-American. Information about which students received free or reduced lunch at school was not collected.

The fourth cohort of teachers was recruited during the 2015-2016 school year. A sample of 14 teachers from four elementary schools in the Chapel Hill-Carrboro City Schools participated. All of the teachers were female – 12 Caucasian and 2 African-American. Eleven of the 14 teachers completed demographic information. Their average age was 35 years (range 24-53) and they had a mean of 11.8 years of teaching experience (range 2-49) with an average of 5.5 years of teaching first grade. Seven of the teachers held Bachelor's degrees, and four of the them had Masters degrees. A final cohort of 87 first-graders, 45 boys and 42 girls, was recruited during the 2015-2016 school year. At the initial time point the children's mean age was 82 months (range 71-90) or 6 years, 9 months. The diversity of the sample was representative of the Chapel-Hill Carrboro City Schools from which the participants were drawn, with 66% of families describing their ethnicity as Caucasian, 11% as Mixed ethnicity, 7% as Asian, and 6% as African-American. Information about which students received free or reduced lunch at school was not collected.

A total of 60 first-grade teachers participated between the years of 2002 and 2016. All of the teachers were female – 49 Caucasian, 10 African-American, 1 Hispanic, and 1 Native American – and their average age was 37 years (range 22-62). They had a mean of 11 years of teaching experience (range 1-38), with an average of 7 years of teaching first grade. A total of 424 first-graders, 213 boys and 211 girls, participated between the years of 2002 and 2016. At the initial time point the children's mean age was 80 months (range 60-95) or 6 years, 7 months. Fifty-two percent of families described their ethnicity as Caucasian, 18% as African-American, 10% as Hispanic, 10% as Asian, and 10% as Mixed ethnicity. As indicated above, information about whether students qualified for free or reduced lunch was only collected in Cohorts 1 and 2, in which a total of 55 children (23%) received such assistance.

Procedures and Measures

Extensive classroom observations were carried out in each of the participating first-grade classrooms, focusing on the extent to which teachers used language that encouraged deep levels of processing and metacognitive awareness during instruction. Consistent with previous work, a total of 60 minutes of mathematics instruction and 60 minutes of language arts instruction was observed in each classroom (Coffman et al., 2008). Across the four cohorts, individual mathematics and language arts lessons ranged from 2 .5 to 30 minutes, and it took from two to nine visits per subject, per classroom to complete the observations. Lesson topics in mathematics ranged from counting and addition to patterns and shapes. In language arts lesson topics range from verb tenses and digraphs to main ideas and using graphic organizers. This variety allowed for a broad characterization of teachers' instructional language across different topics and content areas. In Cohort 1, classroom observations were live coded, whereas in Cohorts 2, 3, and 4 classroom observations were videotaped for subsequent analysis using the *Taxonomy of Teachers Behaviors*.

Taxonomy of Teacher Behaviors. Using the *Taxonom*y, it is possible to make judgments about the degree to which each teacher incorporates language that promotes deep levels of processing and metacognition. As can be seen in Table 2, the *Taxonomy* classifies teachers' instructional language into four broad categories: instructional activities (providing information), cognitive structuring activities (fostering engagement with the materials in order to promote encoding and retrieval of information), memory requests (asking students to retrieve previously acquired information or plan for future assessment), and metacognitive information (providing or requesting information that might facilitate future performance on cognitive tasks in the classroom).

Instructional activities. Instruction codes were used in intervals when the teacher provided the class with *specific task information* or instructions for performing a particular activity, such as how to form the letter W; *general information* through the presentation of new factual material, such as defining what a homonym is; a *prospective summary* of upcoming events or lessons, such as describing the order of the morning's activities; or if she engaged in *book reading* or read from any other written source, such as a poem or class worksheet, for more than a single phrase.

Cognitive structuring activities. Cognitive structuring activities codes were employed when teacher instruction directed student attention to the ongoing lesson and encouraged student engagement with materials in ways that are known to facilitate deeper processing, which is associated with improved recall (Craik & Lockhart, 1972). *Attention regulation* was coded when teachers focused the attention of children for the purpose of behavioral control or in the service of imminent instructional goals. For example, "eyes on me" would be coded as an instructional goal where as "everyone should be sitting criss-cross applesauce" would be coded as a behavioral goal. *Massed repetition* was used any time a teacher instructed students to perform an activity in unison, such as singing a song or reading a phrase together. The *identifying features* code described instances in which the teacher or children discussed features of a category, such as the defining characteristics of a square. *Categorization* codes referred to the classification of conceptual or material items into at least two categories, by either the teacher or students, such as deciding if numbers are odd or even. *Identifying relationships* was coded when the teacher or children compared or contrasted two concepts or objects, such as determining which number in a series is the greatest. *Making connections with personal experiences* was used when a teacher asked children to associate a prior experience to a current one, for example asking students for names of common items from home to include in a poem. *Drawing inferences* codes were used when a teacher asked children to predict an outcome that had not yet been explained, such as predicting what will happen next in a story. Finally, *visual imagery* was coded when a teacher asked the children to create visual images relevant to the current discussion, such as, "imagine yourself as an animal."

Memory requests. These codes were employed when a teacher asked students to remember information about past experiences, retrieve previously learned facts, or to prepare for future memory assessments. Each memory request was first categorized by the type of question. Reports of prior knowledge or experiences could be *semantic* (as in the report of an already learned fact, such as "How many sides does a triangle have?"), *episodic* (as in recall of a specific event, such as, "What did you do at your birthday party?"), or *procedural* (as in recalling how to perform a series of activities, such as, "What is the first step in solving this problem?"). Memory requests involving future activities were coded as *prospective* when the teacher assigned a behavioral goal (such as, "Don't forget to skip lines when you are writing."), or *anticipated* when the teacher stated a learning goal (such as, "I want you to remember that a sentence always starts with a capital letter."). Additionally, each memory request could be coded as either an *expressed* or *implicit* demand. *Expressed* demands directly reference memory (such as, "Who remembers what 2 + 2 is?") whereas *implicit* demands do not (such as, "How many are 2 and 4 all together?").

Metacognitive information. These codes were used in intervals when teachers provided metacognitive information through direct instruction in the form of a strategy suggestion or provision of a rationale for the use of a particular strategy. Metacognitive information codes were also employed when teachers solicited metacognitive information from the students. When teachers recommended that children adopt a method or procedure to help them remember or process information a *suggestion* was coded (such as, "Look at the picture to help you figure out the word."). *Metacognitive Rationale* codes were used when the teacher provided justification for using a specific strategy, for example "Using a weekly planner will help you organize your assignments so you can remember all of them." *Suppression* codes were used in intervals when the teacher instructed children to refrain from using an ineffective strategy, such as "Don't just guess." When teachers recommended an alternative, more effective strategy, such as "Instead of just guessing about what will happen next, look carefully at the pictures and think about what they tell you about the story," a *replacement* was coded. In addition to providing metacognitive information, teachers also solicited it from their students. When teachers asked children to provide a potential strategy or a rationale for a strategy they have used (such as, "How did you know you were supposed to subtract?") a *metacognitive questioning* code was used.

Cognitive Processing Language. Using the Taxonomy, a total of 60 minutes of mathematics and 60 minutes language arts instruction was coded in 30-second intervals. In any given interval, any code can occur no more than once. This enables a characterization of teachers' instructional language that reflects the percentage of intervals in which any of the 26 codes or types of instructional language occurred. In previous work, Coffman et al. (2008)

identified a subset of codes that characterize the extent to which teachers incorporate Cognitive Processing Language, or language that encourages deep levels of processing and metacognitive awareness during instruction. As can be seen in Table 3, two individual codes and three composite codes are combined to yield an index of CPL: (1) *strategy suggestions*, (2) *metacognitive questioning*, (3) the co-occurrence of *deliberate memory demands* and *instructional activities*, (4) the co-occurrence of *deliberate memory demands* and *cognitive structuring activities*, and (5) the co-occurrence of *deliberate memory demands* and *metacognitive information.* Traditionally, CPL scores have been calculated by standardizing the percent of intervals in which each code or composite code occurred according to its mean and standard deviation. The resulting *T* scores for each of the five components of CPL are then averaged to generate an overall score. Importantly, standardizing each code impedes the examination of mean-level differences. Accordingly, a composite CPL score was calculated that is conceptually consistent with the one presented by Coffman et al. (2008) but avoids standardization, which gives equal weight to each of the components. Given that each individual component of CPL could occur no more than once per interval, there are a total of five possible opportunities to use CPL in any given 30-second interval, resulting in 600 possible instances of CPL across 120 intervals of mathematics or language arts instruction (or 1200 instances of 240 intervals of mathematics and language arts instruction combined). The total number of observed instances of CPL was divided by the number of total possible instances, resulting in a percentage of the possible instances of CPL that were actually observed.[3]. For an example of instruction that is rich in CPL during mathematics and language arts instruction, see Table 4.

[3] This is mathematically equivalent to taking the average number of intervals that contain a CPL code.

Eight videotaped lessons totaling 59 minutes of instruction were used as training files to familiarize coders with the use of the *Taxonomy*[4]. The first two files were considered practice files, and each coder was required to reach 100% reliability with the master file. On the remaining six training files coders were required to attain a criterion of at least 80% reliability with the master file. Following this certification process, reliable coders began coding independently. To ensure reliability, 25% of intervals and were randomly selected to be compared with the files of a master coder. Inter-rater reliability scores for Cohort 1 using the *Taxonomy* ranged from 80 to 96%, with an average of 87%; from 80% to 100%, with an average of 85% for Cohort 2; from 80% to 100%, with an average of 85% for Cohort 3, and from 80% to 95%, with an average of 85% for Cohort 4.

Child assessments. Child assessments were conducted twice, once at the beginning of the first-grade year (Time 1) and then again at the end of the year (Time 2). Each assessment took place at the children's school and lasted between 45 and 60 minutes per session. An experienced and trained research assistant worked one-on-one with each participant to administer a battery of tasks designed to measure a range of cognitive and academic skills. In this investigation, the focus is on children's strategic behaviors on both memory and mathematics tasks. All procedures were videotaped for subsequent analysis[5].

Object Memory Task (Baker-Ward et al.,1984). This task was designed to assess children's deliberate techniques for remembering, including behavioral and linguistic strategies

[4] In Cohort 1 classroom observations were live coded. Prior to the start of data collection each observer had extensive exposure to videotapes of teacher-led instruction and independently coded 50 30-second intervals of instruction using the *Taxonomy*. To assess reliability, observers' files were compared to those of the master coder and measured in terms of percent agreement. Reliability scores ranged from 80 to 96% with an average of 87%.
[5] In Cohort 1 child assessments were conducted four times throughout the year, and in Cohort 2 child assessments were conducted three times throughout the year. In Cohorts 3 and 4, assessments were conducted twice, once in the fall of the year and once in the spring. In order to be consistent across cohorts, two time points, the fall and the spring assessments, were selected for analyses across all cohorts.

displayed while attempting to remember a set of stimulus objects. Multiple sets of 15 familiar and colorful objects were created and counterbalanced. The sets contained a range of items such as plastic toy animals, vehicles, and household items (e.g., gardening tools and dishes). All sets were designed so that the items were unrelated and were not able to be easily grouped by color, function, or semantic category. At each assessment, the children were given a 2-minute study period and told that they should “work to remember these things.” At the conclusion of the study period, the examiner hid the objects under a cloth and asked for recall. Recall measures were obtained by documenting how many items the child was able to remember. Children’s task-oriented or helpful strategic behaviors, such as visual examination, pointing, and naming, were videotaped for subsequent coding and analysis. For a complete list and definition of strategic behaviors see Table 1. The proportion of time that children spend engaged in task-oriented strategic behaviors was calculated to provide an index of children’s strategy use. Two coders independently scored 25% of the records. Kappas for Cohort 1 ranged from .72 to .98, with an average of .84; scores for Cohort 2 ranged from .72 to .94, with an average of .80, whereas those for Cohorts 3 and 4 ranged from .70 to .95, with an average of .80, and from to .72 to .95, with an average of .84, respectively. This task was administered in the fall and spring in Cohorts 1, 2, 3, and 4.

Free Recall with Organizational Training Task (Moely et al., 1992). A free recall task with a training component utilizing semantically-related items allowed for the assessment of organization strategies at both input (e.g., sorting or grouping) and output (e.g., category clustering). At each trial, children were presented with 16 cards with easy-to-label line drawings drawn from four closely related taxonomic categories (e.g., fruits, animals, transportation groups).

Items were presented one at a time in a random order, and then left visible so that the child could prepare for recall. When the child signaled that he or she was ready, the cards were removed and recall was assessed. At the first assessment point in the fall of grade 1 each child received three trials – baseline, training, and generalization – whereas in the spring of grade 1 only non-instructed generalization trials were administered. In the baseline trial, children were told they could pick the cards up, move them around, or do anything else they wanted to work to remember the cards but no explicit instructions for strategy use were given so that children's spontaneous use of organizational techniques could be observed. In the training trial, the experimenter grouped the cards for the children and told them that if they put the cards into groups (e.g., fruit, transportation categories) and try to remember the groups, that they will remember the cards better. The generalization trial was with a new set of stimulus cards and contained no strategy instruction. By contrasting children's spontaneous organization on the baseline trial with that observed on the generalization trial after explicit mnemonic training it is possible to examine children's application of strategy training to a new group of stimuli without any further instruction, allowing for the measurement of the continued use of strategies in the absence of specific instructions to do so.

In addition to measures of recall, an indicator of children's strategy use was also obtained. A standard index of categorical grouping, the Adjusted Ratio of Clustering (ARC) score (Roenker, Thompson, & Brown, 1971) was calculated to characterize the extent to which children in engaged in organizational sorting, or categorical grouping at input. The ARC scores ranged from -1 (below chance organization) to 0 (chance) to 1 (complete organization). Two coders independently scored all records, with any discrepancies being resolved through examination of the original videotape. This task was administered in the fall and spring in Cohorts 1, 2, 3, and 4.

Math Problem Solving Task (Siegler & Jenkins, 1989). This arithmetic task was administered in order to characterize children's addition skills. An experimenter read 10 single-digit addition problems aloud and instructed the child to solve each problem using any method they preferred. Immediately after the child provided a response, the examiner asked for a retrospective report of how the problem had been solved. Both the answer and the strategy reported by the child were recorded. Children's performance on each trial resulted in two measures, one for accuracy, defined in terms of the number of problems answered correctly, and the other for the effectiveness of strategy use, defined as the percentage of the correctly solved problems that involved the use of a strategy. A complete list of definitions and examples for each strategy are provided in Table 5. Two coders independently scored 25% of all records based on the behaviors that could be observed and the children's self-reported strategy nominations. This task was administered in the fall and spring in Cohorts 2, 3, and 4. It was not administered in Cohort 1.

Head Toes Knees Shoulders (Ponitz, McClelland, Jewkes, Connor, Farris & Morrison, 2008). Children's behavioral regulation was assessed using the Head Toes Knees Shoulders (HTKS) task. The task involves 20 test trials in which children were required to use four rules to respond to the experimenter's verbal commands. The participants were given a command (e.g., "touch your head" or "touch your knees") and instructed to perform the opposite action (e.g., touch their toes or touch their shoulders, respectively). Two points were awarded for a correct response and one point was given if the child initially moved toward to the wrong body part, but then self-corrected. Incorrect responses received zero points. Scores were calculated based on children's performance on the 20 test trials. Possible scores thus range from 0 to 40 and represent children's abilities to remember instructions and inhibit the dominant response in order to respond correctly. Two coders independently scored all records, with any discrepancies being resolved through the examination of the original videos. This task was administered once during the year in Cohorts 2, 3, and 4. It was not administered in Cohort 1.

Digit Span (Jacobs, 1887; McCarthy, 1972). The digit span is an assessment of children's short-term and working memory skills and is used as a measure of basic memory capacity. Following standardized administration procedures, experimenters verbally presented strings of numbers of increasing length to the children. On the forward version of the task, children were instructed to repeat the presented string of numbers, whereas on the backward version they were instructed to repeat the string of numbers in backward order. When a child responded incorrectly, another string of the same length was presented, and after two incorrect responses within the same span length, the task was discontinued. A child's performance was characterized by two scores, a forward and a backward span, each of which reflects the longest string of digits that could be recalled correctly in the specified direction. Two coders independently scored all records with any discrepancies being resolved through the examination of the original videos. This task was administered once during the first-grade year in all cohorts.

RESULTS

In the sections that follow, I first present the children's performance on the memory and mathematics tasks across the first-grade year in a descriptive fashion[6]. A list of each of the tasks with corresponding measures and definitions is provided in Table 6. I then describe variation in the types of language that teachers use during mathematics and language arts instruction and characterize the classroom context in terms of teachers' use of *Cognitive Processing Language* (CPL)[7]. In the final section, hierarchical linear models are used to examine the hypothesized association between CPL, self-regulation and children's memory and mathematical skills at the end of first-grade. It was expected that children who were exposed to greater amounts of CPL across the year would demonstrate greater strategy use and increased performance on both memory and mathematics at the end of the year, compared to their peers who experienced less CPL. Furthermore, an interaction between self-regulation and CPL was expected to be observed, such that those children who scored lower on measures of behavioral regulation were expected to benefit the most from instruction that was rich in CPL.

Aim 1: Describing Children's Memory and Math Performance

The overall sample means are presented in Table 7 to provide an initial overview of the children's performance on each of the tasks in the fall and spring of the first-grade year. To

---

[6] Note that the sample size for each outcome measure varies slightly. Not all tasks were administered in each of the four cohorts. In addition, some data are missing due to lost videos or because children did not complete particular tasks.

[7] Note that 60 minutes of mathematics instruction were collected for the full sample of 60 teachers but 60 minutes of language arts instruction was only collected for 57 teachers. Accordingly, the combined 120 minutes of mathematics and language arts instruction is based on the sample of 57 teachers.

provide additional descriptive information about the distribution of performance, a series of box plots are presented in Figures 1 to 16. As can be seen in the box plots presented in Figure 1, there were only modest increases in children's recall on the Object Memory Task in the spring (*M* = 8.13, *SD* = 2.07) compared to the fall (*M* = 7.90, *SD* = 2.06). In order to characterize the strategies children used while trying to remember the objects, the amount of time spent engaged in task-oriented strategic behaviors (e.g., pointing, naming) during the two-minute study period was coded (a complete list of the specific task-oriented strategic behaviors are displayed in Table 1). As shown in Figure 2, the amount of time that children spent engaged in helpful behaviors was relatively the same in the fall (*M* = 115.94, *SD* = 28.29) and spring (*M* = 118.03, *SD* = 27.47).

Additional measures of children's deliberate memory skills were obtained from the Free Recall with Organizational Training Task. In the fall of first grade, children completed a baseline trial, a training trial and a generalization trial following a 15-minute delay. As displayed in Figure 3, children recalled a similar number of cards on the fall baseline (*M* = 8.96, *SD* = 2.54) and generalization trials (*M* = 9.44, *SD* = 3.42) in the fall. Recall improved slightly throughout the year as children remembered 10.49 cards (*SD* = 3.26) on the generalization trial administered in the spring. A standard index of categorical grouping, the adjusted ratio of clustering (ARC) score (Roenker et al., 1971), was calculated to characterize children's sorting during the study period. ARC scores range from -1 (below chance organization) to 0 (chance) to 1 (complete categorization). As can be seen in Figure 4, the children's use of an organizational sorting strategy improved following training, as evidenced by an increase in children's sorting ARC scores from the baseline (*M* = -.07, *SD* = .40) trial to the generalization trial in the fall (*M* = .43, *SD* = .60) and in the spring, (*M* = .47, *SD* = .58)

In addition to measures of children's memory skills, a single-digit addition task was administered to assess children's accuracy and strategy use in the context of arithmetic. As can be seen in Figure 5, there were modest increases in children's accuracy across the year. Overall, the children answered more problems correctly at the end of the first-grade year (*M* = 8.45, *SD* = 2.13) than at the beginning (*M* = 7.57, *SD* = 2.65). An individual strategy code (e.g., sum, min, decomposition) was assigned for each of the ten problems included in the addition task. In order to assess effective strategy use, a composite variable was created that represents the percentage of times that any strategy was used to produce a correct answer. As shown in Figure 6, at the end of the first-grade children used strategies to solve problems correctly (*M* = 85%, *SD* = 23%) more often than they did at the beginning of the year. (*M* = 80%, *SD* = 25%).

Also displayed in Table 7 are the results of the backward version of the Digit Span Task, an indicator of children's working memory skills, and the Head Toes Knees Shoulders task, a measure of children's behavioral self-regulation skills, both administered in the fall. On average, children were able to recall approximately three numbers in backwards order (*M* = 3.20, *SD* = .88) and had an average score of 33.34 (range = 0 - 40, *SD* = 5.69) on the HTKS. In sum, these findings provide support for the hypothesis that children's performance, in terms of recall, accuracy, and strategy use, would improve over the course of the year.

Aim 2: Examining Associations in Performance within and across Tasks

Correlations between strategy use and performance. In both the fall and spring, significant correlations were observed between children's strategy use and performance on each of the three outcome measures. As can be seen in Table 8, the strength of the association between children's strategy use and their recall on the OBJ task remained relatively the same from the fall ($r = .21$, $p < .01$) to the spring ($r = .24$, $p < .01$). On the FRT task, however, strategy use and recall were more strongly correlated in the spring ($r = .62$, $p <$ 05) compared to the baseline ($r = .43$, $p < .05$) and generalization ($r = .52$, $p < .05$) trials in the fall, as can be seen in Table 9. The opposite pattern was observed on the mathematics task, such that children's strategy use and accuracy were more strongly correlated in the fall ($r = .69$, $p < .05$) than they were in the spring ($r = .57$, $p < .05$), as can be seen in Table 10. It was hypothesized that strategy use and performance (e.g., recall and accuracy) would be more strongly correlated in the spring than the fall as strategy use should be become less cognitively demanding over time. The data reported here provide partial support for this hypothesis but only in the context of the FRT task.

Correlations between recall in memory and accuracy in mathematics within domain. Correlations in performance across the two memory tasks were also examined. In the fall of first grade, significant associations were observed between children's recall on both memory tasks. As can be seen in Table 11, children's recall during the OBJ task was moderately correlated with recall on the FRT task at both the baseline ($r = .26$, $p < .05$) and generalization ($r = .25$, $p < .05$) trials. Within-task correlations were also observed in the fall, such that children's recall on the baseline and generalization trials of the free recall with training task was moderately correlated ($r = .31$, $p < .05$). In the spring, the correlation between children's recall on OBJ and FRT was stronger ($r = .40$, $p < .05$) than it was in the fall.

In addition to concurrent associations within memory tasks, longitudinal associations were also observed between children's initial recall and their recall at the end of the year. Recall on OBJ in the fall was moderately associated with recall during OBJ ($r = .37$, $p < .05$) and FRT ($r = .28$, $p < .05$) in the spring. Similarly, children's recall on the baseline trial of FRT in the fall was associated with OBJ recall ($r = .29$, $p < .05$) and their recall on (the generalization trial of the) FRT ($r = .32$, $p < .05$) in the spring. Children's recall on the generalization trial of FRT in the fall was also associated with recall during OBJ (r = .32, p < .05) and FRT (r = .48, p < .05) in the spring. Finally, children's accuracy on the mathematics task in the fall was strongly correlated with accuracy in the spring (r = .75, p < .05).

Correlations between recall in memory and accuracy in mathematics across domains. At the beginning of first grade, children's recall on the baseline (r = .12, p < .05) and generalization (r = .24, p < .05) trials of FRT was weakly but significantly correlated with accuracy in mathematics. Recall during OBJ and accuracy on the mathematics task were not associated. A similar pattern was observed at the end of the year, such that in the spring children's accuracy in mathematics was correlated with recall during FRT (r = .21, p < .05), but not OBJ. Longitudinal associations between recall and accuracy were also observed. Children's recall on the generalization trial of FRT in fall was associated with their accuracy on the mathematics task in the spring OBJ (r = .19, p < .05). Taken together, these findings provide general support for the hypothesis that correlations between recall would be more strongly correlated both within task and at the end of the year compared to the beginning.

Correlations between strategy use within domain. Significant correlations between children's strategy use during OBJ and FRT were not observed in the fall. A weak but positive association (r = .10, p < .05) was observed between the amount of time children spent engaged in helpful behaviors during OBJ and their sorting ARC scores on the FRT task in the spring. Concurrent associations in strategy use during FRT were observed, as can be seen in Table 12. Children's sorting ARC scores on the baseline and generalization trials were weakly but positively correlated (r = .27, p < .05).

In addition to concurrent associations between memory tasks, longitudinal associations between children's strategy use between tasks were also observed. To illustrate, the amount of time that children spent engaged in task-oriented strategic behaviors during OBJ in the fall was moderately associated with the duration of time they spent engaged in strategic behaviors during OBJ in the spring ($r = .41$, p < .05). In addition, children's sorting ARC scores on the baseline trial of FRT in the fall were significantly associated with sorting ARC scores ($r = .28$, $p < .05$) in the spring. Finally, children's strategy use in mathematics was in the fall was significantly correlated with their strategy use in mathematics in the spring ($r = .23$, $p < .05$).

Correlations between strategy use across domains. Fewer associations in strategy use were observed across memory and mathematics tasks. In the fall, children's strategy use on the mathematics task was significantly associated with their sorting ARC scores on the generalization trial of FRT ($r = .18$, $p < .05$). In the spring there were no concurrent associations in children's strategy use across domains. However, children's sorting ARC scores on the generalization trial of FRT in the fall continued to be correlated with their strategy use on the mathematics task in the spring ($r = .15$, $p < .05$). These findings provide additional support for the hypothesis that children's performance would be more strongly associated within domain and over time.

Aim 3: Characterizing the Classroom Context

Teacher-led whole-group instruction was observed during regularly scheduled mathematics and language arts lessons in each of the 60 first-grade classrooms. The data reported in Table 13 reflect the variety of types of instructional language captured by the *Taxonomy of Teacher Behaviors* during 60 minutes of instruction in each subject area averaged across all classrooms. Each code in the *Taxonomy* could occur no more than once in any 30-second interval. Thus, inspection of the table reveals the percentage of the 120 30-second intervals in which each of the codes was assigned.

For example, as can be seen in the far-right column of Table 13, across 120 minutes of mathematics and language arts lessons, the full sample of teachers spent a majority of time engaged in instructional activities, with 82.65% of intervals containing some form of instructional language. More specifically, teachers frequently provided general information (64.39%) and specific task information (35.45%). In addition, teachers also devoted a considerable amount of time to cognitive structuring activities, as 45.00% of intervals contained a cognitive structuring code. Indeed, teachers emphasized student attention (20.18%) and focused on relating current classroom activities to previous experiences (11.07%). Memory requests were also frequent, occurring in 55.59% of intervals. Teachers asked students to recall semantic or factual information (52.13%) and to remember academic information in anticipation of future demands (5.58%). In contrast to instructional activities, cognitive structuring activities, and memory requests, teachers provided metacognitive information relatively infrequently (11.15%). The widespread variability in each of the codes across classrooms is also reflected in Table 13.

Cognitive Processing Language. Based on these individual codes, a composite measure of teachers' *Cognitive Processing Language* was developed to reflect the extent to which teachers provided metacognitively-rich language during instruction that promotes the deep processing of information by encouraging students to actively monitor and assess their own thinking and understanding. CPL consists of the five following components: (1) *strategy suggestions*, (2) *metacognitive questioning*, (3) the co-occurrence of *deliberate memory demands* and *instructional activities*, (4) the co-occurrence of *deliberate memory demands* and *cognitive structuring activities*, and (5) the co-occurrence of *deliberate memory demands* and *metacognitive information.* The data reported in Table 14 characterize the provision of CPL for the full sample of first-grade teachers. As can be seen in the far-right column of Table 14, during mathematics and language arts instruction, teachers provided strategy suggestions (6.02%) and metacognitive questions (5.40%) relatively infrequently. Similarly, memory requests and metacognitive information did not co-occur often (6.54%). In contrast, memory requests and cognitive structuring activities (27.38%) and memory requests and instructional activities (47.39%) co-occurred much more frequently.

In previous research (Coffman et al., 2008, Grammer et al., 2016, Hudson et al., 2018) standard scores for each component of CPL were computed in order to combine codes with substantially different rates of occurrence into a composite index. Each of the five codes was standardized according to its mean and standard deviation and the resulting *T* scores were averaged to generate a composite CPL score that could be used to compare across classrooms. Although teachers' provision of CPL is distributed continuously, relatively small sample sizes have prompted researchers to assign teachers with contrasting levels of CPL to one of two groups, either high or low, on the basis of a median split.

In the sections that follow, I take advantage of a larger sample size of first-grade teachers by examining CPL in a continuous fashion. Rather than standardizing each code, which impedes examining mean-level differences, I created a composite score that reflects the extent to which teachers made use of any of the individual component codes during the course of instruction. In any given interval, a teacher could utilize as few as none of the component codes or as many as five, for a total of 600 possible instances of CPL across 120 intervals of mathematics or language arts instruction (or 1200 instances across 240 intervals of mathematics and language arts instruction combined). This method of characterizing the data enables a comparison of CPL between cohorts as well as across subject matter. I begin by considering differences in teachers' instructional language across cohorts and then subject matter by comparing the provision of CPL during mathematics and language arts lessons.

It should be noted that the provision of CPL was not associated with teacher's level of education or years of teaching experience. Indeed, a series of one-way ANOVAs revealed that teachers' use of CPL during mathematics lessons $F(1, 53) = .05$, $p = .82$, language arts lessons $F(1, 51) = .11$, $p = .74$, and mathematics and language arts lessons combined $F(1, 51) = .03$, $p = .87$, did not vary as a function of whether teachers had a Bachelors' degree or a Masters' degree. Similarly, teachers' use of CPL during mathematics $F(3, 51) = .58$, $p = .63$, language arts lessons $F(3, 49) = .55$, $p = .65$, and mathematics and language arts lessons combined $F(3, 49) = .45$, $p = .72$, was not associated with years of teaching experience[8].

Aim 4: Examining CPL by Cohort

In order to consider potential cohort differences in the provision of CPL, as well as the provision of the individual components of CPL, during a combined two-hours of mathematics and language arts instruction, a set of boxplots was created. The average amount of CPL during mathematics and language arts instruction for each cohort is displayed in Figure 7. As can be seen, teachers in Cohort 1 provided relatively less CPL across mathematics and language arts lessons than teachers in Cohorts 2, 3 and 4.

Further examination of the individual components of CPL indicated that this variation was primarily driven by differences in the co-occurrence of memory requests and instructional activities. Inspection of Figure 8 and Table 15 indicates that teachers in Cohort 1 (M = 37.68%, SD = 8.22%) and Cohort 3 (M = 37.95%, SD = 6.46%) paired memory requests and instructional activities less often than teachers in Cohorts 2 (M = 59.44%, SD = 8.23%) and 4 (M = 50.80%,

[8] Teachers were assigned to one of four groups based on number of years of teaching experience: (1) <5 years, (2) 6-10 years, (3) 11-15 years, (4) 16+ years.

SD = 11.14%). Moreover, memory requests and instructional activities were paired more often in Cohort 2 (*M* = 59.44%, *SD* = 8.23%) than in Cohort 4 (*M* = 37.95%, *SD* = 6.46%). There were no observable differences in the provision of strategy suggestions, metacognitive questions, memory requests and cognitive structuring activities, or memory requests and metacognitive information during two hours of combined mathematics and language arts instruction between cohorts.

CPL during mathematics instruction across cohorts. Additional boxplots were examined to consider the provision of CPL during mathematics lessons across cohorts. Inspection of Figure 9 and Table 16 indicate that the overall provision of CPL during mathematics instruction appeared to increase over time. Further examination of the individual components of CPL indicated that this variation was driven by differences in the co-occurrence of (1) memory requests and instructional activities and (2) memory requests and cognitive structuring activities. Inspection of Figure 10 and Table 17, teachers in Cohort 1 (*M* = 40.18%, *SD* = 9.02%) paired memory requests and instructional activities less often than teachers in Cohorts 2 (*M* = 57.65%, *SD* = 9.11%), 3 (*M* = 57.06%, *SD* = 7.88%), and 4 (*M* = 58.57%, *SD* = 10.59%). In addition, as shown in Figure 11 memory requests and cognitive structuring activities co-occurred less often in Cohort 1 (*M* = 23.21%, *SD* = 8.87%) than in Cohort 4 (*M* = 34.76%, *SD* = 10.99%).

CPL during language arts instruction across cohorts. Finally, a series of boxplots was examined to consider differences in the provision of CPL during language arts instruction. As shown in Figure 12, variation in CPL during language arts instruction was observed across cohorts but difference did not appear to occur in a systematic fashion. Further examination of the individual components of CPL indicated that this variation was a result of differences across a variety of individual codes. As can be seen in Figure 13, teachers in Cohort 2 (*M* = 61.32%, *SD* = 10.92%) paired memory requests and instructional activities more often than teachers in Cohort 1 (*M* = 35.18%, *SD* = 8.34%), 3 (*M* = 43.96%, *SD* = 16.96%) and 4 (*M* = 46.28%, *SD* = 11.84%).

Additional differences were observed between teachers in Cohorts 3 and 4 in the provision of metacognitive questions. As shown in Figure 14, teachers in Cohort 3 (*M* = 5.48%, *SD* = 4.42%) asked more metacognitive questions during language arts instruction than teachers in Cohort 4 (*M* = 2.28%, *SD* = 2.24%). Teachers in Cohort 3 (*M* = 6.39%, *SD* = 4.51%) also paired memory requests and metacognitive information more often than teachers in Cohorts 1 (*M* = 2.86%, *SD* = 3.64%) and 4 (*M* = 2.76%, *SD* = 1.94%), which is presented in Figure 15.

Taken together these data provide descriptive evidence to support the hypothesis that overall levels of CPL would increase across the cohorts in the study (i.e., over time). However, these data do not provide consistent support for the hypothesis that increases in CPL would be due to the more frequent provision of cognitive structuring activities and metacognitive language.

Aim 5: Comparing CPL in Mathematics and Language Arts

Correlations were calculated to compare teachers' use of CPL during mathematics and language arts lessons. Evidence was found to support the hypothesis that CPL would be correlated across subject matter but nonetheless consistently higher in mathematics in comparison with language arts. Indeed, teachers' provision of CPL during mathematics instruction was moderately associated with CPL during language arts instruction ($r = .34$, $p < .05$). However, as shown in Figure 16, the teachers' overall CPL score was higher in mathematics (*M* = 21.17%, *SD* = 5.28%) than language arts (*M* = 17.09%, *SD* = 5.19%), indicating that CPL occurs more frequently during the course of mathematics instruction than in language arts. Further examination of the individual components of CPL indicates that with the exception of memory requests paired with cognitive structuring activities, teachers tended to use each component of CPL more often during mathematics lessons than language arts lessons. On average, the teachers provided more strategy suggestions during mathematics lessons (*M* = 7.59%, *SD* = 5.86%) than language arts lessons (*M* = 4.45%, *SD* = 3.91%). Similarly, teachers asked metacognitive questions more often in mathematics (*M* = 7.66%, *SD* = 4.56%) than language arts (*M* = 3.13%, *SD* = 3.28%). Teachers also paired memory requests and instructional activities together more frequently during mathematics lessons (*M* = 53.07%, *SD* = 11.86%) compared to language arts lesson (*M* = 47.57%, *SD* = 15.49%). Memory requests and metacognitive information also occurred more frequently during mathematics (*M* = 8.67%, *SD* = 5.54%) in language arts (*M* = 4.39%, *SD* = 3.71%).

Aim 6: Linking CPL to Children's Memory and Mathematics Performance

A series of hierarchical linear models (HLMs) (Bryk & Raudenbush, 1992; Bryk & Raudenbush, 1987) was used to investigate the relation between teachers' CPL and first-graders memory and mathematics performance. Prior to fitting HLMs, the data were examined for univariate normality, and graphs for each variable of interest are presented in Appendix A. For each outcome of interest, the first model analyzed was an unconditional model to determine whether variability in the outcome variable was significantly different than zero and to evaluate the within- versus between- classroom differences in student performance at the end of first-grade. A random intercept regression model with level-1 predictors was fit next in order to adjust for children's working memory skills, self-regulation skills, and initial performance at the beginning of first-grade[9]. Children's working memory scores, self-regulation scores, and

[9] Given the variation observed in CPL across cohorts, cohort was also included as a covariate but was not a significant predictor of performance for any outcome variable.

performance at Time 1 were group mean centered. This set of covariates was chosen because each of these child-level variables are theoretically believed to be associated with children's end-of-year performance in memory and mathematics. Given that the emphasis of this investigation involves understanding the nature of the current data, rather than making inferences to a broader population, level-1 predictors were included as fixed effects in each model[10]. Following this, teachers' CPL was included as a level-2 predictor. Finally, a slopes-as-outcomes model was fit to examine the cross-level interaction between children's self-regulation and teachers' CPL on students' performance at the end of first grade. As will be seen in the following sections, the results of the HLMs do not indicate that teachers' CPL was associated with student performance on any of the outcome variables of interest. As such, additional descriptive analyses are subsequently presented in an effort to contextualize the pattern of results that was observed.

Recall during the Object Memory Task. A Random Effects ANOVA model was fit to formally evaluate the within- versus between-classroom differences in children's recall on the OBJ task at the end of first grade. As can be seen in Table 18, the model implied estimate of the number of items recalled at the end of first grade was 8.11 and is significantly different from zero. The variance component at level 1, which represents within-group variance in recall, is 3.86 and is statistically significant. The variance component at level 2, which represents between-group variance in recall, is .44 and is also statistically significant. Using these variance estimates, the Intraclass Correlation Coefficient (ICC) was calculated. The ICC measures the degree of dependence in the data or the strength of the nesting effect. The resulting ICC for the data was .10, which indicates that approximately 10% of the variability in recall is due to

[10]The models were also fit with level-1 predictors included as random effects but failed to converge, even after standardizing each predictor. This is likely due to the fact that the intraclass correlation for each outcome of interest was considerably small.

between-classroom differences. An alternative interpretation of the ICC is that, on average, students' recall scores are correlated 10% within any given classroom. The equations for this model are reported in Appendix B.

Following this, a conditional model with group mean centered level-1 predictors was estimated in order to examine the relationship between children's recall at the end of the year and child-level variables, including children's working memory skills, self-regulation, and recall at the beginning of the year. The only significant predictor of the number of items recalled at the end of first grade was children's recall at the beginning of the year, $F(1, 251) = 38.38$, $p < .01$. Children's working memory and self-regulation skills were not significantly associated with their end-of-the-year recall on OBJ. The model-implied estimate for the number of items recalled at the end of first-grade now reflects the predicted score for a student who scored at his or her classroom average for each predictor. Based on this model, for every one item increase in recall at the beginning of the year, there is a corresponding increase of .34 items at the end of the year. The results of this model are displayed in Table 19 and the corresponding equations are reported in Appendix C.

Subsequently, teachers' CPL was included as a level-2 predictor. The previously observed effect of initial recall remained significant. The fixed effect of children's self-regulation skills is .04 and is marginally significant, $F(1, 236) = 3.83$, $p = .05$. The effect of teachers' use of CPL is not statistically significant. Based on this model, for every item recalled at Time 1 a corresponding .35 increase in items recalled is expected at Time 2. Moreover, for every 1-point increase in a child's self-regulation score, a corresponding .04 increase in items recalled is expected at the end of the year. The results of this model are presented in Table 20 and corresponding equations are reported in Appendix D.

A final model including a cross-level interaction between teachers CPL and children's self-regulation skills was fit. As can be seen in Table 21, the only significant predictor of recall at Time 2 was children's initial recall, $F(1, 236) = 36.38$, $p < .01$. A significant interaction between children's self-regulation skills and teachers' use of CPL was not observed. The equations for this model are presented in Appendix E.

Strategy use during Object Memory Task. A Random Effects ANOVA model was fit to formally evaluate the within- versus between-classroom differences in children's strategy use on the OBJ task at the end of first grade. As can be seen in Table 22, the model implied estimate of the amount of time that children spent engaged in strategic behaviors at the end of first grade was 118.09 seconds and is significantly different from zero. The variance component at level 1, which represents within-group variance in recall, is 729.30 and is statistically significant. The variance component at level 2, which represents between-group variance in recall, is 25.83 and is not statistically significant. Using these variance estimates, the resulting ICC for the data was .03, which indicates that approximately 3% of the variability in recall is due to between-classroom differences. An alternative interpretation of the ICC is that, on average, students' recall scores are correlated 3% within any given classroom. The equations for this model are reported in Appendix F.

Following this, a conditional model with group mean centered level-1 predictors was estimated in order to examine the relationship between children's strategy use at the end of the year and child-level variables, including children's working memory skills, self-regulation, and recall at the beginning of the year. The only significant predictor of the amount of time children spent engaged in strategy use at the end of first grade was the duration of children's strategy use at the beginning of the year, $F(1, 274) = 75.69$, $p < .01$. Children's working memory and self-regulation skills were not significantly associated with their end-of-the-year strategy use on OBJ. The model-implied estimate for the amount of time spent engaged in strategic behaviors at the end of first-grade now reflects the predicted score for a student who scored at his or her classroom average for each predictor. Based on this model, for every one-second increase in strategy use at the beginning of the year, there is a corresponding increase of .35 seconds at the end of the year. The results of this model are reported in Table 23 and the corresponding equations are displayed in Appendix G[11].

Subsequently, teachers' CPL was included as a level-2 predictor. The previously observed effect of initial strategy use remained significant. The fixed effect of children's self-regulation skills is .49 and is trending toward significance, $F(1, 257) = 2.99$, $p = .09$. The effect of teachers' use of CPL is not statistically significant. Based on this model, for every one-second increase in strategy use at Time 1 a corresponding .35-second increase in strategy use is expected at Time 2. The results of this model are shown in Table 24 and the corresponding equations are reported in Appendix H.

A final model including a cross-level interaction between teachers CPL and children's self-regulation skills was fit. As can be seen in Table 25, the only significant predictor of strategy use at Time 2 was children's initial strategy use, $F(1, 256) = 72.17$, $p < .01$. The interaction effect between children's self-regulation skills and teachers' use of CPL is trending towards significance, $F(1,256) = 3.40$, $p = .07$. The equations for this model are reported in Appendix I.

Recall during the Free Recall with Organizational Training Task. The same modeling building steps were followed to examine children's recall on the FRT task. The results

---

[11] It should be noted that the models reported in Appendices G, H, and I allowed for between teacher variance at level-2 but the best estimate was effectively zero.

of the Random Effects ANOVA for recall are displayed in Table 26 and the corresponding equations are reported in Appendix J. The model implied estimate of the number of items recalled at the end of first grade is 10.47 and is significantly different from zero. The within-group variance is 10.39 and the between group variance is .23, which results in an ICC of .02. This indicates that approximately 2% of the variance in recall on FRT is due to between classroom differences.

A conditional model including group mean centered level-1 predictors was estimated in order to examine the relationship between children's recall at the end of the year and children's working memory skills, self-regulation, and initial recall. Children's initial performance at the beginning of the year, $F(1, 259) = 21.99$, $p < .01$ as well as their self-regulation skills, $F(1, 257) = 9.81$, $p < .01$ significantly predicted their recall in the spring. The model-implied estimate for the number of items recalled at the end of first-grade now reflects the predicted score for a student who scored at his or her classroom average for each predictor. Based on this model, for every one item increase in recall at the beginning of the year, there is a corresponding increase of .37 items at the end of the year. In addition, for every one-point increase in performance on the HTKS task, there is a corresponding increase of .11 items recalled at the end of the year. The results of this model displayed in Table 27 and the corresponding equations are reported in Appendix K.

In the following model, teachers' CPL was included as a level-2 predictor. The previously observed effects of initial recall and self-regulation remained significant. The effect of teachers' use of CPL is not statistically significant. The results of this model are shown in Table 28 and the corresponding equations are reported in Appendix L. A final model including a cross-level interaction between teachers CPL and children's self-regulation skills was fit. As can be seen in Table 29, the only significant predictor of recall on the FRT task in the spring was children's recall in the fall. A significant interaction between children's self-regulation skills and teachers' use of CPL was not observed. The equations for this model are shown in Appendix M.

Strategy Use during the Free Recall with Organizational Training Task. The results of the Random Effects ANOVA for children's strategy use during FRT are presented in Table 30 and the corresponding equations are displayed in Appendix N. The model implied estimate of children's sorting ARC scores at the end of first grade was .47 (scores range from -1 to 1 with a score of zero equal to chance) and is significantly different from zero. The within-group variance is .32 and the between group variance is .02, which results in an ICC of .06. This indicates that approximately 6% of the variance in children's organizational sorting on FRT is due to between classroom differences.

In the next step, level-1 predictors were added. Children's initial strategy use, $F(1, 256) = 25.58$, $p < .01$, and self-regulation skills, $F(1, 256) = 4.38$, $p < .05$, were significant predictors of their strategy use in the spring. Children's working memory skills were not significantly associated with their strategy use in the spring during FRT. The model-implied estimate for organizational sorting at the end of first-grade now reflects the predicted score for a student who scored at his or her classroom average for each predictor. Based on this model, for every one-unit increase in sorting ARC scores at the beginning of the year, there is a corresponding .45 increase in sorting at the end of the year. In addition, for every one-point increase in self-regulation, sorting ARC scores are expected to increase by .01. The results of this model are presented in Table 31 and the corresponding equations are reported in Appendix O.

In the following model, teachers' CPL was included as a level-2 predictor. The previously observed effect of initial recall remained significant and self-regulation was marginally significant, $F(1, 241) = 3.54$, $p = .06$. The effect of teachers' use of CPL is not statistically significant. The results of this model are shown in Table 32 and the corresponding equations are reported in Appendix P. A final model including a cross-level interaction between teachers CPL and children's self-regulation skills was fit. As can be seen in Table 33, the only significant predictor of strategy use on the FRT task in the spring was children's strategy use in the fall. A significant interaction between children's self-regulation skills and teachers' use of CPL was not observed. The equations for this model are shown in Appendix Q.

Accuracy during the Mathematics Task. The results of the Random Effects ANOVA for accuracy are displayed in Table 34 and the corresponding equations are reported in Appendix R. The model implied estimate of the number of correct responses provided at the end of first grade is 8.45 and is significantly different from zero. The within-group variance is 4.48 and the between group variance is .04, which results in an ICC of .01. This indicates that approximately 1% of the variance in recall on FRT is due to between classroom differences.

In the next step, a conditional model with level-1 predictors was fit. Children's accuracy in the fall, $F(1, 257) = 157.87$, $p < .01$, working memory skills $F(1, 251) = 3.77$, $p = .05$, and self-regulation $F(1, 252) = 4.81$, $p < .05$, were all significantly associated with children's accuracy in the spring. The model-implied estimate for organizational sorting at the end of first-grade now reflects the predicted score for a student who scored at his or her classroom average for each predictor. Based on this model, for every one-point increase in accuracy at the beginning of the year, there is a corresponding .54 increase in accuracy at the end of the year. In addition, for every one-point increase in working memory skills there is a corresponding .22 increase in accuracy at the end of the year, and for everyone one-point increase in self-regulation, the number of accurate responses in the spring is expected to increase by .04. The results of this model shown in Table 35 and the corresponding equations are reported in Appendix S.

Teachers' provision of CPL was included as a level-2 predictor in the next step. The previously observed effects of initial recall, working memory, and self-regulation remained significant. The effect of teachers' use of CPL is not statistically significant. The results of this model are shown in Table 36 and the corresponding equations are reported in Appendix T. A final model including a cross-level interaction between teachers CPL and children's self-regulation skills was fit. As can be seen in Table 37, the only significant predictor of strategy use on the FRT task in the spring was children's strategy use in the fall. A significant interaction between children's self-regulation skills and teachers' use of CPL was not observed. The equations for this model are displayed in Appendix U.

Strategy Use during the Mathematics Task. The results of the Random Effects ANOVA for strategy use during the mathematics task are presented in Table 38 and the corresponding equations are reported in Appendix V. The model implied estimate of the percentage of effective strategy use at the end of first grade is .85 and is significantly different from zero. The within-group variance is .05 and the between group variance is .00, which results in an ICC of 0. This indicates that none of the variance in strategy use on the mathematics task is due to between classroom differences.

Following this, a conditional model with level-1 predictors was fit. Children's strategy use in the fall $F(1, 248) = 4.87$ $p < .05$ and their working memory skills $F(1, 248) = 7.90$, $p = .01$, were significantly associated with children's strategy use in the spring. Self-regulation was not significantly associated with strategy use. The model-implied estimate for effective strategy use at the end of first-grade now reflects the predicted score for a student who scored at his or her classroom average for each predictor. Based on this model, for every one percentage-point increase in strategy use at the beginning of the year, there is a corresponding .14 increase in strategy use at the end of the year. In addition, for every one-point increase in working memory skills there is a corresponding .05 increase in strategy use at the end of the year. The results of this model are shown in Table 39 and the corresponding equations are reported in Appendix W.

In the next step, teachers' use of CPL was added as a level-2 predictor. The previously observed effects of initial recall and working memory remained significant. The effect of teachers' use of CPL is not statistically significant. The results of this model are displayed in Table 40 and the corresponding equations are reported in Appendix X. A final model including a cross-level interaction between teachers CPL and children's self-regulation skills was fit. As can be seen in Table 41, children's initial performance and working memory skills were the only significant predictors of strategy use on the mathematics task in the spring. A significant interaction between children's self-regulation skills and teachers' use of CPL was not observed. The equations for this model are presented in Appendix Y.

The results of the HLMs generally do not provide support for the hypothesis that student performance would vary at the end of the year in relation to the amount of CPL provided by the teachers. Across each of the different outcome variables the most significant predictor of performance at the end of first-grade was students' performance at the beginning of the year. Partial support was found for the hypothesis that students' self-regulation would be related to their performance, as higher self-regulation scores were associated with (1) greater recall and strategy use during FRT and (2) more accurate responses during the math task. The hypothesis that children's self-regulation skills and teachers' CPL would interact to effect student performance was not supported.

Additional Analyses

Given that previous observational and experimental (Coffman et al., 2008; Coffman, et al., 2017; Hudson et al., 2018; Grammer et al., 2013; Grammer et al., 2016) examinations of teachers' instructional language have consistently demonstrated an effect of CPL on student outcomes, the lack of significant results in this investigation is unexpected. One important distinction between these previously reported findings and the current investigation is the treatment of teachers' provision of CPL as a continuous variable, which enables a comparison of CPL across different cohorts. As reported above in the discussion of findings related to Aim 4, changes were observed over the years in the extent to which teachers provided CPL. More specifically, the overall amount of CPL that teachers provided during mathematics lessons was significantly lower in Cohort 1 than in Cohorts 2, 3, and 4. This observation is important, given that previous studies examining only Cohort 1 (Coffman et al., 2008; Coffman, et al., 2018; Grammer et al., 2016) have demonstrated a significant effect of first-grade teachers' use of CPL on children's (1) strategy use and recall at the end of first grade, (2) mathematics performance in second grade, and (3) study skills in the fourth grade. Importantly, each of these studies classified teachers into two groups, high CPL versus low CPL, on the basis of a median split.

Accordingly, an effort was made to understand how the observed increases in the overall rates of CPL over the years would affect the classification of teachers across the different cohorts as high versus low based on a median split. Inspection of Figure 17 reveals that on average, teachers in Cohort 1 who were classified as *high* provided one or more instances of CPL in approximately 20% of intervals. For comparison, teachers who would be classified as *low* on the basis of a median split in Cohorts 2 (M = 18.46%), 3 (M = 18.64%), and 4 (M = 20.16%) provided approximately the same amount of CPL. In other words, teachers who were classified as high in Cohort 1 are indistinguishable from teachers who were classified as low in Cohorts 2, 3, and 4. Indeed, a two-way ANOVA revealed a significant main effect of Cohort, $F(3, 52) = 13.18$, $p < .01$, and a significant main effect of high versus low classification, $F(1, 52) = 79.83$, $p < .01$.

In addition to notable cohort-level differences in the provision of CPL, variation in child-level performance across cohorts was also observed. To illustrate, consider children's sorting performance during FRT. As can be seen in Figure 18, two distinct patterns emerged across cohorts on the generalization trial at the end of the year. Whereas children in Cohorts 1 and 2 demonstrated increased strategy use over the course of the year, children in Cohorts 3 and 4 demonstrated decreased rates of strategy use as evidenced by lower sorting ARC scores at the end of the year compared to the generalization trial following organizational training in the fall. Importantly, a one-way ANOVA indicated that the cohorts did not differ in the extent to which they acquired the organizational training strategy following training at Time 1 $F(3, 409) = 1.95$, $p = .12$. However, significant differences were observed in the extent to which children maintained the use of the strategy at the end of the year, $F(3, 406) = 11.19$, $p < .01$. A Tukey post-hoc test revealed that sorting ARC scores were significantly lower in Cohorts 3 ($M = .28$, $SD = .58$) and Cohort 4 ($M = .33$, $SD = .59$) compared to Cohorts 1($M = .64$, $SD = .53$) and 2 ($M = .57$, $SD = .56$).

DISCUSSION

The research presented here constitutes the first attempt to examine in an integrative fashion the data from a series of observational studies carried out in first grade classrooms between 2002 and 2016. Each of these studies had been designed to assess children's developing cognitive skills in relation to aspects of the instructional context in which they are embedded. Previous research indicates that children who are exposed to more versus less Cognitive Processing Language (CPL) (based on a median split) differ (1) in the extent to which they spontaneously use strategies and in their response to strategy training in the context of deliberate memory tasks (Coffman et al., 2008) as well as (2) in their effective use of arithmetic strategies to solve single-digit addition problems (Hudson et al., 2018). This investigation was designed to expand on the findings reported by Coffman et al. (2008) and Hudson et al. (2018) in three important ways. First, by utilizing four cohorts of first-grade students and teachers, it was possible to examine teachers' provision of CPL in a continuous fashion rather than dichotomizing teachers into high and low groups. Doing so had the potential to provide a more nuanced and accurate reflection of exposure to CPL during instruction. Second, this study incorporated measures of children's performance across memory and mathematics tasks, which provides a novel opportunity to explore strategy use across domains. Lastly, the inclusion of observational measures of children's self-regulation skills allowed questions about the individual contribution of child-level differences, as well as possible interactive effects between child-level and environmental-level factors to be explored.

In the following sections, the findings will be discussed according to each of the six aims of the study. Subsequently, limitations of the current investigation, including possible measurement issues at both the child- and classroom-level will be addressed in an effort to understand why an association was not observed between teachers use of CPL and student performance. In light of these findings, directions for future research, including the role of metacognition and self-regulated learning, will also be considered.

Children's Performance in Memory and Mathematics

Descriptive analyses of children's performance were carried out to address Aims 1 and 2. Consistent with the hypotheses presented in Aim 1, children's recall, accuracy, and strategy use improved on each of the memory and mathematics tasks over the course of the first grade. Importantly, however, these improvements were minimal and although statistically significant, they may not represent substantively meaningful changes in performance.

Aim 2 was designed to examine the correlation between children's recall, accuracy, and strategy use, both within and across domains, over the course of the year. Children's recall and strategy use during the Object Memory Task (OBJ) were significantly and positively correlated in the fall and spring, such that as indicators of strategy use increased so did the number of items recalled. The strength of the association between strategy use and recall on the OBJ task did not increase over the course of the year, as had been expected. On the Free Recall with Organizational Training (FRT) task, children's recall and strategy use were significantly correlated on each trial during the fall and spring. As hypothesized, the association between strategy use and recall was stronger in the spring than it was in the fall, indicating that strategy use and recall were more closely linked in the spring of first-grade than they were in the fall. Finally, children's strategy use and accuracy on the mathematics task were significantly and positively correlated in both the fall and the spring, but contrary to what was hypothesized, the association was weaker in the spring than in the fall.

Thus, only partial support was found for the hypothesis that children's strategy use would be more strongly associated with performance at the end of the year. It is possible that an increase in the association between strategy use and recall over the year was observed on FRT but not OBJ because the items in the FRT task are highly associated and consequently can be readily organized into taxonomic categories, whereas the items in the OBJ task are low associated and cannot be easily grouped. Thus, strategy use during OBJ may require substantially more effort than strategy use during FRT, thereby limiting the cognitive resources available for encoding and storing the items in memory.

The observed decrease in the strength of the association between strategy use and accuracy on the mathematics task from the fall to the spring is likely due to the fact that the percentage of trials on which children used retrieval increased over the course of the year. Retrieval codes were assigned when children (1) did not appear to "think" or engage in any observable strategic behaviors (i.e. counting), (2) responded immediately (within 4-seconds of the problem being presented) and (3) explained that they knew the answer (e.g., "I knew it" or "That's easy. I already learned it."). Indeed, higher rates of retrieval in the spring (M = 31%) compared to the fall (M = 21%) indicate that children were able to directly retrieve the answers from memory, without having to make use of a strategy to solve the addition problems. As a result of more frequent attempts to retrieve answers, there are fewer chances to observe children's deliberate strategy use. Moreover, it is possible that instances of strategy use were reserved for more difficult problems on which the likelihood of making a computational error or executing the strategy incorrectly was greater, resulting in a lower correlation between strategy use and performance in the spring compared to the fall.

In addition to within-task correlations between recall and strategy use, associations were also observed between performance on the two memory tasks, OBJ and FRT. More specifically, the number of items that children recalled during OBJ was correlated with their recall on the baseline and generalization trials of FRT in the fall, and the strength of this association increased over the course of the year, such that recall across the two memory tasks was more strongly correlated in the spring than in the fall. Despite these associations in recall across the two memory tasks, strategy use within domain was not correlated. As mentioned above, this may be due to in part to the fact that the items presented during the OBJ task were low-associated items and thus did not readily fit into taxonomic groups whereas the items presented during FRT were high-associated items and specifically designed to fit into four semantic categories (e.g., types of clothing, toys, playground equipment). As a result, the relatedness of the items during FRT may have been so strong that the use of an organizational strategy was obvious.

Taken together, these findings provide consistent support for the hypothesis that children's strategy use and performance are associated within-task, but only partial support was found to support the hypothesis that the strength of this association would increase over time. Furthermore, these findings do not indicate that over time children's strategic skill would be generalized across memory tasks. It is worth noting that the FRT, the task on which the association between strategy and recall strengthened over the year, involved a training component in which children were (1) taught how organize items into semantic groups, (2) provided a rationale for doing so (i.e., it will be easier to remember them later), and (3) given an opportunity to practice. Perhaps such scaffolding reduced the cognitive effort required provided the necessary support needed for strategy use to benefit children's recall. By comparison, the use of low-associated items and the lack of strategy training during the OBJ task may have placed too great a burden on children's limited cognitive resources.

Importantly, however, two key findings emerged among the cross-domain associations between children's recall and strategy use during FRT and their accuracy and strategy use on the mathematics tasks. In the fall children's accuracy on the math task was significantly correlated with the number of items recalled on the baseline trial in the fall and the generalization trials of FRT and the strength of this association was comparable in the spring. Similarly, in the fall children's strategy use during the generalization trial during FRT was associated with their strategy use on the math task and similar concurrent associations between strategy use during FRT and the math task were observed in the spring. Moreover, longitudinal associations emerged between children's strategy use during FRT in the fall and their strategy use on the math task in the spring. That is, children's strategy use on the generalization trial of FRT during the fall remained significantly associated with their strategy use on the mathematics task in the spring. The association between children's strategy use on the generalization trial during FRT and strategy use in the context of the mathematics task in both the fall and spring suggests that the extent to which children are able to generalize the organizational training provided on the training trial of FRT may be an important factor underlying the integration of strategy use across domains over time.

The Classroom Context and Teachers' Provision of Cognitive Processing Language (CPL).

In addition to examining differences in child-level performance, Aim 3 called for descriptive analyses of the instructional context of the first-grade classroom. The results reported here are consistent with previous findings (Coffman et al., 2008; Hudson et al., 2018[12]), such that during both mathematics and language arts lessons teachers spent a majority of instructional time providing general information and task specific instruction to students. Teachers also spent a considerable amount of time asking students to recall previously learned information, and also emphasizing the importance of remembering current information for future assessments. To a lesser extent, the teachers provided attention regulation cues and asked students to make connections between current topics and previously learned material. Importantly, teachers engaged in metacognitively-rich instruction comparatively less often during the course of instruction. Overall, these findings suggest that the instructional context of the first-grade classroom predominantly consists of teacher-led provision of new information and solicitation of previously learned information. By comparison, less emphasis is placed on higher-order cognitive skills such as asking students to explain their problem-solving strategies or to provide rationales for their strategy choice.

CPL across Cohorts and Subject Matter

In addition to broadly describing the instructional context of first-grade classrooms, teachers' use of CPL was also examined. In particular, variation in the provision of CPL across cohorts (Aim 4) and subject matter (Aim 5) was considered. Overall rates of CPL were expected to increase over time due to the implementation of Common Core Standards (NGABCP, 2010). The North Carolina State Board of Education voted to adopt the Standards in June of 2010 and based on this information, increases in CPL were expected to be seen in as early as Cohort 2 (2010-2011). Although the Common Core does not specify any particular curriculum or pedagogy, an emphasis is placed on mathematical understanding and higher-ordering thinking

[12] The findings reported by Coffman et al., 2018 were based on Cohort 1 exclusively and the findings reported by Hudson et al., 2018 were based on a sample of kindergarten students and teachers.

skills, as opposed to procedural skill. As such, increases in the provision of metacognitive and cognitive structuring codes, which provide rich opportunities for children to make connections, and explain and justify their problem-solving approaches were expected to lead to higher rates of overall CPL.

Different patterns in the provision of CPL were observed in mathematics and language arts instruction between the 2002-2003 and 2015-2016 school years. In mathematics, the provision of CPL appeared to increase in a linear fashion as hypothesized, but the only cohort that was significantly lower in CPL than any other cohort was Cohort 1. That is, although CPL during mathematics instruction continued to increase over the years there was not a significant increase between Cohorts 2, 3, and 4. This supports the hypothesis that CPL would increase after the implementation of Common Core, at least in the context of mathematics. As expected, the increase in CPL over time was partially driven by the frequency with which teachers paired memory requests and cognitive structuring activities. However, teachers did not demonstrate greater use of metacognitive language over time.

A different pattern emerged during language arts instruction, such that teachers in Cohort 2 provided significantly more CPL than teachers in Cohorts 1 and 4, but did not differ from Cohort 3. These findings do not provide evidence to suggest that consistent gains in the provision of CPL during language arts instruction occurred following the implementation of Common Core. Moreover, contrary to the hypothesis that changes in CPL would be result from increases in metacognitive language, the variation observed in CPL during language arts instruction was due to significantly more frequent memory requests and instructional activities in Cohort 1.

In addition to examining CPL across cohorts, comparisons were also made across subject matter (Aim 5) As expected, teachers' provision of CPL was moderately correlated across subject matter, but consistently higher during mathematics instruction compared to language arts.

Linking CPL to Children's Performance

One of the primary aims of this study was to examine the changes in students' strategy use and performance in relation to the instructional context of their first-grade classroom. In doing so, a continuous measure of teachers' instructional language was developed based on the raw scores for each component of CPL. Raw scores were used, as opposed to standardized scores which have been used in previous studies (Coffman et al., 2008; Grammer et al., 2016; Hudson et al., 2018), in order to be able to compare CPL across cohorts. Due to the lack of significant findings reported here, it remains to be seen whether the continuous measure that was developed is a reasonable alternative to the traditional median split.

Based on previous findings from both observational and experimental work (Coffman et al., 2008; Grammer et al., 2016, Grammer, Coffman, & Ornstein, 2013; Hudson et al., 2018), children were expected to demonstrate greater strategic skill in the domains of memory and mathematics when exposed to higher levels of CPL. However, evidence to support this hypothesis was not observed for any of the outcome measures included in this investigation. As a result, additional descriptive analyses of teachers' use of CPL and children's performance were carried out in order to examine cohort-level differences that may have contributed to the unexpected pattern of results.

It is important to note that these preliminary analyses were presented in order to provide additional descriptive information about variation in both teacher- and child-level measures across each of the individual cohorts. Further examination of the differences that exist across cohorts provide insight into the lack of significant findings relating children's performance to aspects of the instructional context in which they are embedded. At the classroom level, the provision of CPL increased between the 2002-2003 and 2015-2016 school years in such a way that teachers in Cohorts 2, 3, and 4 who would be classified as low in the provision of CPL on the basis of a median split were indistinguishable from teaches in Cohort 1 who would be classified as high. This distinction is critical because previous analyses using Cohort 1 data have consistently demonstrated an effect of teacher CPL on a range of student outcomes in first, second, and fourth grade (Coffman et al., 2008; Grammer et al., 2016; Coffman et al., 2018). One possible interpretation of this observation is that a certain amount of CPL is required for student performance to increase. If teachers in Cohorts 2, 3, and 4 exceeded this threshold then presumably all students would stand to benefit, making it difficult to detect a difference in performance in relation to exposure to CPL.

In addition to classroom-level differences across cohorts, differences in child-level performance were also observed that may have further contributed to the lack of significant findings. Indeed, two distinct patterns in strategy use during the FRT task emerged across cohorts. In Cohorts 1 and 2, children demonstrated continued growth as evidenced by higher sorting ARC scores on the generalization trial in the spring compared to the baseline and generalization trials in the fall. Children in Cohorts 3 and 4, however, demonstrated comparable uptake of organizational training as evidenced by similar sorting ARC scores on the generalization trial in the fall but failed to maintain the use of this strategy as evidenced by a marked decrease in their sorting ARC scores in the spring. Although it is difficult to speculate about why this unexpected decrease in performance occurred, it nonetheless provides insight into the complex differences that exist across different cohorts of first-graders. In light of these complex and unexpected differences in both child- and teacher-level measures it seems unreasonable to expect that a true effect of teacher CPL on student performance exists to be detected in the current data set.

Limitations

As highlighted above, the failure to establish linkages between teachers' use of CPL and children's performance on memory and mathematics was unexpected given CPL has been consistently associated with children's performance in a number of previous observational and experimental investigations (Coffman et al., 2008; Grammer et al., 2016, Grammer, Coffman, & Ornstein, 2013; Hudson et al., 2018). This lack of significant findings highlights a number of limitations that require careful consideration and may be especially important in future research that aims to understand how children's strategic skills develop across different domains in relation to aspects of the educational environment in which they are embedded. In the sections that follow, key considerations related to measuring child-level performance and observing and characterizing classroom instruction are discussed.

Child-level considerations. It is important to note that at the child-level many of the significant gains in performance that were observed at the end of the year were very small and may not be substantively meaningful in terms of growth. Furthermore, such limited growth over the course of the year undermines attempts to identify linkages between teachers' use of CPL and student performance. This highlights the need to carefully consider whether the tasks being used are developmentally appropriate and sensitive enough to capture change over the course of one academic year. In regard to the math task, it is important to note that a ceiling effect was observed for children's accuracy and strategy use in the fall of first-grade. This suggests that the single-digit addition problems that were administered were not challenging enough to accurately capture variability in children's knowledge about addition facts and strategies. In order to be able to observe improvements in skill across the year it may be necessary to use a more developmentally appropriate assessment that includes more difficult items. According to the Common Core Standards by the end of first-grade students should be able to add and subtract within 20. The current measure used in this investigation included 10 single-digit addition problems. Four of the problems had sums below 10 and the remaining six problems had sums that did not exceed 18. With this in mind, it seems necessary to adjust the current measure to include addition problems with larger addends as well as subtraction problems.

In addition to providing different types of arithmetic problems, it may also be worthwhile to consider (1) different ways of characterizing children's use of addition strategies as well as (2) alternative techniques for assessing and classifying strategy use. In this investigation, an emphasis was placed on the deliberate and effortful nature of children's strategy use. As such, instances of retrieval, (i.e., the recall of a previously learned math fact), were not included in our analyses, as retrieval reflects an automatized process that develops as a consequence of repeated experience rather that deliberate engagement in goal-direct problem-solving behaviors. In light of the ceiling effect observed on the math task, it seems likely that considering children's use of retrieval, especially on single-digit problems with addends less than 10, may provide additional insight into their developing mathematical knowledge. That is, in addition to focusing on children's strategy use it may also be important to capture their knowledge for math facts as indexed by instances of retrieval. Doing so would provide a more complete picture of children's strategic repertoire as well as their general arithmetic knowledge. Additional information may also be gained by supplementing measures of effective strategy use with indicators of strategy sophistication. The same strategies identified in the current coding protocol could be further classified as either overt or covert and rank ordered from least (e.g., counting all) to most advanced (e.g., decomposition).

Another possibility for capturing a more accurate representation of children's strategy use involves using a think-aloud protocol in which students provide verbal reports while working on problems in real time (Pressley & Afflerbach, 1995). In the current investigation, strategy use was determined by observing children's overt counting behaviors during the problem-solving process in conjunction with the retrospective report children provided of their own behaviors immediately after providing an answer. At times, children seemed unable to remember how they solved the problem or explained a different strategy than the other they were observed using. Using a think-aloud protocol may help circumvent this problem and provide a richer account of children's strategic behaviors that would allow for a more qualitative assessment of children's understanding of how to execute strategies.

Classroom-level considerations. In addition to carefully evaluating the child-level tasks used in this study, it is equally important to reflect on the observational measure used to characterize teachers' instructional language. The *Taxonomy* is used to identify the presence or absence of different types of instructional language that occur in 30-second intervals, and in its current form, any given code can occur no more than once in any 30-second interval. As such, the current measure of CPL reflects the percentage of intervals in which a code occurs, which is assumed to be a proxy for the relative frequency with which that code occurred over the entire course of instruction. To illustrate, as the measure is currently used, a teacher who provides one strategy suggestion (e.g., "you can use a tens frame to help you") in a 30-second interval would receive the same score as another teacher who provided two or more strategy suggestions (e.g., “you can use a tens frame to help you, or if you prefer to use a part-part-whole map you can use that tool instead”).

Moreover, when examining composite variables such as the co-occurrence of memory requests and metacognitive information, a teacher who paired a memory request with one metacognitive question (e.g.., “who remembers the strategy we used to solve this problem yesterday?”) would receive the same score as a teacher who provided a memory request paired with a metacognitive question, a metacognitive rationale, and a strategy suggestion (e.g., “who remembers the strategy we used to solve this problem yesterday? How does a tens frame help us count-on? Today I’m going to show you another way to count-on using a part-part whole map.”) As the *Taxonomy* is currently used, a teacher receives credit if any metacognitive code and memory code is paired in the same interval but does not receive any additional credit for further instances of the same type of language. Simply put, this technique merely detects the presence of certain types of language but does not necessarily account for overall quantity. This is an important distinction, as it seems likely that the cognitive demands placed on students by a teacher who pairs multiple instances of metacognitive language together may be substantially different from those that arise from the provision of a single metacognitive question. Indeed, a teacher who asks multiple, related metacognitive questions promotes a cognitively rich dialogue that supports students’ abilities to reflect on multiple ways to solve problems, as well as to consider rationales for why different strategies may be more or less well-suited to specific problem types.

With this in mind, it is imperative that any effort to revise the current coding scheme be able to accurately distinguish between instances of unique occurrences of codes versus repetitions. For example, in the context of a review lesson on using a tens-frame a teacher may ask several different students to explain how they used the tool to help them solve a given problem. In this case, it seems reasonable to assign a metacognitive question code each time the teacher poses the question to a new child because each child may articulate a different way that a tens-frame can be used to count (e.g., make a 10, make a 5, count-on). In contrast, in the context of a lesson on greater than and less than symbols, it does not seem reasonable to assign an identifying relationship code or a semantic memory code each time that a teacher asks the class to make a decision regarding two pairs of numbers. In this case, the repetition in questioning is simply designed to give children an opportunity to repeated practice deciding which inequality symbol is correct.

Finally, providing qualitative evaluations when assigning codes may serve as an additional way to depict nuances in teachers' instructional language that are not currently captured using the *Taxonomy*. Some codes in the Taxonomy are assigned when teachers use certain key phrases, such as "ready" or "eyes on me." Each of these examples would receive an attention regulation code as these phrases are designed to instruct or guide students' attention to a task. However, an attention regulation code would also be assigned when teachers seek to highlight important parts of lessons that students should focus on, such as "I want you to pay very close attention to how these two shapes are the same and how they are different." Anecdotally, it seems to be the case that teachers often use key phrases as fillers and although the same code is assigned in both instances, they likely have very different substantive meanings in terms of the way in which students' attention or orientation to the task at hand is affected.

Future Directions

In addition to considering the limitations identified above, it is also important for future research to attempt to consider (1) when during the elementary school years it is most appropriate to begin studying children's strategy use and (2) how additional measures of children's metacognition and self-regulated learning may be able to reconcile some of the null findings reported here and contribute to a more complete understanding of children's strategy use and performance in relation to early educational experiences.

The "new" first grade. The instructional climate and demands of the first grade have changed dramatically during the span of 14 years in which the data for this investigation were collected. In 1995, Morrison and colleagues made use of established cut-off dates for school entry in order to create a "natural" experiment, taking advantage of the fact that those children whose birth dates preceded the cutoff date entered the first grade and those who missed the cutoff date were enrolled in kindergarten. In this way, it was possible to isolate the specific effect of educational experiences on children's development, given that children in both groups were effectively the same age. One set of findings from this line of work highlighted differences in children's recall in relation to the educational context in which they were embedded. Specifically, children in first-grade classrooms demonstrated significant gains in recall on deliberate memory tasks over the course of the year whereas their same-age peers in kindergarten did not. Based on these findings, which suggest that something about the first-grade experience is tied to the development of children's memory skills, the initial Classroom Memory Study (Cohort 1, 2002-2003) was designed to focus on the experiences of first-grade students and teachers.

However, more recent research examining kindergarten and first grade classrooms in 1998 and 2010 documents a number of changes that provide empirical evidence to support the popular notion that "kindergarten is the new first grade" (Bassok, Latham, & Rorem, 2016) Indeed, kindergarten classrooms today are characterized by an increasingly rigorous focus on academic skill building and test preparation. More specifically, in 1998, 13% of kindergarten teachers reported that children should be able to count to 20 at kindergarten entry. By 2010, the percentage increased to 35%. In addition to increases in academic expectation, corresponding changes in teachers' instructional practices were also observed. In 1998 83% of kindergarten teachers reported daily math instruction compared to 91% in 2010. Moreover, the percentage of teachers who reported spending more than 3 hours per day in whole-group teacher-led instruction increased from 15% in 1998 to 32% in 2010. Similarly, reported daily use of textbooks more than doubled and daily use of worksheets increased by 15%. Further comparisons demonstrated that these increases in academic rigor observed in 2010 resulted in kindergarten classrooms that were more similar in structure in and focus to first-grade classrooms of the late '90s.

In addition to substantial changes at the classroom level, Bassok and Latham (2014) have also demonstrated that kindergarteners are entering school with substantially greater early academic skills than before. Taken together, these findings provide a rationale for examining children's strategy use and performance at kindergarten entry. Doing so would provide an opportunity to capture initial variation in child-level performance during the transition to formal schooling and provide a foundation for examining (1) the potential long-term impact of early educational experiences on the emergence and subsequent refinement of children's strategic behaviors as well as (2) longitudinal associations between early instantiations of children's strategic knowledge and later indicators of academic achievement.

The role of metacognition. In order to more thoroughly understand the association between children's strategy use and performance as well as strategy use across different domains it is important for future studies to include measures of children's domain-specific metacognitive knowledge as well as domain-general indicators of children's self-regulated learning. Metacognition broadly refers to higher order declarative knowledge about cognition as well as the procedural knowledge regarding the regulation of one's cognitive activities (Flavell, 1979; Schneider & Loffler, 2016) and is believed to exert an influence on children's performance in two ways. First, metacognition affects children's initial decision to be strategic as it produces an understanding that performance will generally improve if sufficient effort is put forth in strategy selection and deployment. Second, it facilitates effective implementation of strategies and monitoring activities that enable students to evaluate performance and adapt their approach as necessary.

Consider the unexpected pattern observed in Cohorts 3 and 4, whereby students were unable to maintain the organizational training provided during the FRT task in the fall, unlike like their peers in Cohorts 1 and 2 who generalized the training after a 15-minute delay and continued to demonstrate gains in organized sorting at the end of the year. Incorporating measures of children's domain specific metacognition may provide key insight into the differential sorting patterns following training observed across cohorts. Indeed, previous microgenetic research demonstrates that the extent to which children acquired the use of an organizational strategy over an 11-week period was related to their initial metamemory knowledge prior to the start of the study. Moreover, those children who failed to acquire the use of an organizational strategy during the course of the study, but demonstrated considerable metacognitive growth, were observed to have adopted the use of the organizational strategy in a follow-up study a few months later (Schlagmüller & Schneider, 2002).

Additional longitudinal studies have enabled researchers to further consider (1) whether metacognition is a developmental antecedent of strategy use or (2) whether metacognition and strategy use develop reciprocally. In one study, Grammer and her colleagues (Grammer, Purtell, Coffman, & Ornstein, 2011) examined the longitudinal relation between children's use of an organizational strategy during the FRT task and children's metamemory (children's metacognition related specific to memory) skills and observed that children's metamemory skills preceded their strategic sorting behaviors and also predicted subsequent strategy use. Given that this research has been carried out in the context of children's memory skills, it is important for future studies to address children's domain-general metacognitive skills and strategy use more broadly by considering whether the same developmental trajectory of the metacognition-strategy use relationship emerges across different domains, including mathematics.

One possibility is that children's strategy use precedes the development of their metacognitive knowledge in domains like mathematics where specific strategies are explicitly taught. In contrast, perhaps metacognition precedes strategy use in domains such as memory where specific strategies are not often explicitly taught. Embedding assessments of children's metacognition and strategy use in microgentic and longitudinal studies designed to follow children from the transition to formal schooling throughout early elementary school will be critical to addressing these questions.

In the context of mathematics, researchers have demonstrated success using structured interviews to assess students' procedural, declarative, and situational metacognitive knowledge. This information can be gained by asking to students to explain how they solved the problem (procedural knowledge), to identify alternative ways to solve the same problem (declarative knowledge), and to provide rationales for their strategy choice (situational knowledge) (Throndsen, 2011; Carr & Jessup, 1997). In the context of memory, researchers have used a variety of techniques, including think-aloud protocols, retrospective reports, and peer tutoring paradigms (Best & Ornstein, 1986), to assess children's (1) knowledge about their own cognitive states, demands of the cognitive task at hand and relevant strategies (declarative metamemory), as well as their (2) understanding of when and why strategies are effective (conditional metamemory), and (3) ability to monitor and adapt behaviors and strategies (procedural metamemory) (Cavanaugh & Perlmutter, 1982; Schneider 2010).

The role of self-regulated learning. Future research efforts designed to examine the association between strategy use across domains, may benefit from including measures of children's self-regulated learning. Self-regulated learning (SRL) refers to the cyclical process of deliberately planning, monitoring, and regulating that students engage in in order to accomplish an academic task of goal (Zimmerman, 1989). Although SRL involves metacognitive knowledge, it is distinct from metacognition in so far as SRL is a more dynamic process whereby metacognitive knowledge and experience interact with behavioral and motivational factors to guide student learning. Indeed, SRL reflects the complex coordination between metacognitive, behavioral, and motivational components rather than their isolated contributions (Dinsmore, Alexander & Loughlin, 2008).

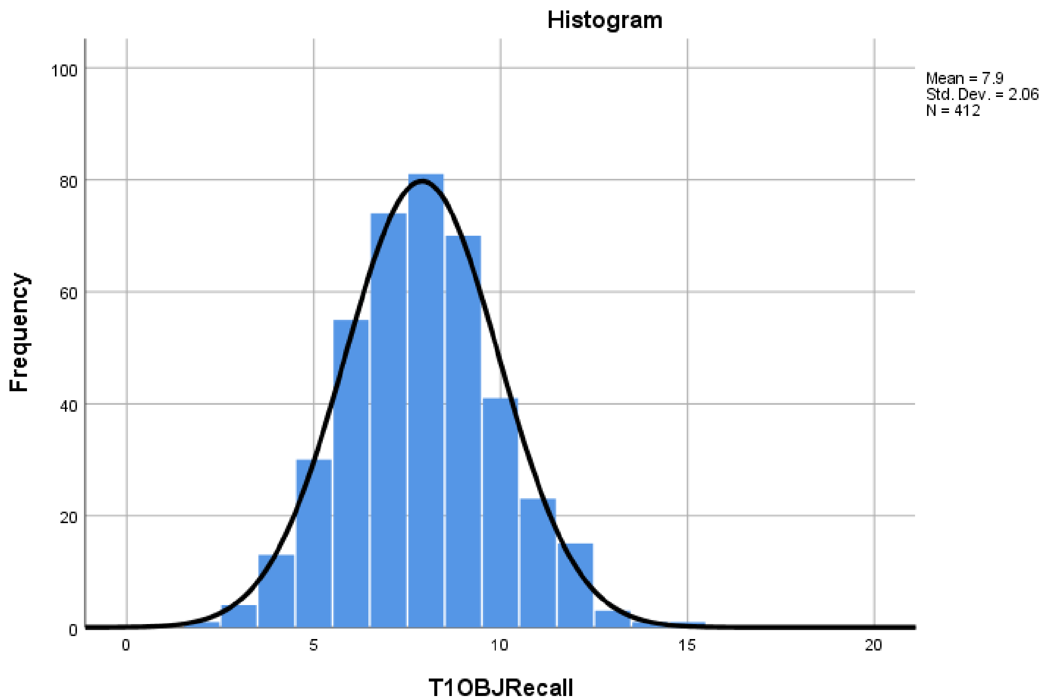
Very preliminary evidence from this investigation suggests that children's strategic knowledge may generalize across domains, as evidenced by the significant correlation between children's sorting scores on the generalization trial of FRT (following organizational training) at the beginning of first-grade and their strategy use in mathematics in the fall and spring. Whereas domain specific indicators of metacognition may be relevant to understanding children's strategy use *within* a domain, measures of self-regulated learning may be critical to understanding the complex integration and coordination of strategic skills across tasks or domains.

The construct of self-regulated learning may also be better suited to capturing individual differences in children's ability to competently direct and regulate their own actions toward learning goals than the measure of behavioral regulation used in the current study. In this study, the Head Toes Knees Shoulders was used to characterize children's self-regulation. Although successful performance on this task involves coordination of working memory, attentional flexibility, and inhibition, which are necessary for engaging in self-regulated learning, this behavioral regulation measure fails to assess the higher-order cognitive, metacognitive, and motivational components that are critical to the learning process. Measures of children's self-regulated learning, which account for their ability to control learning-directed behaviors through planning, monitoring, and evaluating are therefore more likely to be associated with children's strategy use and performance than more laboratory-based measures of children's behavioral control. One current obstacle in the field of SRL is the reliance on offline measures, such as questionnaires or interviews, which are typically collected independently from a learning task and may not correspond to children's actual behaviors (Dinsmore et al., 2008; Veenman, Van Hout-Wolters, Afflerbach, 2006). Furthermore, self-report questionnaires and interviews are difficult to complete with very young children who may be unable to understanding the questions being asked or to accurately verbalize their thoughts and behaviors. The use of observational measures of students' SRL would more accurately reflect the processes and behaviors students engage in while completing learning tasks.

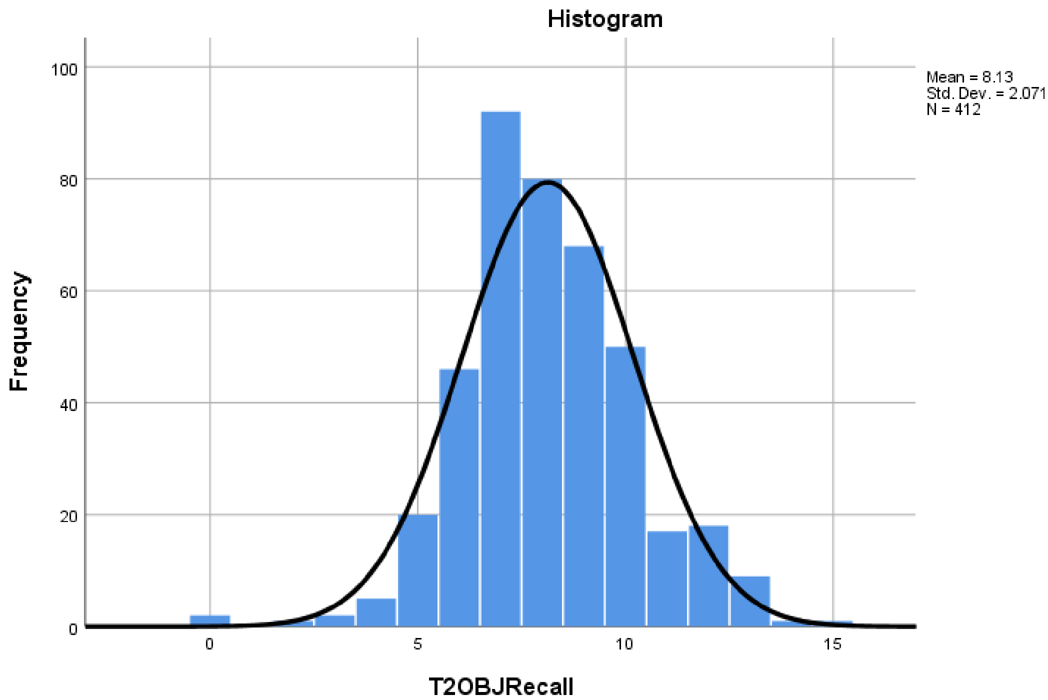
Finally, much remains to be learned about how teachers promote self-regulated learning, especially among elementary school-aged children. Pintrich (2002) suggests that metacognitive and self-regulated learning strategies be embedded in content-driven lessons and explicitly taught through modeling, shared discussion, and authentic practice. It seems possible that teachers who encourage deep levels of processing and incorporate more metacognitively rich instruction into their lessons, as indexed by CPL, may facilitate growth in their students self-regulated learning skills and consequently their strategy use and performance. In addition to examining the ways in which teachers promote the development of SRL in classroom, longitudinal work is needed to examine how children's SRL skills develop over time and across contexts.

In conclusion, the results of this study highlight the complexities of conducting longitudinal research in classrooms when the educational context itself is undergoing change. In order to more fully understand the ways in which children's strategic skills develop over time and across different academic domains future research needs to ensure that individual tasks are appropriately challenging so as to be able to detect variation in skill as well as change over time. In addition, including measures of children's domain specific metacognitive knowledge as well as indicators of their self-regulated learning skills will be critical to examining the association between strategic behaviors and subsequent performance.

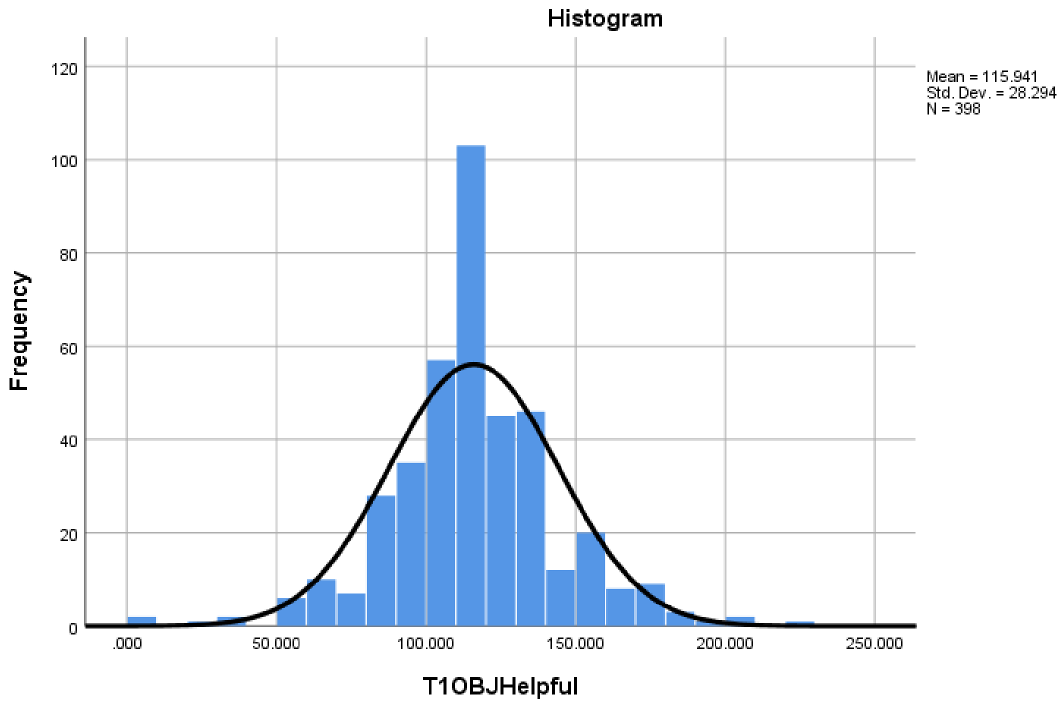
APPENDIX A: EXPLORATORY ANALYSES



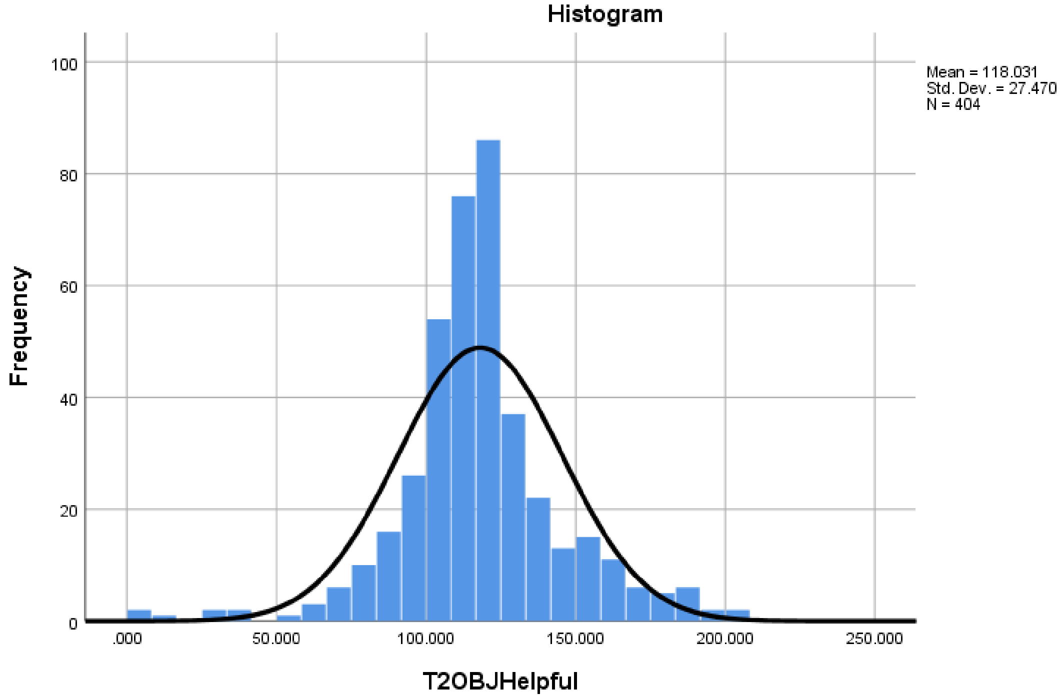
Recall on the Object Memory Task at Time 1



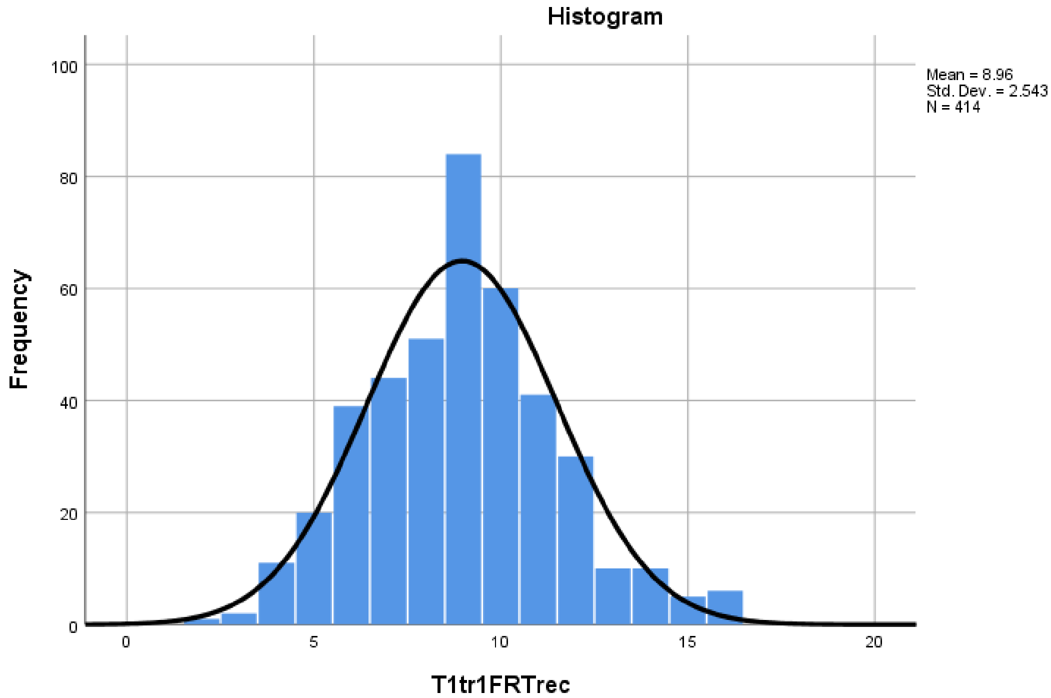
Recall on the Object Memory Task at Time 2



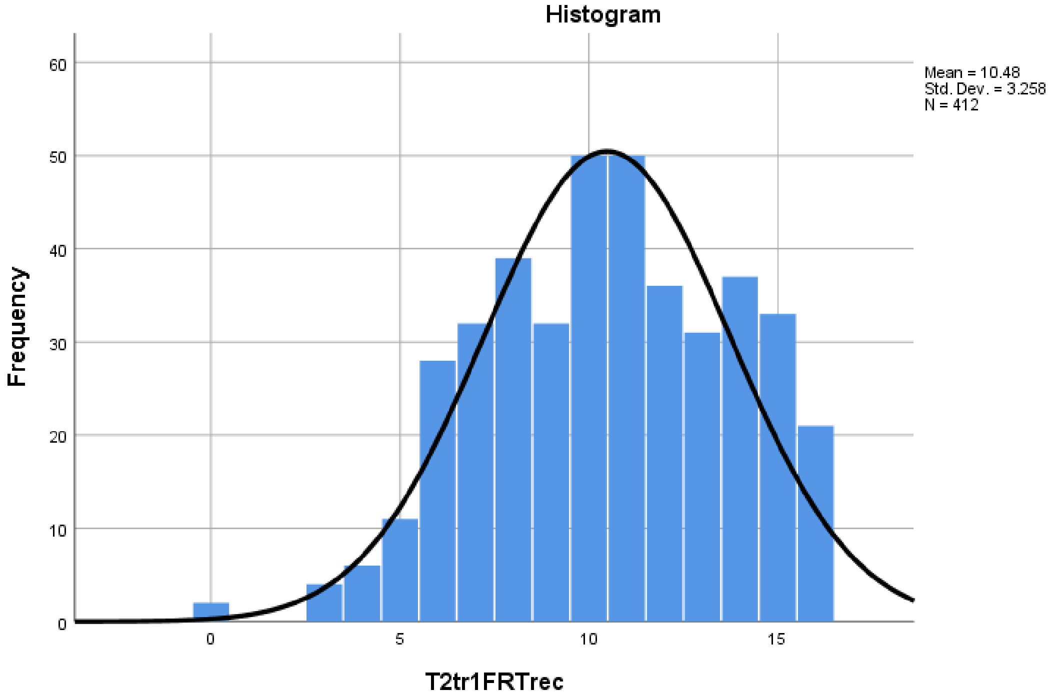
Strategy Use on the Object Memory Task at Time 1



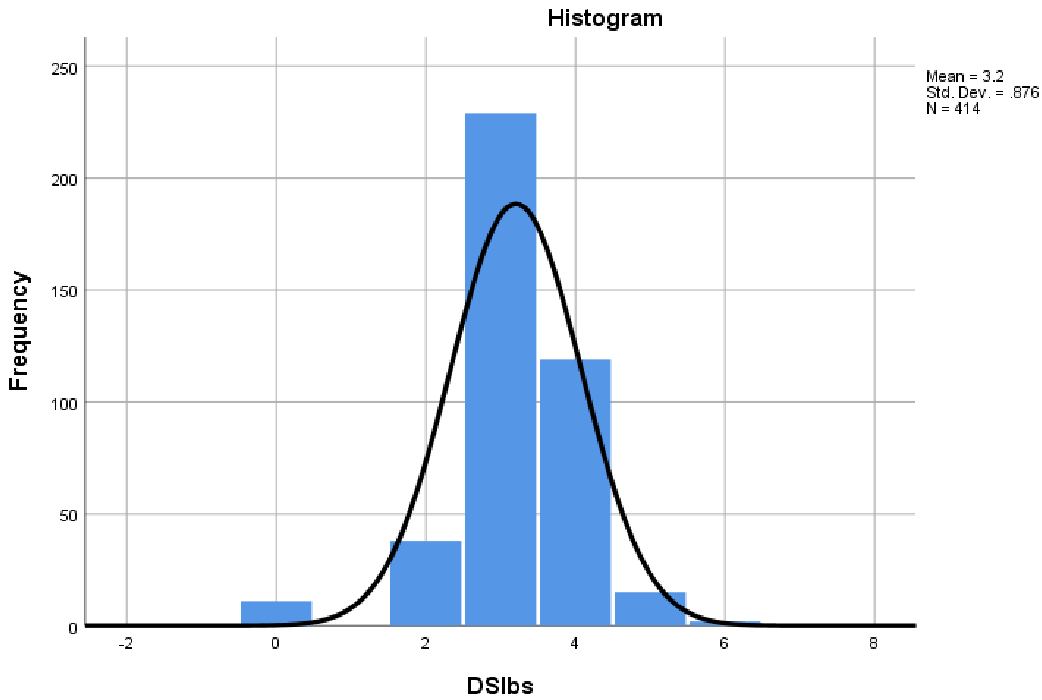
Strategy Use on the Object Memory Task at Time 2



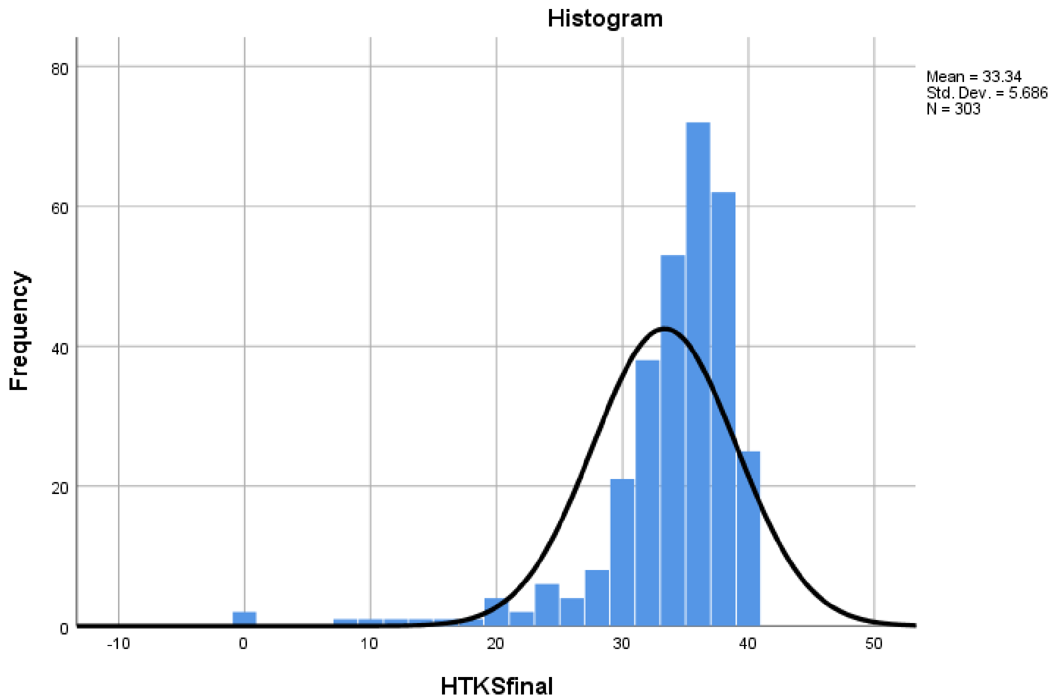
Recall on the Free Recall with Organizational Training Task at Time 1



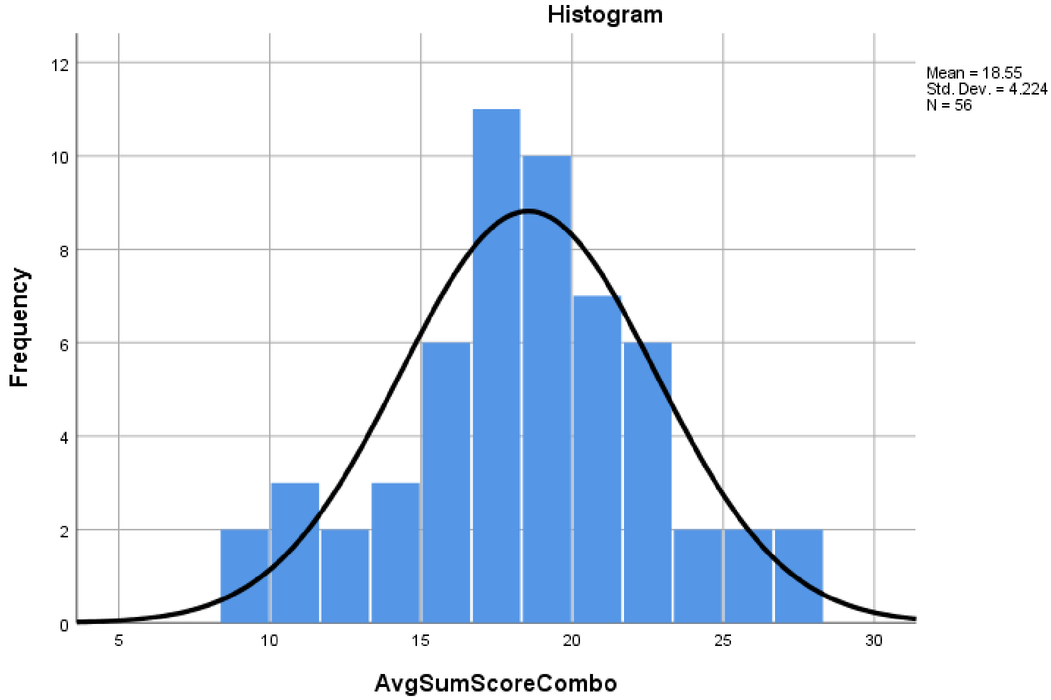
Recall on the Free Recall with Organizational Training Task at Time 2



Number of Digits Correctly Recalled in Backward Order



Total score on the Head Toes Knees Shoulders Task



Average CPL Score during Mathematics and Language Arts Instruction

APPENDIX B: RECALL IN OBJECT MEMORY TASK UNCONDITIONAL MODEL

Level 1 (within-classroom differences): deviation of child from the class average

$RecallT2_{ij} = \beta_{0j} + r_{ij}$

Level 2 (between-classroom differences): deviation of a given classroom

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

Reduced-form:

$RecallT2_{ij} = \gamma_{00} + \mu_{0j} + r_{ij}$

ICC:

.44/(.44 + 3.86) = .1023

APPENDIX C: RECALL IN OBJECT MEMORY TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 PREDICTORS

Level 1:

$$RecallT2_{ij} = \beta_{0j} + \beta_{1j}\text{RecallT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

Reduced-form:

$$RecallT2_{ij} = \gamma_{00} + \gamma_{10}\text{RecallT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$$

APPENDIX D: RECALL IN OBJECT MEMORY TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 AND LEVEL 2 PREDICTORS

Level 1:

$RecallT2_{ij} = \beta_{0j} + \beta_{1j}\text{RecallT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$RecallT2_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\text{ RecallT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij}$
$+ \mu_{0j} + r_{ij}$

APPENDIX E: RECALL IN OBJECT MEMORY TASK SLOPES-AS-OUTCOMES MODEL WITH CROSS-LEVEL INTERACTION

Level 1:

$\mathit{RecallT2}_{ij} = \beta_{0j} + \beta_{1j}\,\text{RecallT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfReg}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + \gamma_{11}\text{CPL}_{j} + \mu_{1j}$

Reduced-form:

$\mathit{RecallT2}_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\,\text{RecallT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{11}\text{CPL}_{j}\text{SelfReg}_{ij} + \mu_{0j} + \mu_{1j}\text{SelfReg}_{ij} + r_{ij}$

APPENDIX F: STRATEGY USE IN OBJECT MEMORY TASK UNCONDITIONAL MODEL

Level 1 (within-classroom differences): deviation of child from the class average

$StrategyUseT2_{ij} = \beta_{0j} + r_{ij}$

Level 2 (between-classroom differences): deviation of a given classroom

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

Reduced-form:

$StrategyUseT2_{ij} = \gamma_{00} + \mu_{0j} + r_{ij}$

ICC:

25.83/(25.83 + 729.30) = .03

APPENDIX G: STRATEGY USE IN OBJECT MEMORY TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 PREDICTORS

Level 1:

$\textit{StrategyUseT2}_{ij} = \beta_{0j} + \beta_{1j}\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$\textit{StrategyT2}_{ij} = \gamma_{00} + \gamma_{10}\text{StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$

\

APPENDIX H: STRATEGY USE IN OBJECT MEMORY TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 AND LEVEL 2 PREDICTORS

Level 1:

$StrategyUseT2_{ij} = \beta_{0j} + \beta_{1j}\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$StrategyUseT2_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \gamma_{10}\ \text{StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$

APPENDIX I: STRATEGY USE IN OBJECT MEMORY TASK SLOPES-AS-OUTCOMES MODEL WITH CROSS-LEVEL INTERACTION

Level 1:

$StrategyUseT2_{ij} = \beta_{0j} + \beta_{1j}\,\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfReg}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + \gamma_{11}\text{CPL}_{j} + \mu_{1j}$

Reduced-form:

$StrategyUseT2_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\text{StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{11}\text{CPL}_{j}\text{SelfReg}_{ij} + \mu_{0j} + \mu_{1j}\text{SelfReg}_{ij} + r_{ij}$

APPENDIX J: RECALL IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK UNCONDITIONAL MODEL

Level 1 (within-classroom differences): deviation of child from the class average

$RecallT2_{ij} = \beta_{0j} + r_{ij}$

Level 2 (between-classroom differences): deviation of a given classroom

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

Reduced-form:

$RecallT2_{ij} = \gamma_{00} + \mu_{0j} + r_{ij}$

ICC:

.23/(.23 + 10.39) = .0216

APPENDIX K: RECALL IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 PREDICTORS

Level 1:

$\mathit{RecallT2}_{ij} = \beta_{0j} + \beta_{1j}\text{RecallT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$\mathit{RecallT2}_{ij} = \gamma_{00} + \gamma_{10}\text{RecallT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$

APPENDIX L: RECALL IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 AND LEVEL 2 PREDICTORS

Level 1:

$$RecallT2_{ij} = \beta_{0j} + \beta_{1j}\text{RecallT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

Reduced-form:

$$RecallT2_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\text{ RecallT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$$

APPENDIX M: RECALL IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK SLOPES-AS-OUTCOMES MODEL WITH CROSS-LEVEL INTERACTION

Level 1:

$\mathit{RecallT2}_{ij} = \beta_{0j} + \beta_{1j}\,\text{RecallT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfReg}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + \gamma_{11}\text{CPL}_j + \mu_{1j}$

Reduced-form:

$\mathit{RecallT2}_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \gamma_{10}\,\text{RecallT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{11}\text{CPL}_j\text{SelfReg}_{ij} + \mu_{0j} + \mu_{1j}\text{SelfReg}_{ij} + r_{ij}$

APPENDIX N: STRATEGY USE IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK UNCONDITIONAL MODEL

Level 1 (within-classroom differences): deviation of child from the class average

$StrategyUseT2_{ij} = \beta_{0j} + r_{ij}$

Level 2 (between-classroom differences): deviation of a given classroom

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

Reduced-form:

$StrategyUseT2_{ij} = \gamma_{00} + \mu_{0j} + r_{ij}$

ICC:

.02/(.02 + .32) = .06

APPENDIX O: STRATEGY USE IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 PREDICTORS

Level 1:

$StrategyUseT2_{ij} = \beta_{0j} + \beta_{1j}\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$StrategyT2_{ij} = \gamma_{00} + \gamma_{10}\text{StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$

APPENDIX P: STRATEGY USE IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 AND LEVEL 2 PREDICTORS

Level 1:

$\textit{StrategyUseT2}_{ij} = \beta_{0j} + \beta_{1j}\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$\textit{StrategyUseT2}_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\text{ StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$

APPENDIX Q: STRATEGY USE IN FREE RECALL WITH ORGANIZATIONAL TRAINING TASK SLOPES-AS-OUTCOMES MODEL WITH CROSS-LEVEL INTERACTION

Level 1:

$\textit{StrategyUseT2}_{ij} = \beta_{0j} + \beta_{1j}\,\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfReg}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + \gamma_{11}\text{CPL}_j + \mu_{1j}$

Reduced-form:

$\textit{StrategyUseT2}_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \gamma_{10}\text{StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{11}\text{CPL}_j\text{SelfReg}_{ij} + \mu_{0j} + \mu_{1j}\text{SelfReg}_{ij} + r_{ij}$

APPENDIX R: ACCURACY UNCONDITIONAL MODEL

Level 1 (within-classroom differences): deviation of child from the class average

$AccuracyT2_{ij} = \beta_{0j} + r_{ij}$

Level 2 (between-classroom differences): deviation of a given classroom

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

Reduced-form:

$AccuracyT2_{ij} = \gamma_{00} + \mu_{0j} + r_{ij}$

ICC:

.04/(.04 + 4.48) = .0088

APPENDIX S: ACCURACY RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 PREDICTORS

Level 1:

$$AccuracyT2_{ij} = \beta_{0j} + \beta_{1j}\text{AccuracyT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$

Reduced-form:

$$AccuracylT2_{ij} = \gamma_{00} + \gamma_{10}\text{AccuracyT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$$

APPENDIX T: ACCURACY RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 AND LEVEL 2 PREDICTORS

Level 1:

$$\mathit{AccuracyT2}_{ij} = \beta_{0j} + \beta_{1j}\text{AccuracyT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$

Reduced-form:

$$\mathit{AccuracyT2}_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\,\text{AccuracyT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$$

APPENDIX U: ACCURACY SLOPES-AS-OUTCOMES MODEL WITH CROSS-LEVEL INTERACTION

Level 1:

$AccuracyT2_{ij} = \beta_{0j} + \beta_{1j}\text{ AccuracyT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfReg}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + \gamma_{11}\text{CPL}_j + \mu_{1j}$

Reduced-form:

$AccuracyT2_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_j + \gamma_{10}\text{ AccuracylT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} +$
$\gamma_{11}\text{CPL}_j\text{SelfReg}_{ij} + \mu_{0j} + \mu_{1j}\text{SelfReg}_{ij} + r_{ij}$

APPENDIX V: STRATEGY USE IN MATHEMATICS UNCONDITIONAL MODEL

Level 1 (within-classroom differences): deviation of child from the class average

$StrategyUseT2_{ij} = \beta_{0j} + r_{ij}$

Level 2 (between-classroom differences): deviation of a given classroom

$\beta_{0j} = \gamma_{00} + \mu_{0j}$

Reduced-form:

$StrategyUseT2_{ij} = \gamma_{00} + \mu_{0j} + r_{ij}$

ICC:

.00/(.00 + .05) = 0

APPENDIX W: STRATEGY USE IN MATHEMATICS RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 PREDICTORS

Level 1:

$$\textit{StrategyUseT2}_{ij} = \beta_{0j} + \beta_{1j}\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

Reduced-form:

$$\textit{StrategyT2}_{ij} = \gamma_{00} + \gamma_{10}\text{StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$$

APPENDIX X: STRATEGY USE IN MATHEMATICS RANDOM INTERCEPT REGRESSION MODEL WITH LEVEL 1 AND LEVEL 2 PREDICTORS

Level 1:

$\textit{StrategyUseT2}_{ij} = \beta_{0j} + \beta_{1j}\text{StrategyUseT1}_{ij} + \beta_{2j}\text{WorkingMemory}_{ij} + \beta_{3j}\text{SelfRegulation}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced-form:

$\textit{StrategyUseT2}_{ij} = \gamma_{00} + \gamma_{01}\text{CPL}_{j} + \gamma_{10}\text{ StrategyUseT1}_{ij} + \gamma_{20}\text{WorkingMemory}_{ij} + \gamma_{30}\text{SelfRegulation}_{ij} + \mu_{0j} + r_{ij}$

APPENDIX Y: STRATEGY USE IN MATHEMATICS SLOPES-AS-OUTCOMES MODEL WITH CROSS-LEVEL INTERACTION

Level 1:

$StrategyUseT2_{ij} = \beta_{0j} + \beta_{1j}\,\mathrm{StrategyUseT1}_{ij} + \beta_{2j}\mathrm{WorkingMemory}_{ij} + \beta_{3j}\mathrm{SelfReg}_{ij} + r_{ij}$

Level 2:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\mathrm{CPL}_{j} + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + \gamma_{11}\mathrm{CPL}_{j} + \mu_{1j}$

Reduced-form:

$StrategyUseT2_{ij} = \gamma_{00} + \gamma_{01}\mathrm{CPL}_{j} + \gamma_{10}\mathrm{StrategyUseT1}_{ij} + \gamma_{20}\mathrm{WorkingMemory}_{ij} + \gamma_{11}\mathrm{CPL}_{j}\mathrm{SelfReg}_{ij} + \mu_{0j} + \mu_{1j}\mathrm{SelfReg}_{ij} + r_{ij}$

Table 1

Task-Oriented Strategic Behaviors Coded in the Object Memory Task

Behavior	Definition
Association	A child verbalizes an association with or elaboration about an object (e.g., "I have a necklace like this").
Categorization	A child groups two or more items together, either verbally or physically. The presence of the group category must be either obvious to the observer or identified verbally by the child.
Covert mnemonic activity	A child's behavior suggests studying, as in moving the lips as if rehearsing or alternating between closing the eyes and looking at the objects, as in self-testing.
Manipulation	A child makes any type of manual contact with the objects that does not involve their unique properties.
Naming	A child provides any label – conventional or personal – for an object, without further description.
Object talk	A child discusses physical properties or an object (e.g., "These glasses are green").
Pointing	A child points to the objects without touching or moving them.
Visual examination	A child scans the objects without touching any of them.

Table 2

Code Definitions for the Taxonomy of Teacher Behaviors

Category	Definition
Non-instruction	The teacher is not engaged in a memory or instructional activity
Instructional activities	
Book reading	Reading aloud to the group
General information giving	Presentation of factual information
Prospective summary	Description of upcoming events
Specific task information	Instructions for performing a particular task
Cognitive structuring	
Attention regulation:	Directing or focusing student's attention
Behavioral goal	to reprimand or guide behavior
Instructional goal	to instruct or guide attention to do a task
Massed repetition	Performance of an activity in unison
Identifying features	Generating features of a category i.e.: parts of a bug
Categorization	Verbally or physically putting class material into categories
Identifying relationships	Comparison of at least 2 items, emphasizing similarities and differences
Connections to personal experiences:	Associating a prior experience to a current classroom activity
Home	outside of school
School	in school
Drawing inferences	Predicting an outcome or intentions or desires of another
Visual imagery	Creating visual mental images that relates to the material
Memory requests	
Episodic	Retrieval of a specific past event in or out of the classroom
Semantic	Retrieval of an already learned fact, idea, or object

Procedural	How to perform a series of activities with a behavioral goal
Prospective	Non-instructional task to be completed in the future
Anticipated	Expectation for child to remember information w/o a given strategy
Metacognitive instruction	
Metacognitive rationale	Provides rationale for strategy use or for organizing or self-regulation
Metacognitive questioning	Asks child to provide potential strategy or rationale for strategy choice
Suggestion	Recommends a method for remembering information
Suppression	Asks student to refrain from using an unhelpful or inappropriate method
Replacement	Recommends an alternative strategy

Table 3

Component Codes from the Taxonomy Used to Index Cognitive Processing Language

Code	Definition	Example
Strategy Suggestion	Recommending that a child adopt a method or procedure for remembering or processing information	"Use your doubles facts to help you solve this problem"
Metacognitive Question	Requesting that a child provide a potential strategy, a utilized strategy, or rationale for a strategy they have indicated using	"Who can think of different strategy for solving this problem?"
Co-occurrence of Deliberate Memory and Instructional Activities	Requesting information from children's memory while also presenting instructional information	"Today we are going to continue our lesson on even and odd numbers. Who remembers what 'even' means?"
Co-occurrence of Deliberate Memory and Cognitive Structuring Activities	Requesting information from children's memory while simultaneously facilitating encoding and processing by focusing attention or organizing material	"Yesterday we talked about similarities and differences between squares and rectangles. Who can remind me what squares and rectangles have in common?"
Co-occurrence of Deliberate Memory and Metacognitive Information	Requesting information from children's memory while providing or soliciting metacognitive information	"Before we get started let's make a list to remind ourselves of the different strategies and tools that we can use to help us solve these kinds of equations."

Table 4

Sample Instruction from High CPL Classrooms in Mathematics and Language Arts

High CPL Mathematics Example	High CPL Language Arts Example
In this example, the teacher is showing the class two different ways to solve subtraction problems. T: You can think about addition to help you subtract. That's one way to solve a subtraction problem. And that's what we've been doing with our part-part-whole maps. So if you want to think addition – if that helps you – and you're doing "6 -2" – the way that we think addition to help us subtract is by putting our whole at the top of our part -part-whole map. Raise your hand if you can tell me what the whole is in the subtraction sentence? S1: Six. T: You got it. Six is the biggest number so it is the whole. What is the part that we know? S2: Two. T: Two. So I can draw two circle or I can put the number 2 up there. To think addition to help us subtract we can count on. We are trying to get to 6 and the part that we know is two so we can count on, 3, 4, 5, 6 (places four counters on the part-part-whole map). What is the missing part? S3: 4. T: So that's one way to"6-2". The other way is to "take away" because that's what it means to subtract – to take away. I don't care which way you solve the problem, whichever way makes more sense to you. To take away we take the whole and we put that many counters on our table. I'm going to put 6 because it's my biggest number. The subtraction sign	*In this example, the teacher is showing students cards with words written on them and asking the students to decide where each word belongs on the word wall.* T: Raise your hand if you can read this word for me. Alandra? Who can help Alandra read this word? Prophet? S1: "Here". T: Not here. It doesn't have an "r" in it. Think about the sounds that you're making. What's this word, Anthony? S2: "His." T: "His." If I cover up the "h" what word do you see? All: Is. T: That should help you when you get to this word and it's kind of tricky. We know "is" that's a pretty easy word. So when we see "is" and we put the "h" in front of it, it is going to say "his." Where does "his" need to go on our word wall? S3: On the "N's" T: Close to an "N" but this letter has a little bit longer line. Who can help out – what letter is this? S4: "H" T: We have two words under "h" let's read our two words. All: "he" "his" T: Ok, last word. If you aren't sure try to put those sounds together. That's what we do

means take away. So we are going to take away how many, everybody? All: Two. T: You can either flip them over to yellow to show that you are taking them away or you can take them off your sheet altogether. How many are left? All: Four. T: Did we get the same answer? All: Yes. T: Yes. These are just two different ways to solve the same problem.	when we’re not sure about a word. What’s this word, Alexandra? S5: Her. T: Her. Thank you for thinking. Let’s spell it. All: H-e-r “her”. T: Where does “her” need to go on the word wall? S5: H T: So we have three “h” words now. We made books last week with those words on them. Please make sure that you are using your books that you make here to practice your words during the week so that when we get ready to put them up you know them and can read them and write them.

Table 5

Children's Strategies (adapted from Siegler & Jenkins, 1989)

Strategy	Typical Use of Strategy to Solve 3 + 5
Sum	Put up 3 fingers and count "1, 2, 3" and put up 5 fingers and count "1, 2, 3, 4, 5". Then count by saying "1, 2, 3, 4, 5, 6, 7, 8".
Short-cut Sum	Count by saying "1, 2, 3, 4, 5, 6, 7, 8" while simultaneously putting up one finger on each count.
Min	Count by saying "5, 6, 7, 8" or 6, 7, 8" perhaps while simultaneously putting up one finger on each count beyond 5.
Max	Count by saying "3, 4, 5, 6, 7, 8" or "4, 5, 6, 7, 8" perhaps simultaneously putting up one finger on each count beyond 3.
Finger Recognition	Put up 3 fingers, put up 5 fingers, say "8" without counting.
Decomposition	Say "3 + 5 is like 4 + 4, so it's 8"

Table 6

Child-level Tasks and Measures

Task	Measure	Definition
Object Memory Task	Recall	Total number of items recalled
	Helpful Behaviors	Total amount of time spent engaged in strategic behaviors during the study period
Free Recall Task	Recall	Total number of items recall
	Sorting ARC score	An adjusted ratio of categorical grouping at input in which chance grouping is set at zero and perfect grouping at one
	Strategy Set Size	Total number of strategies used during study period
Math Problem Solving Task	Accuracy	Total number of correct answers
	Effective Strategy Use	Percentage of time an addition strategy was used to produce a correct answer
Digit Span	Working Memory	The longest string of digits produced without error in backward order
Head-Toes-Knees-Shoulders	Self-regulation	Ability to remember instructions and inhibit dominate response in order respond correctly

Table 7

Summary of Overall Child-level Performance in the Fall and Spring

	Fall	Spring
Measures	*M(SD)*	*M(SD)*
Memory Measures		
Object Memory Task		
Recall	7.90 (2.06)	8.13 (2.07)
Helpful Behaviors	115.94 (28.29)	118.03 (27.47)
Free Recall Task		
Recall		
Baseline	8.96 (2.54)	-
Generalization	9.44 (3.42)	10.49 (3.26)
Sorting		
Baseline	-.07 (.40)	-
Generalization	.43 (.60)	.47 (.58)
Math Measures		
Addition Strategy Task		
Accuracy	7.57 (2.65)	8.45 (2.13)
Strategy use	.80 (.25)	.85 (.23)
Ancillary Measures		
Digit span		
Longest backward string	3.20 (.88)	-
Head Toes Knees Shoulders		
Behavioral Regulation	33.34 (5.69)	-

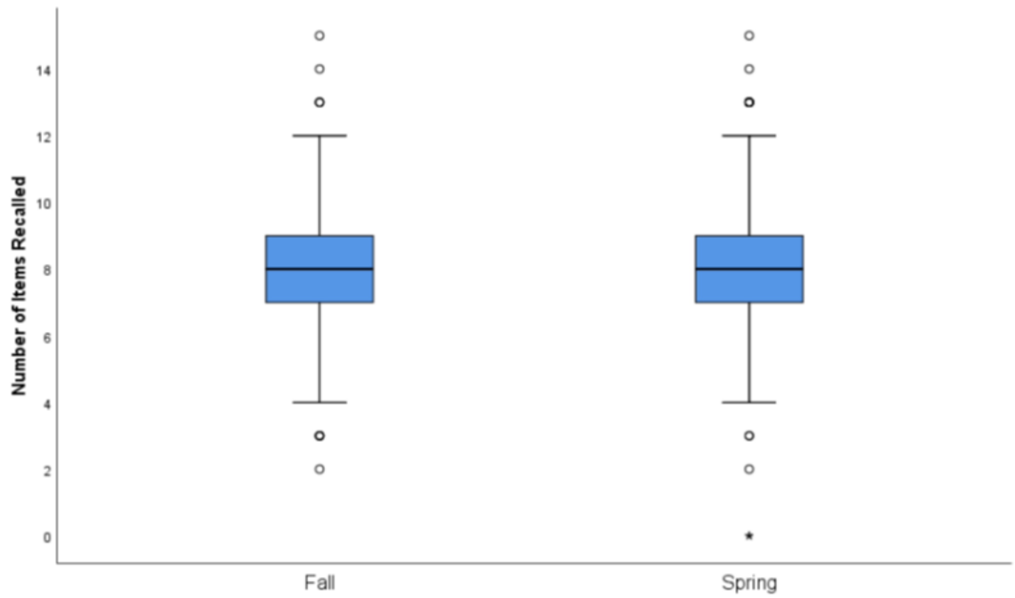


Figure 1: *Number of Items Recalled during the Object Memory Task in the Fall and Spring*

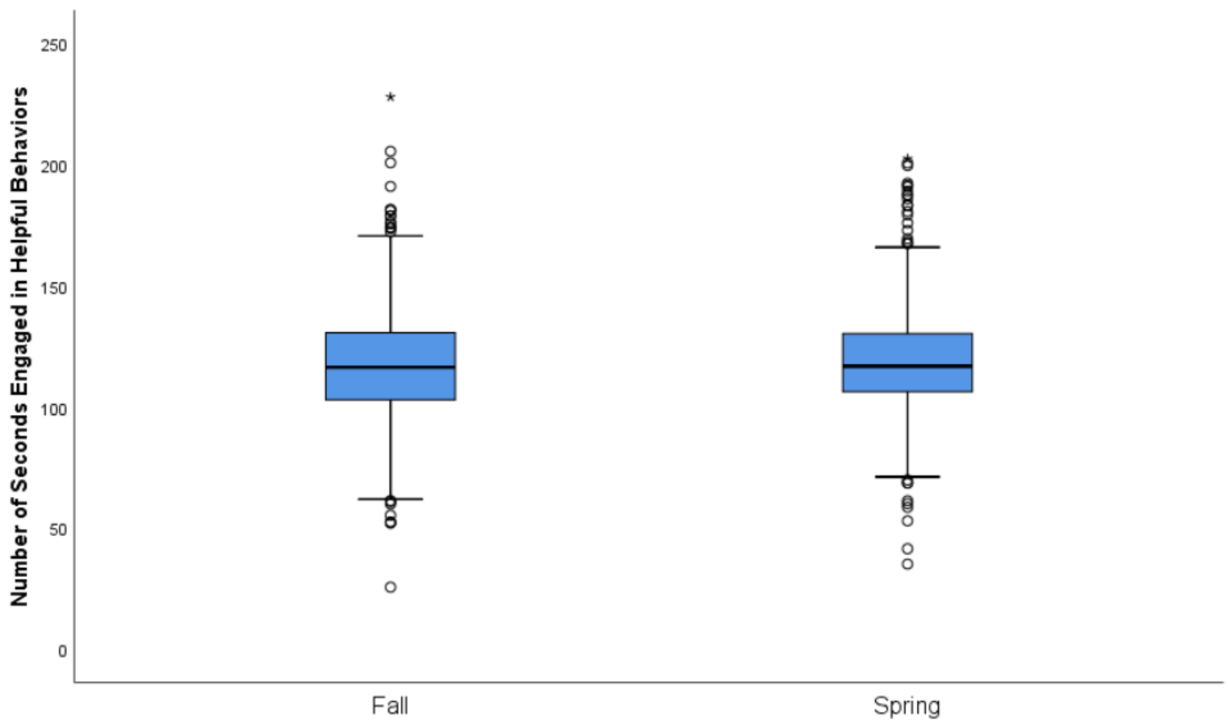


Figure 2: *Amount of Time Spent Engaged in Helpful Behavior during the Object Memory Task in the Fall and Spring*

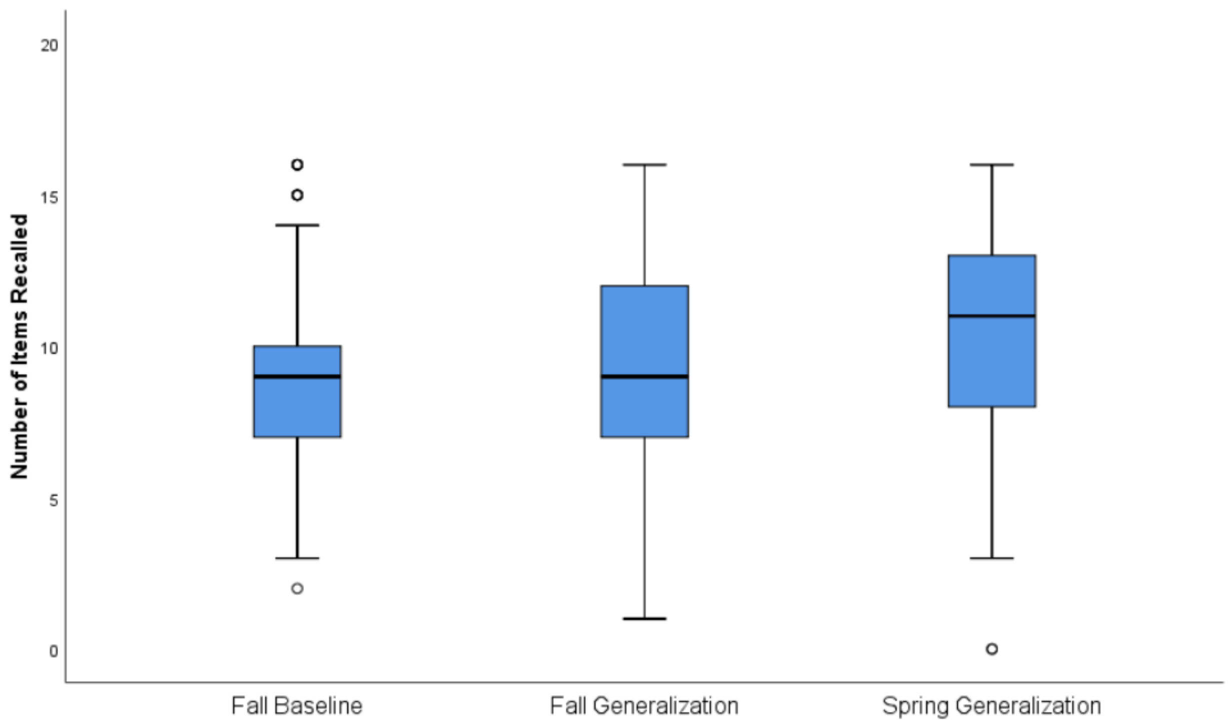


Figure 3: *Number of Items Recalled during the Free Recall with Organizational Training Task in the Fall and Spring*

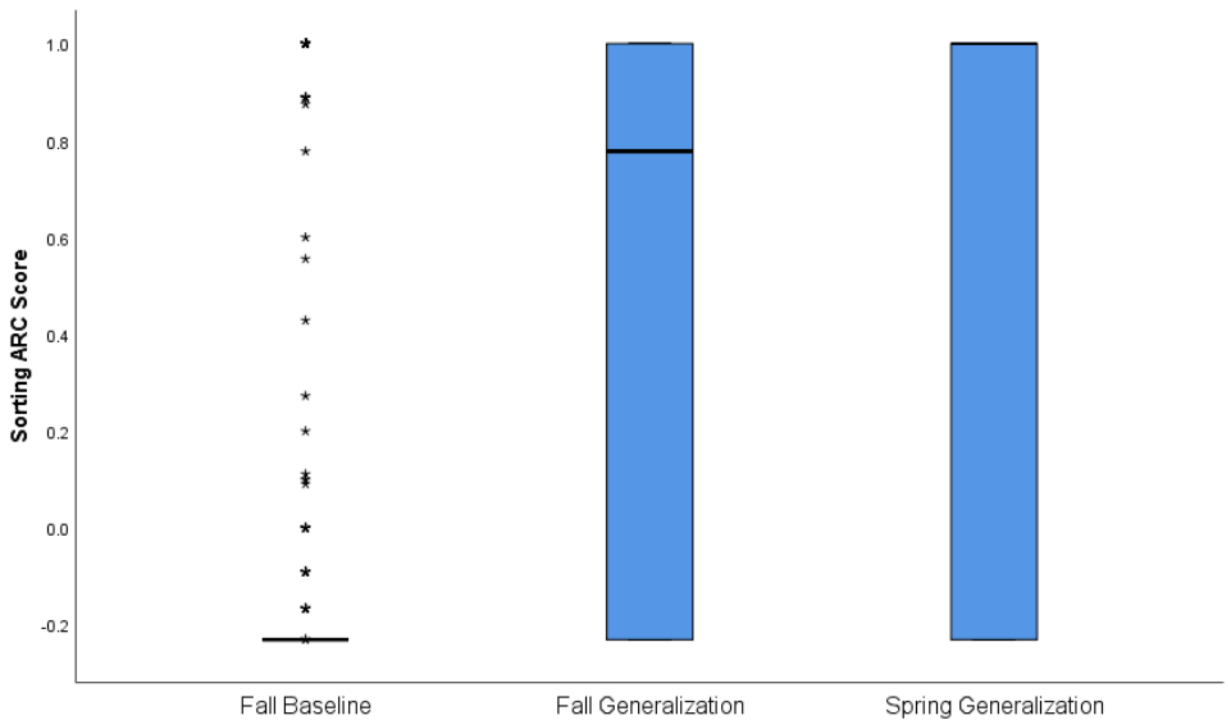


Figure 4: *Sorting ARC Scores during the Free Recall with Organizational Training Task in the Fall and Spring*

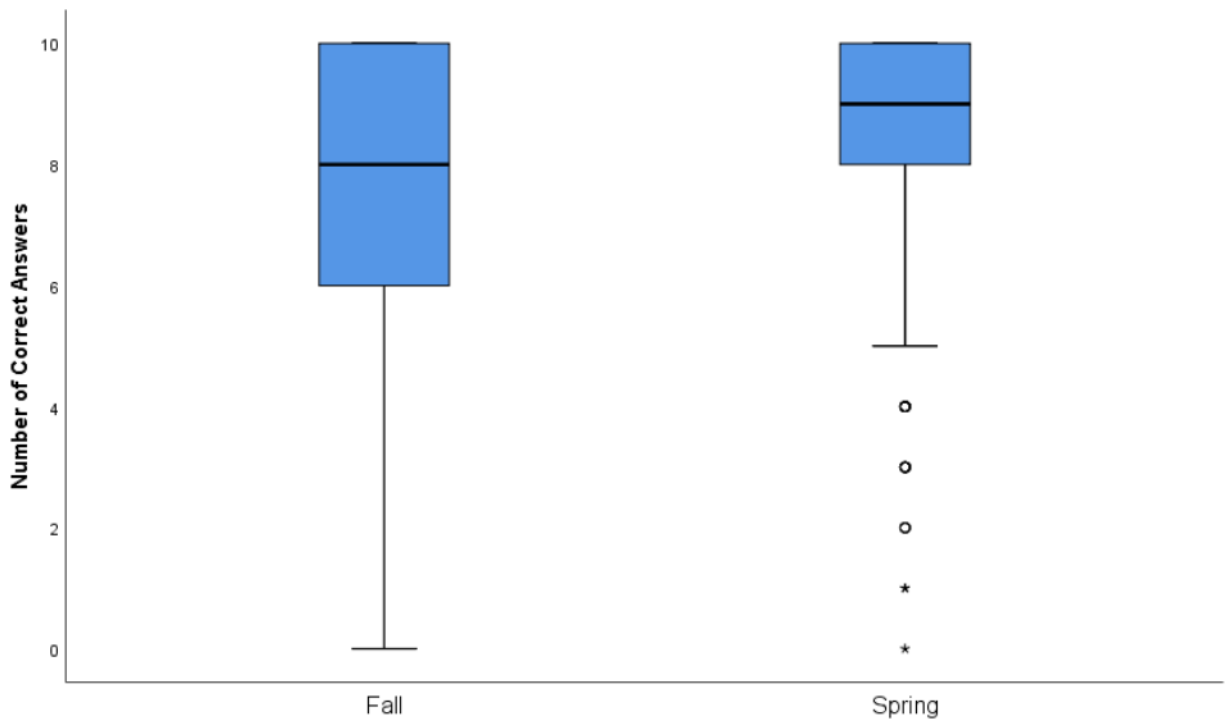


Figure 5: *Number of Correct Answers Provided during the Mathematics Task*

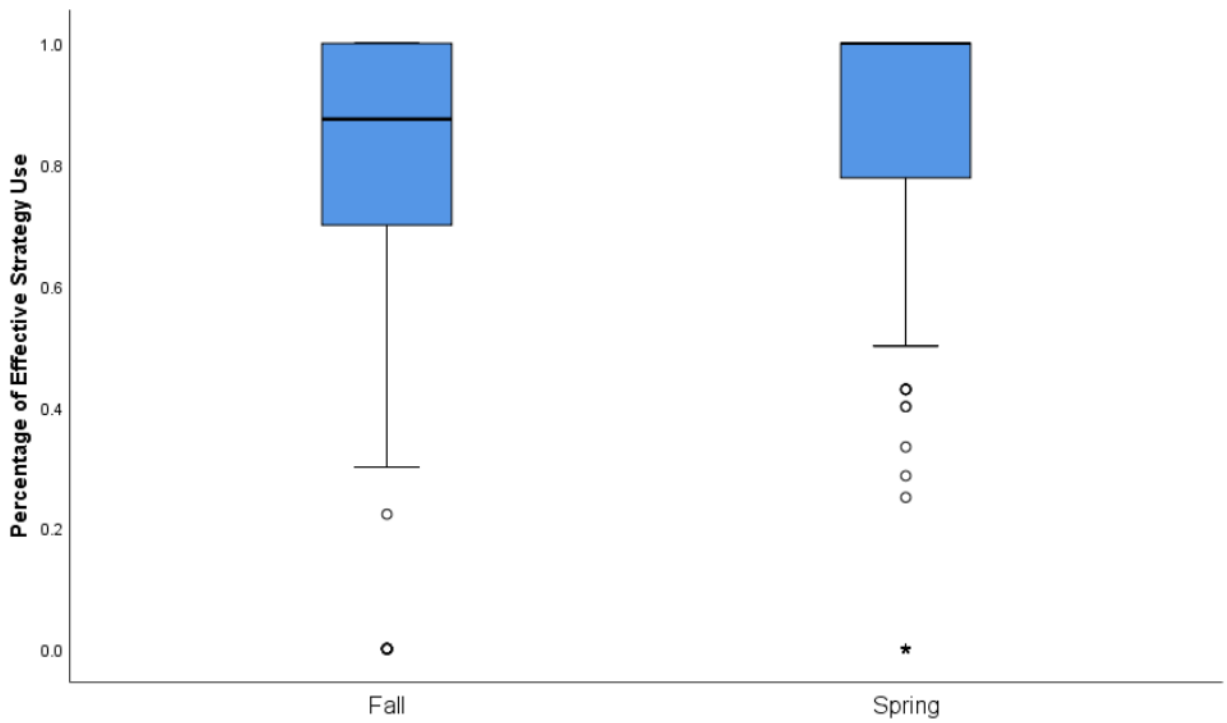


Figure 6: *Percentage of Effective Strategy Use during the Mathematics Task*

Table 8

Summary of Correlations between Recall and Strategy use during the Object Memory Task during in the Fall and Spring

		1	2	3	4	5	6
1.	Recall (Fall)	-					
2.	Helpful Behaviors (Fall)	.21*	-				
3.	Recall (Spring)	.37*	.09	-			
4.	Helpful Behaviors (Spring)	.13*	.41*	.24*	-		
5.	Self-regulation	.08	.06	.16*	.11	-	
6.	Working Memory	.09	.04	.13*	.11*	.33*	-

*p<.05

Table 9

Summary of Correlations between Recall and Strategy use in Free Recall with Organizational Training Task in the Fall and Spring

		1	2	3	4	5	6	7	8
1.	Baseline Recall (Fall)	-							
2.	Baseline Sorting (Fall)	.43*	-						
3.	Generalization Recall (Fall)	.31*	.15*	-					
4.	Generalization Sorting (Fall)	.16*	.27*	.52*	-				
5.	Generalization Recall (Spring)	.32*	16.*	.48*	38.*	-			
6.	Generalization Sorting (Spring)	.23*	.28*	.29*	.40*	.62*	-		
7.	Self-regulation	.18*	.12*	.16*	.14*	.26*	.21*	-	
8.	Working Memory	.15*	.08	.16*	.21*	.23*	.20*	.33*	-

*p<.05

Table 10

Summary of Correlations between Accuracy and Strategy use in Mathematics in the Fall and Spring

		1	2	3	4	5	6
1.	Accuracy (Fall)	-					
2.	Effective Strategy Use (Fall)	.69*	-				
3.	Accuracy (Spring)	.75*	.49*	-			
4.	Effective Strategy Use (Spring)	.32*	.23*	.57*	-		
5.	Self-regulation	.44*	.37*	.44*	.17*	-	
6.	Working Memory	.44*	.32*	.40*	.27	.33*	-

*$p<.05$

Table 11

Summary of Correlations between Recall in the Object Memory Task, the Free Recall with Organizational Training Task and Accuracy in Math

		1	2	3	4	5	6	7
1.	OBJ Recall (Fall)	-						
2.	FRT Baseline Recall (Fall)	.26*	-					
3.	FRT Generalization Recall (Fall)	.25*	.31*	-				
4.	Math Accuracy (Fall)	.06	.12*	.24*	-			
5.	OBJ Recall (Spring)	.37*	.29*	.32*	.12*	-		
6.	FRT Generalization Recall (Spring)	.28*	.32*	.48*	.25*	.40*	-	
7.	Math Accuracy (Spring)	.06	.09	.19*	.75*	.10	.21*	-

*p<.05

Table 12

Summary of Correlations between Strategy Use in the Object Memory Task, the Free Recall with Organizational Training Task, and the Math Task

		1	2	3	4	5	6	7
1.	OBJ Helpful (Fall)	-						
2.	FRT Baseline Sorting ARC (Fall)	.02	-					
3.	FRT Generalization Sorting ARC (Fall)	.06	.27*	-				
4.	Math Strategy Effectiveness (Fall)	.11	-.03	.18*	-			
5.	OBJ Helpful (Spring)	.41*	.05	.02	.08	-		
6.	FRT Generalization Sorting ARC (Spring)	-.01	.28*	.40*	.07	.10*	-	
7.	Math Strategy Effectiveness (Spring)	-.04	-.07	.15*	.23*	.01	.06	-

*p<.05

Table 13

Mean and Range of Percent Occurrence of Teacher Behaviors for Whole Sample Across Subjects

Taxonomy Codes	Mathematics	Language Arts	Mathematics and Language Arts
Non-Instructional/Non-Memory Relevant	2.54% (0.00% - 9.17%)	2.43% (0.00% - 10.00%)	2.56% (0.42% - 7.08%)
Instructional Activities – Category Total	87.68% (60.83% - 98.30%)	90.13% (72.50 – 99.17%)	82.65% (53.33% - 98.75%)
Book Reading	2.29% (0.00% - 20.80%)	13.42% (0.00% - 38.33%)	7.69% (0.00% - 25.83%)
General Information Giving	71.64% (26.67% - 95.00%)	74.33% (35.00% - 95.83%)	64.39% (31.67% - 95.42%)
Prospective Summary	5.04% (0.00% - 16.67%)	7.16% (0.00% - 22.50%)	6.28% (0.83% - 13.33%)
Specific Task Information	38.75% (16.67% - 73.30%)	31.79% (11.67% - 55.83%)	35.45% (16.25% - 63.75%)
Cognitive Structuring – Category Total	45.29% (16.67% - 69.17%)	44.70% (18.33% - 68.33%)	45.00% (20.42% - 65.42%)
Attention Regulation- Behavioral Goal	15.25% (1.67% - 35.83%)	17.10% (0.83% - 44.17%)	16.65% (1.67% - 37.50%)
Attention Regulation- Instructional Goal	22.48% (7.50% - 53.33%)	18.11% (4.17% - 45.00%)	20.18% (6.67% - 46.67%)
Massed Repetition	5.86% (0.00% - 27.50%)	9.48% (0.00% - 40.83%)	7.64% (0.00% - 29.58%)

Identifying Features	3.33% (0.00% - 35.83%)	2.80% (0.00% - 26.67%)	3.14% (0.00% - 17.92%)
Categorization	1.24% (0.00% - 11.70%)	3.78% (0.00% - 23.33%)	2.49% (0.00% - 11.67%)
Identifying Relationships	12.01% (0.00% - 29.17%)	5.60% (0.00% - 15.00%)	8.67% (0.83% - 19.58%)
Connections- Personal Experiences at Home	0.68% (0.00% - 5.83%)	1.99% (0.00% - 10.00%)	1.34% (0.00% - 5.83%)
Connections- Personal Experiences at School	10.01% (3.33% - 25.83%)	11.71% (0.00% - 21.67%)	11.07% (1.67% - 22.08%)
Drawing Inferences	.58% (0.00% - 5.83%)	3.10% (0.00% - 15.83%)	1.83% (0.00% - 7.92%)
Visual Imagery	.12% (0.00% - 1.67%)	.36% (0.00% - 5.00%)	0.24% (0.00% - 2.50%)
Memory Requests – Category Total	62.14% (35.85% - 83.30%)	50.68% (5.00% - 85.00%)	55.59% (32.50% - 80.42%)
Episodic	1.77% (0.00% - 11.67%)	3.93% (0.00% - 15.00%)	2.87% (0.00% - 10.42%)
Semantic	57.99% (29.17% - 81.67%)	47.16% (18.17% - 80.83%)	52.13% (24.58% - 77.08%)
Procedural	1.46% (0.00% - 8.30%)	1.00% (0.00% - 9.17%)	1.27% (0.00% - 4.58%)
Prospective	0.84% (0.00% - 8.30%)	1.21% (0.00% - 5.00%)	1.03% (0.00% - 4.17%)
Anticipated	3.88% (0.00% - 19.17%)	7.13% (0.00% - 20.832%)	5.58% (0.42% - 14.17%)

Metacognitive Instruction – Category Total	14.51% (0.00% - 40.83%)	7.60% (0.00% - 19.17%)	11.15% (2.50% - 25.83%)
Metacognitive Rationale	2.17% (0.00% - 10.80)	1.64% (0.00% - 8.33%)	1.94% (0.00% - 9.17%)
Metacognitive Questioning	7.58% (0.00% - 18.33%)	3.13% (0.00% - 14.17%)	5.40% (0.42% - 12.50%)
Suggestion	7.56% (0.00% - 25.00%)	4.45% (0.00% - 14.17%)	6.02% (0.00% - 15.83%)
Suppression	0.25% (0.00% - 3.33%)	0.09% (0.00% - 0.83%)	0.18% (0.00% - 1.67%)
Replacement	0.22% (0.00% - 5.00%)	0.06% (0.00% - 0.83%)	0.15% (0.00% - 2.50%)

Table 14

Mean and Range of Cognitive Processing Codes for Whole Sample across Subjects

Code	Mathematics	Language Arts	Mathematics & Language Arts
Strategy Suggestions	7.59%	4.45%	6.02%
	(0.00% - 25.00%)	(0.00% - 14.17%)	(0.00% - 15.83%)
Metacognitive Questions	7.66%	3.13%	5.40%
	(0.00% – 18.33%)	(0.00% - 14.17%)	(0.42% - 12.50%)
Memory Requests & Instructional Activities	53.07%	47.57%	47.39%
	(25.83% - 80.00%)	(17.50% - 84.17%)	(25.83% - 79.17%)
Memory Requests & Cognitive Structuring Activities	28.87%	25.94%	27.38%
	(8.33% - 54.17%)	(5.83% - 53.33%)	(9.58% - 46.67%)
Memory Requests & Metacognitive Information	8.67%	4.39%	6.54%
	(0.00% - 24.17%)	(0.00% - 14.17%)	(1.25% - 18.33%)

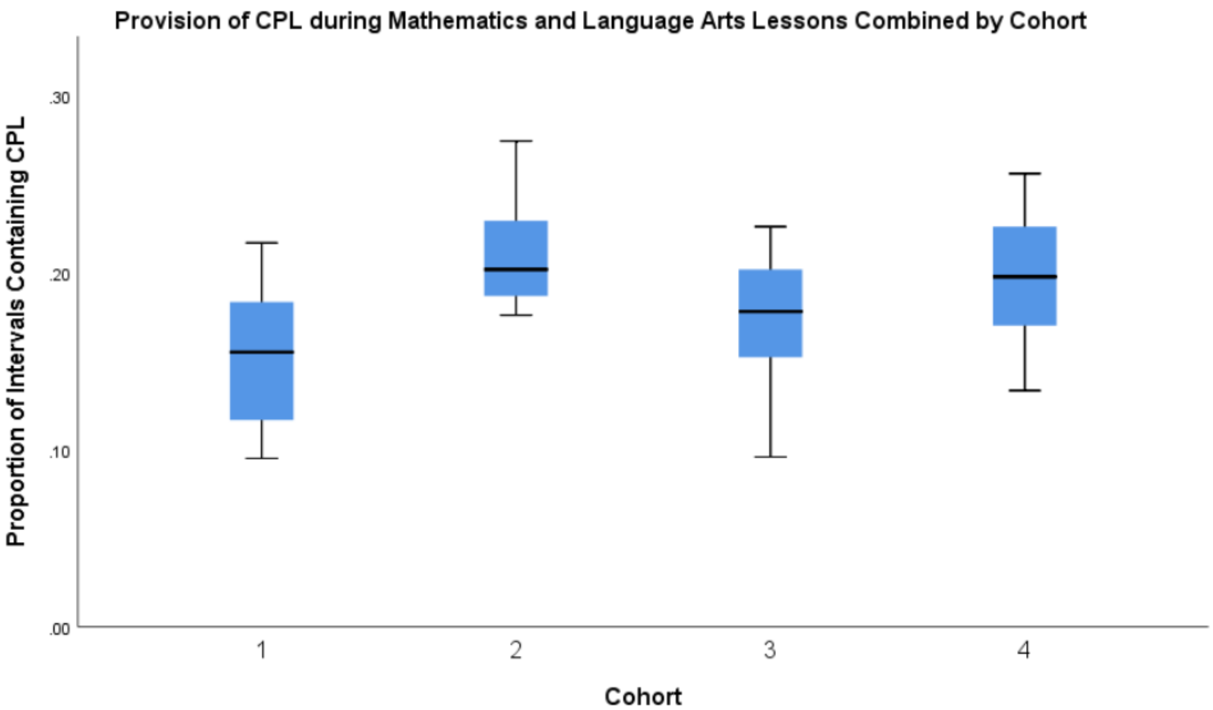


Figure 7: *Provision of CPL during Mathematics and Language Arts Lessons Combined across Cohorts*

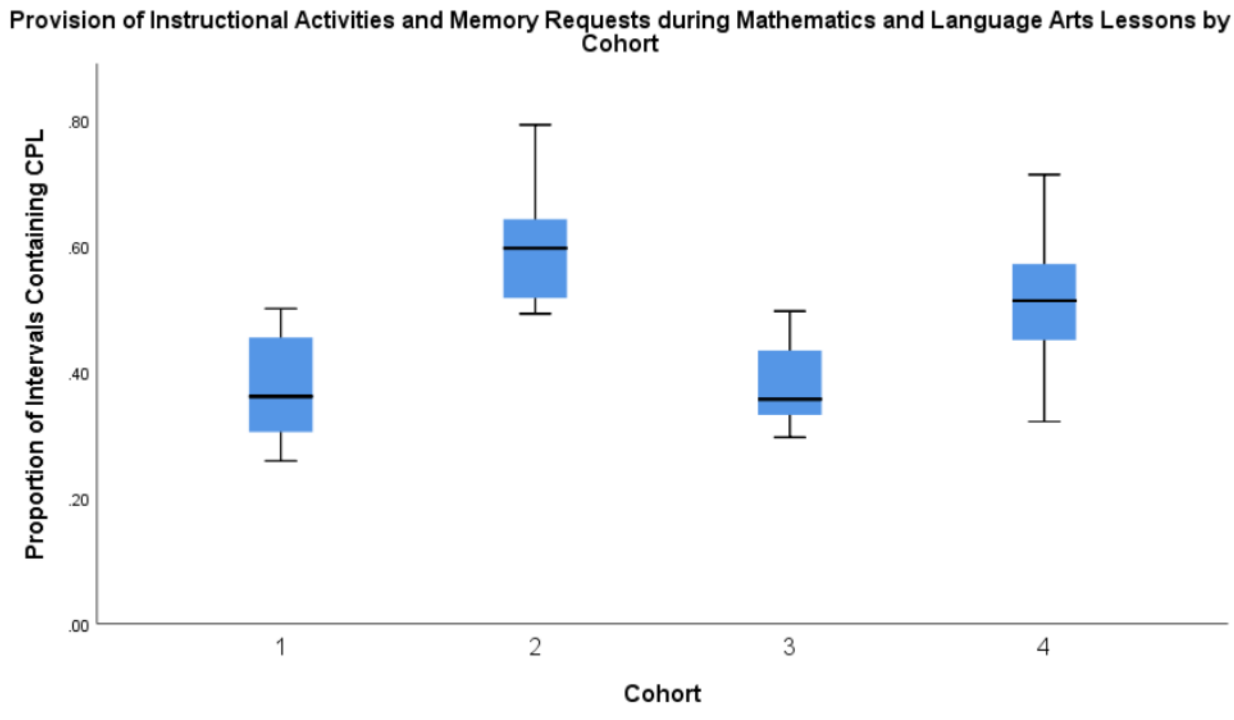


Figure 8: *Provision of Instructional Activities and Memory Requests during Mathematics and Language Arts Lessons by Cohort*

Table 15
Means and ANOVA Summary Statistics for CPL during Mathematics and Language Arts across Cohorts

	Cohort 1 *M(SD)*	Cohort 2 *M(SD)*	Cohort 3 *M(SD)*	Cohort 4 *M(SD)*	*F*	*p*
CPL Score	15.36% (3.74%)	21.22% (3.29%)	17.22% (3.73%)	19.71% (3.77%)	7.72	<.01
Strategy Suggestions	4.91% (3.64%)	5.78% (3.61%)	7.15% (3.44%)	6.51% (3.36%)	.98	=.41
Metacognitive Questions	4.88% (2.66%)	5.00% (1.92%)	6.70% (3.57%)	5.26% (3.13%)	1.15	=.34
Memory Requests and Instructional Activities	37.68% (8.22%)	59.44% (8.23%)	37.95% (6.46)	50.80% (11.14%)	22.19	<.01
Memory Requests and Cognitive Structuring Activities	23.45% (8.13)	28.99% (7.68%)	27.15% (8.94%)	29.71% (8.06%)	1.66	=.19
Memory Requests and Metacognitive Information	5.86% (3.77%)	6.89% (3.72%)	7.15% (3.56%)	6.25% (3.63%)	.35	=.79

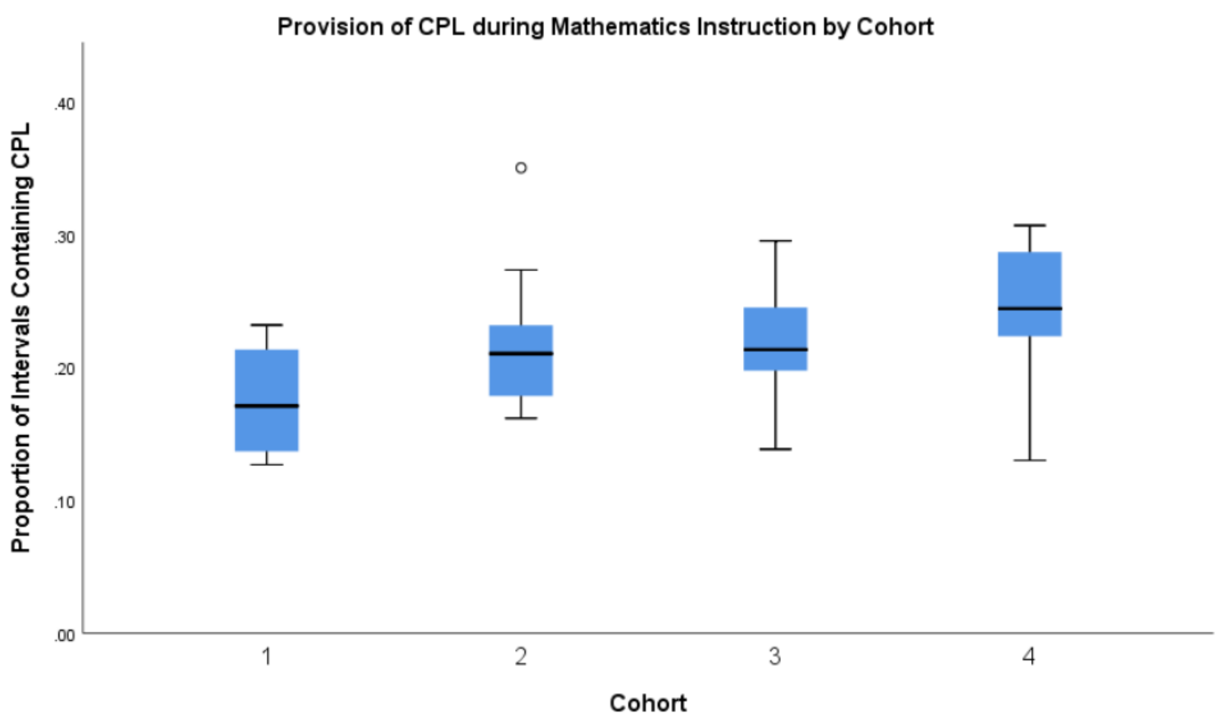


Figure 9: *Provision of CPL during Mathematics Instruction by Cohort*

Table 16

Means and ANOVA Summary Statistics for CPL during Mathematics across Cohorts

	Cohort 1 M(SD)	Cohort 2 M(SD)	Cohort 3 M(SD)	Cohort 4 M(SD)	*F*	*p*
CPL Score	17.20% (3.72%)	21.54% (4.76%)	21.93% (4.56%)	24.19% (5.19%)	5.66	<.01
Strategy Suggestions	6.67%	6.39%	7.72%	9.70%	1.01	=.40
Metacognitive Questions	7.08% (5.22%)	7.45% (3.45%)	7.50% (3.98)	8.33% (5.45%)	.19	=.90
Memory Requests and Instructional Activities	40.18% (9.02%)	57.65% (9.11%)	57.06% (7.88%)	58.57% (10.59%)	18.18	<.01
Memory Requests and Cognitive Structuring Activities	23.21 (8.87%)	27.99% (9.10%)	29.50% (9.87%)	34.76% (10.99%)	3.37	<.05
Memory Requests and Metacognitive Information	8.87% (6.43%)	8.23% (5.33%)	7.89% (3.91%)	9.58% (6.02%)	.27	=.84

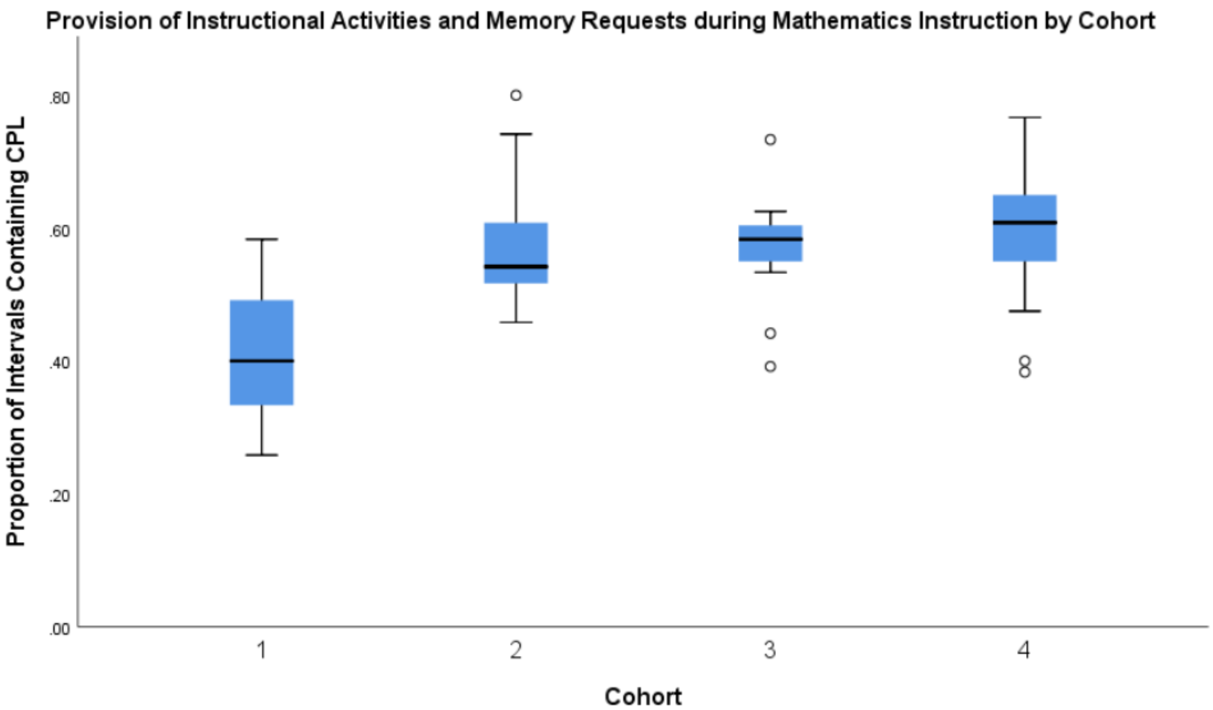


Figure 10: *Provision of Instructional Activities and Memory Requests during Mathematics Instruction by Cohort*

Table 17

Means and ANOVA Summary Statistics for CPL during Language Arts across Cohorts

	Cohort 1 M(SD)	Cohort 2 M(SD)	Cohort 3 M(SD)	Cohort 4 M(SD)	*F*	*p*
CPL Score	13.51% (4.58%)	20.93% (3.94%)	17.32% (5.41%)	15.73% (3.94)	7.66	<.01
Strategy Suggestions	3.15% (4.26%)	5.15% (4.05%)	6.04% (4.07%)	3.46% (2.70%)	1.69	=.18
Metacognitive Questions	2.68% (3.92%)	2.55% (1.37%)	5.48% (4.42%)	2.28% (2.24%)	2.98	<.05
Memory Requests and Instructional Activities	35.18% (8.34%)	61.32% (10.92%)	43.96% (16.96%)	46.28% (11.84%)	12.55	<.01
Memory Requests and Cognitive Structuring Activities	23.69% (11.05%)	30.15% (10.54%)	24.72% (10.36%)	23.97% (8.17%)	1.43	=.25
Memory Requests and Metacognitive Information	2.86% (3.64%)	5.49% (3.35%)	6.39% (4.51%)	2.76% (1.94%)	3.80	<.05

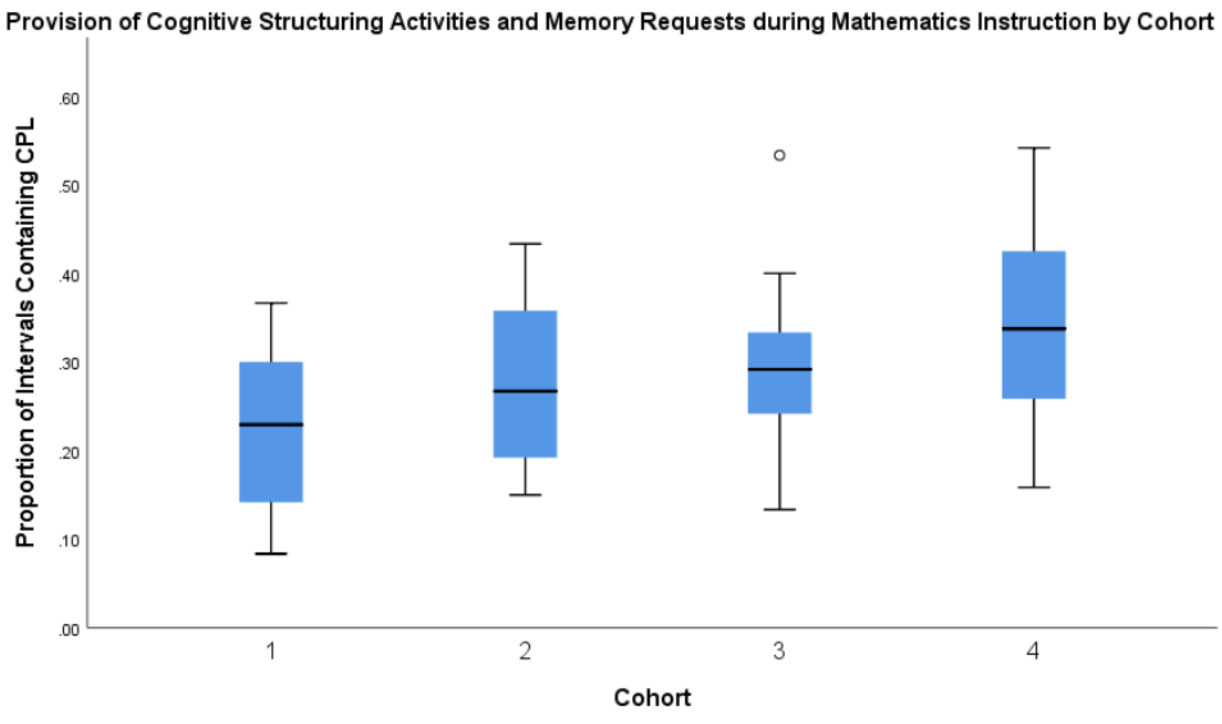


Figure 11: *Provision of Cognitive Structuring Activities and Memory Requests during Mathematics Instruction by Cohort*

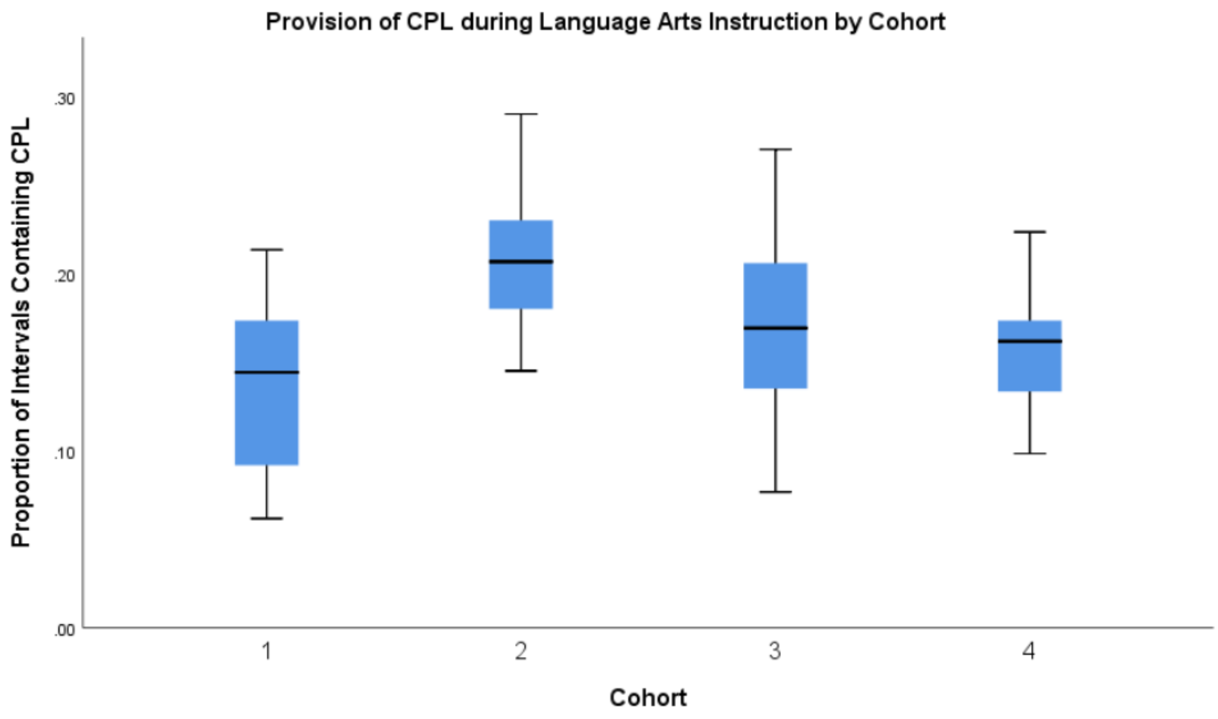


Figure 12: *Provision of CPL during Language Arts Instruction by Cohort*

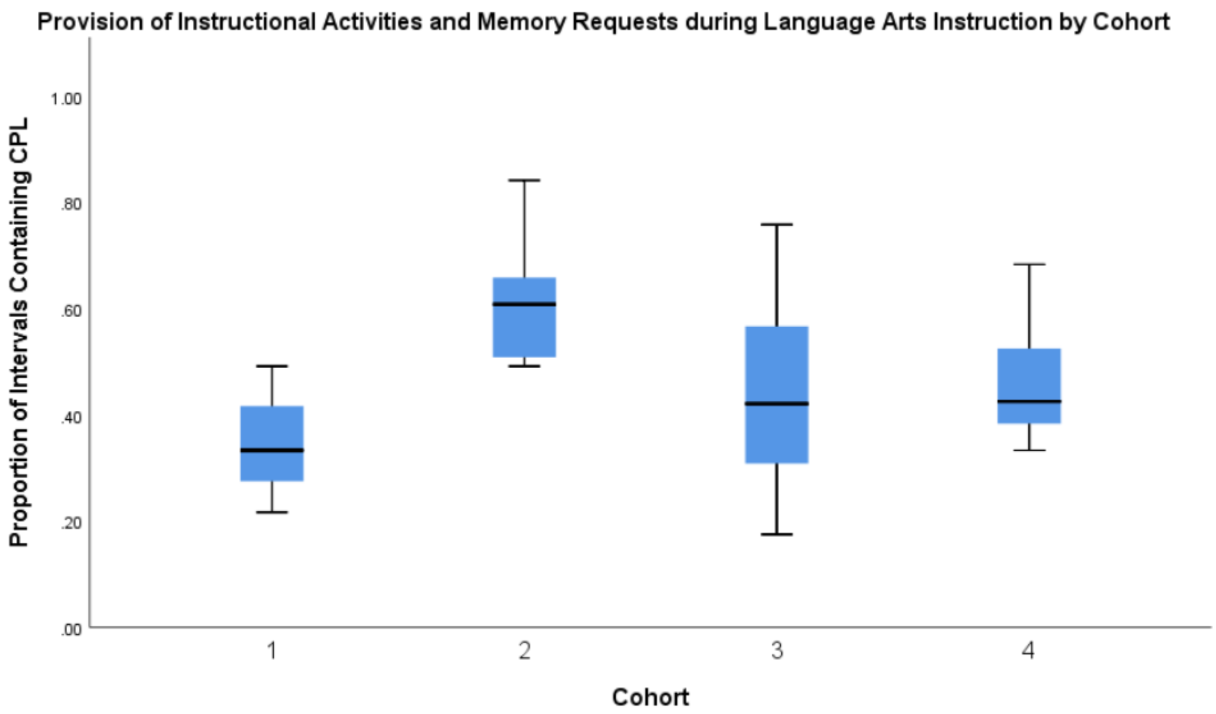


Figure 13: *Provision of Instructional Activities and Memory Requests during Language Arts Instruction by Cohort*

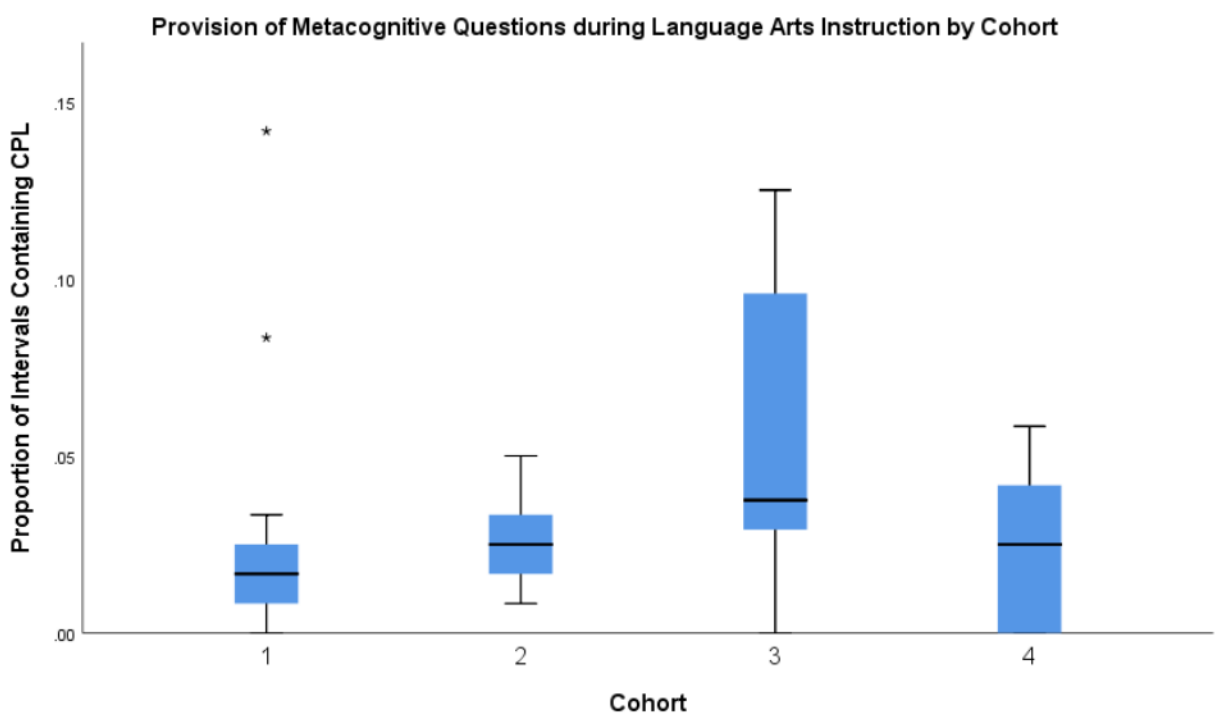


Figure 14: *Provision of Metacognitive Questions during Language Arts Instruction by Cohort*

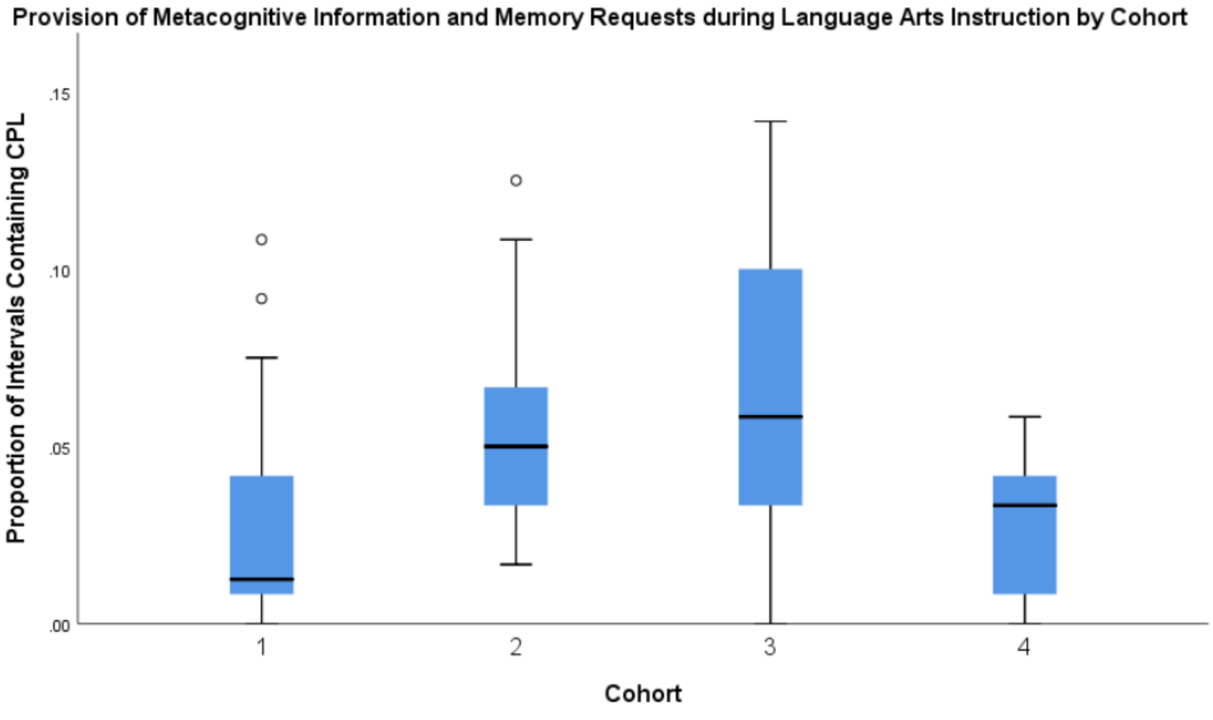


Figure 15: *Provision of Metacognitive Information and Memory Requests during Language Arts Instruction by Cohort*

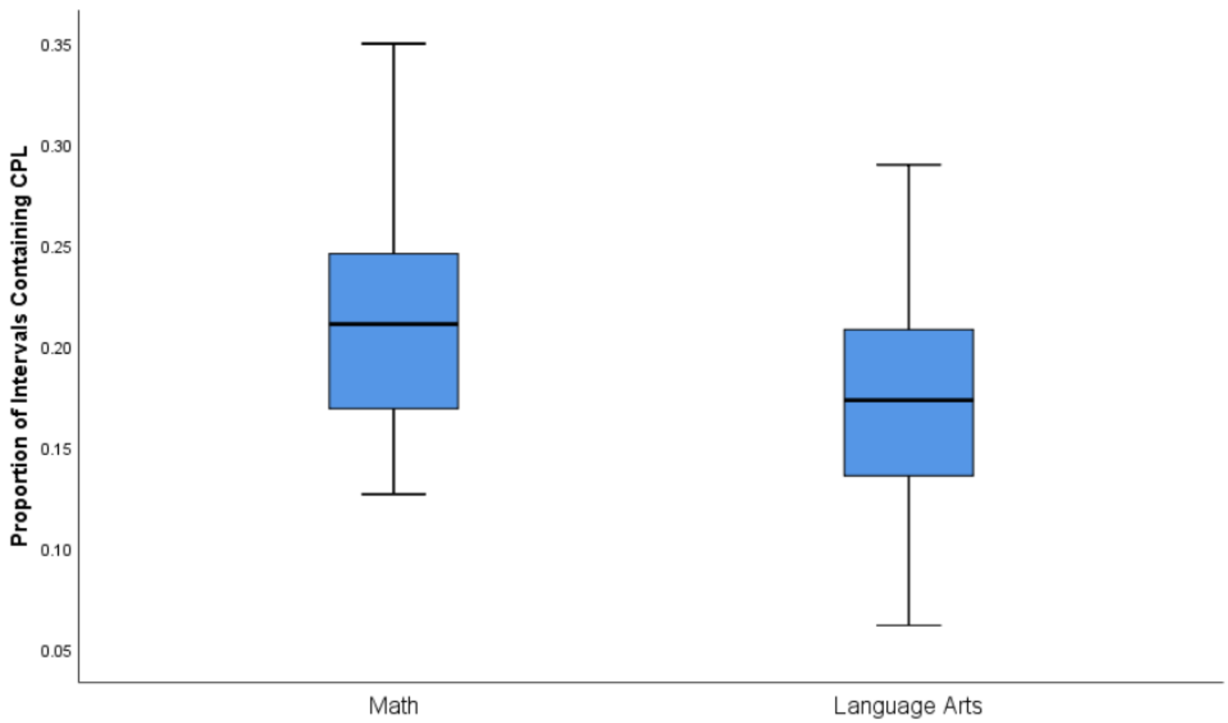


Figure 16: *Provision of CPL during Mathematics and Language Arts Instruction*

Table 18

Summary of Unconditional HLM Analysis for Recall in the Object Memory Task

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	8.11	.13	61.50	53.99	.00	7.84	8.37
Variance estimates							
Level-1	3.85	.29			.00	3.33	4.47
Level-2	.44	.19			.02	.19	1.04

Table 19

Summary of Condition HLM Analysis for Recall in the Object Memory Task with Level-1 Predictors

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	8.09	.16	52.14	39.38	.00	7.78	8.40
Initial performance	.34	.05	6.20	251.25	.00	.23	.44
Working memory	.17	.14	1.25	253.64	.21	-.09	.44
Self-regulation	.04	.02	1.72	251.17	.09	.00	.09
Variance estimates							
Level-1	3.28	.29			.00	2.76	3.90
Level-2	.55	.24			.02	.24	1.30

Table 20

Summary of Conditional HLM Analysis for Recall in the Object Memory Task with Level-1 and Level-2 Predictors

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	8.11	.85	9.57	36.06	.00	6.39	9.82
Initial performance	.35	.06	6.07	236.78	.00	.23	.46
Working memory	.16	.14	1.12	237.81	.26	-.12	.44
Self-regulation	.04	.02	1.96	236.53	.05	.00	.09
CPL	.00	.04	.00	35.87	.99	-.09	.09
Variance estimates							
Level-1	3.26	.30			.00	2.72	3.90
Level-2	.57	.26			.03	.23	1.38

Table 21

Summary of Conditional HLM Analysis for Recall in the Object Memory Task with Cross-level Interaction

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	8.11	.85	9.58	36.05	.00	6.39	9.82
Initial performance	.34	.06	6.03	235.77	.00	.23	.45
Working memory	.16	.14	1.09	236.83	.27	-.13	.44
Self-regulation	.12	.14	.82	235.46	.41	-.17	.41
CPL	0.00	.04	.00	35.86	.99	-.08	.08
CPL*Self-regulation	0.00	.00	-.53	235.45	.59	-.02	.01
Variance estimates							
Level-1	3.26	.30			.00	2.72	3.91
Level-2	.57	.26			.03	.23	1.37

Table 22

Summary of Unconditional HLM Analysis for Strategy Use in the Object Memory Task

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	118.09	1.51	78.07	51.25	.00	115.05	121.13
Variance estimates							
Level-1	729.30	55.35			.00	628.59	846.36
Level-2	25.83	25.83			.32	3.63	183.35

Table 23

Summary of Conditional HLM Analysis for Strategy Use in the Object Memory Task with Level-1 Predictors

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	115.15	1.29	89.10	274	.00	112.61	117.69
Initial performance	.35	.04	8.70	274	.00	.27	.43
Working memory	1.32	1.66	.79	274	.43	-1.95	4.59
Self-regulation	.44	.27	1.61	274	.11	-.10	.98
Variance estimates							
Level-1	463.70	39.62			.00	392.20	548.22
Level-2	0	0			.	.	.

Table 24

Summary of Conditional HLM Analysis for Strategy Use in the Object Memory Task with Level-1 and Level-2 Predictors

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	121.20	7.31	16.58	257	.00	106.81	135.60
Initial performance	.35	.04	8.60	257	.00	.27	.43
Working memory	1.07	1.78	.60	257	.55	-2.43	4.57
Self-regulation	.49	.28	1.73	257	.09	-.07	1.05
CPL	-.31	.37	-.86	257	.40	-1.03	.40
Variance estimates							
Level-1	481.78	42.50				405.28	572.71
Level-2	0	0			.	.	.

Table 25

Summary of Conditional HLM Analysis for Strategy Use in the Object Memory Task with Cross-level Interaction

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	120.56	7.29	16.55	256	.00	106.21	134.91
Initial performance	.35	.04	8.50	256	.00	.27	.43
Working memory	1.28	1.77	.72	256	.47	-2.21	4.78
Self-regulation	-2.83	1.83	-1.55	256	.12	-6.43	.76
CPL	-.28	.36	-.78	256	.43	-1.00	.43
CPL*Self-regulation	.17	.09	1.84	256	.07	-.01	.34
Variance estimates							
Level-1	477.32	42.19			.00	401.39	567.60
Level-2	0	0			.	.	.

Table 26

Summary of Unconditional HLM Analysis for Recall in the Free Recall with Organizational Training Task

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	10.47	.17	60.92	53.90	.000	10.13	10.82
Variance estimates							
Level-1	10.39	.78			.000	8.97	12.02
Level-2	.23	.32			.475	.01	3.58

Table 27

Summary of Condition HLM Analysis for Recall in the Free Recall with Organizational Training Task with Level-1 Predictors

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	10.41	.20	52.31	39.77	.00	10.01	10.81
Initial performance	.37	.08	4.69	258.61	.00	.21	.52
Working memory	.28	.23	1.23	260.68	.22	-.17	.73
Self-regulation	.11	.04	3.13	257.03	.00	.04	.19
Variance estimates							
Level-1	9.14	.81			.00	7.69	10.87
Level-2	.38	.38			.33	.05	2.79

Table 28

Summary of Conditional HLM Analysis for Recall in the Free Recall with Organizational Training Task with Level-1 and Level-2 Predictors

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	9.63	1.11	8.67	39.12	.00	7.39	11.88
Initial performance	.38	.08	4.57	243.92	.00	.21	.54
Working memory	.36	.24	1.52	244.34	.13	-.11	.84
Self-regulation	.12	.04	3.17	242.40	.00	.04	.19
CPL	.04	.06	.79	38.62	.49	-.07	.15
Variance estimates							
Level-1	9.30	.85			.00	7.78	11.12
Level-2	.44	.42			.30	.07	2.91

Table 29

Summary of Conditional HLM Analysis for Recall in the Free Recall with Organizational Training Task with Cross-level Interaction

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	9.64	1.11	8.66	39.14	.00	7.39	11.89
Initial performance	.37	.08	4.44	242.95	.00	.20	.53
Working memory	.35	.24	1.47	243.34	.14	-.12	.82
Self-regulation	.38	.25	1.53	241.12	.13	-.11	.87
CPL	.04	.06	.69	38.70	.49	-.07	.15
CPL*Self-regulation	-.01	.01	-1.06	241.10	.29	-.04	.01
Variance estimates							
Level-1	9.30	.85			.00	7.78	11.11
Level-2	.44	.42			.30	.07	2.90

Table 30

Summary of Unconditional HLM Analysis for Strategy Use in the Free Recall with Training Task

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	.47	.03	13.65	51.71	.00	.40	.54
Variance estimates							
Level-1	.32	.02			.00	.28	.37
Level-2	.02	.01			.10	.01	.07

Table 31

Summary of Condition HLM Analysis for Strategy Use in the Free Recall with Organizational Training Task with Level-1 Predictors

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	.42	.04	10.91	39.76	.00	.34	.49
Initial performance	.45	.09	5.06	255.72	.00	.27	.63
Working memory	.05	.04	1.32	259.23	.19	-.03	.13
Self-regulation	.01	.01	2.09	255.75	.04	.00	.03
Variance estimates							
Level-1	.30	.03			.00	.25	.35
Level-2	.02	.01			.17	.00	.08

Table 32

Summary of Conditional HLM Analysis for Strategy Use in the Free Recall with Organizational Training Task with Level-1 and Level-2 Predictors

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	95% Confidence Interval Upper
Fixed effects							
Intercept	.32	.22	1.46	39.15	.15	-.12	.76
Initial performance	.44	.09	4.82	241.37	.00	.26	.62
Working memory	.07	.05	1.64	242.90	.10	-.01	.15
Self-regulation	.01	.01	1.88	241.22	.06	.00	.03
CPL	.00	.01	.44	38.64	.66	-.02	.03
Variance estimates							
Level-1	.29	.03			.00	.25	.35
Level-2	.02	.02			.12	.01	.09

Table 33

Summary of Conditional HLM Analysis for Strategy Use in the Free Recall with Organizational Training Task with Cross-level Interaction

Parameter	Coefficient	SE	t	df	P	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	.32	.22	1.47	39.15	.15	-.12	.76
Initial performance	.44	.09	4.80	240.38	.00	.26	.62
Working memory	.07	.04	1.61	241.91	.11	-.02	.15
Self-regulation	.03	.04	.71	240.09	.48	-.06	.12
CPL	.00	.01	.44	38.64	.66	-.02	.03
CPL*Self-regulation	.00	.00	-.43	240.08	.67	-.01	.00
Variance estimates							
Level-1	.29	.03			.00	.25	.25
Level-2	.02	.02			.13	.01	.09

Table 34

Summary of Unconditional HLM Analysis for Accuracy in Mathematics

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	8.45	.13	67.15	34.11	.00	8.19	8.71
Variance estimates							
Level-1	4.48	.40			.00	3.77	5.32
Level-2	.04	.17			.80	0.00	86.51

Table 35

Summary of Condition HLM Analysis for Accuracy in Mathematics with Level-1 Predictors

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	8.46	.13	64.55	40.71	.00	8.19	8.73
Initial performance	.54	.04	12.57	256.62	.00	.46	.63
Working memory	.22	.11	1.93	250.80	.05	.00	.44
Self-regulation	.04	.02	2.19	252.20	.03	.00	.08
Variance estimates							
Level-1	1.93	.17			.00	1.61	2.30
Level-2	.46	.17			.00	.22	.94

Table 36

Summary of Conditional HLM Analysis for Accuracy in Mathematics with Level-1 and Level-2 Predictors

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	8.85	.62	14.22	40.15	.00	7.59	10.10
Initial performance	.54	.04	12.57	255.97	.00	.46	.63
Working memory	.22	.11	1.94	250.40	.05	.00	.44
Self-regulation	.04	.02	2.18	251.69	.03	.00	.07
CPL	-.02	.03	-.63	40.98	.53	-.07	.04
Variance estimates							
Level-1	1.93	.17			.00	1.62	2.30
Level-2	.46	.17			.01	.22	.97

Table 37

Summary of Conditional HLM Analysis for Accuracy in Mathematics with Cross-level Interaction

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	95% Confidence Interval Upper
Fixed effects							
Intercept	8.84	.62	14.23	40.14	.00	7.59	10.10
Initial performance	.55	.04	12.55	254.97	.00	.46	.63
Working memory	.22	.11	1.94	249.41	.05	.00	.44
Self-regulation	.00	.08	-.01	249.62	.99	-.15	.15
CPL	-.02	.03	-.63	40.98	.53	-.07	.04
CPL*Self-regulation	.00	.00	.55	249.75	.58	-.00	.01
Variance estimates							
Level-1	1.93	.17			.00	1.62	2.31
Level-2	.46	.17			.01	.22	.96

Table 38

Summary of Unconditional HLM Analysis for Strategy Use in Mathematics

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	.85	.01	60.61	36.30	.00	.82	.87
Variance estimates							
Level-1	.05	.00			.82	.05	.06
Level-2	.00	.00				.00	2.81

Table 39

Summary of Condition HLM Analysis for Strategy Use in Mathematics with Level-1 Predictors

Parameter	Coefficient	SE	t	df	p	95% Confidence Interval Lower	Upper
Fixed effects							
Intercept	.85	.01	58.73	36.88	.00	.82	.88
Initial performance	.14	.06	2.21	248.08	.02	.02	.27
Working memory	.05	.02	2.81	248.58	.00	.02	.08
Self-regulation	.00	.00	.05	246.89	.95	-.01	.01
Variance estimates							
Level-1	.05	.00			.00	.04	.06
Level-2	.00	.00			.54	.00	.03

Table 40

Summary of Conditional HLM Analysis for Strategy Use in Mathematics with Level-1 and Level-2 Predictors

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	.83	.07	12.10	39.35	.00	.69	.97
Initial performance	.15	.06	2.21	247.77	.03	.02	.27
Working memory	.05	.02	2.80	247.25	.01	.02	.09
Self-regulation	.00	.00	.05	246.60	.96	-.01	.02
CPL	.00	.00	.24	41.10	.81	-.01	.01
Variance estimates							
Level-1	.05	.00			.00	.04	.06
Level-2	.00	.00			.50	.00	.03

Table 41

Summary of Conditional HLM Analysis for Strategy Use in Mathematics with Cross-level Interaction

						95% Confidence Interval	
Parameter	Coefficient	SE	t	df	p	Lower	Upper
Fixed effects							
Intercept	.83	.07	12.03	39.36	.00	.69	.97
Initial performance	.14	.06	2.22	246.76	.03	.02	.27
Working memory	.05	.02	2.81	246.28	.01	.02	.09
Self-regulation	.00	.01	-.24	245.22	.81	-.03	.02
CPL	.00	.00	.24	41.12	.81	-.01	.01
CPL*Self-regulation	.00	.00	.26	245.71	.80	.00	.00
Variance estimates							
Level-1	.05	.00			.00	.04	.06
Level-2	.00	.00			.51	.00	.03

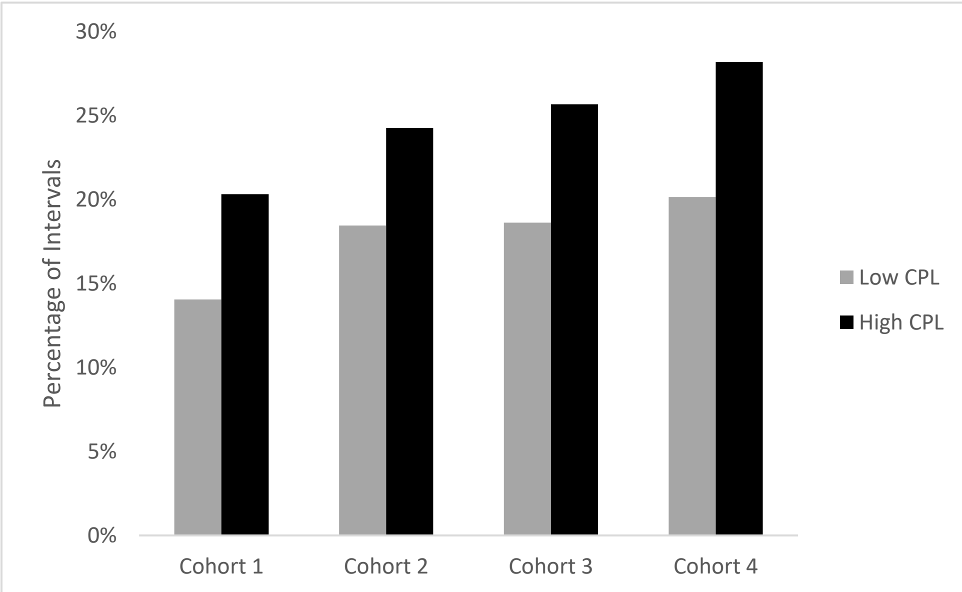


Figure 17: *Percentage of Intervals Containing CPL among "Low" and "High" Teachers across Cohorts*

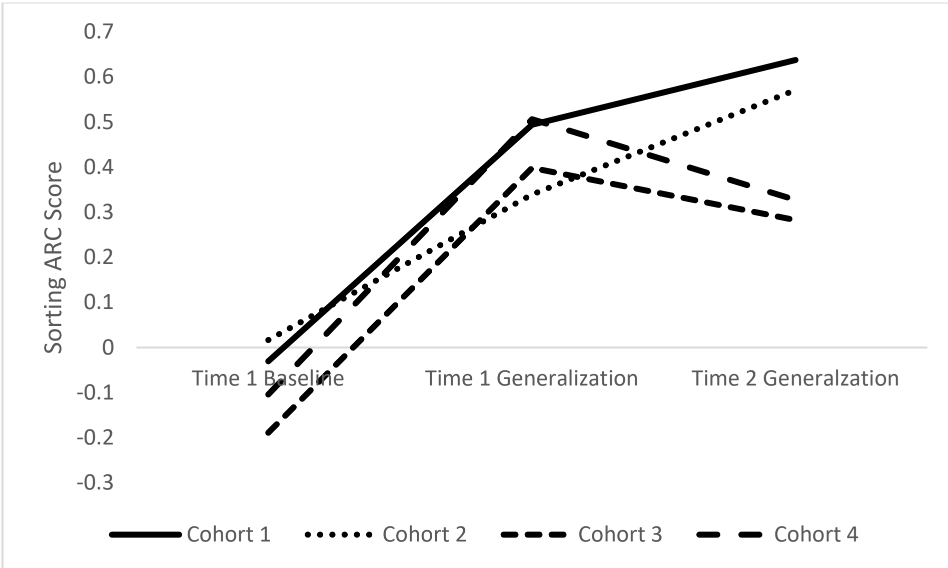


Figure 18: *Children's Sorting ARC Scores across Cohorts*

REFERENCES

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2*, 213-236

Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology* (pp. 302 – 338). New York, NY: Springer.

Baker-Ward, L., Ornstein, P.A., & Holden, D. J. (1984). The expression of memorization in early childhood. *Journal of Experimental Child Psychology, 37*, 555-575.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65-94.

Bassok, D., & Latham S. (2014). *Kids today: Changes in school readiness in an early childhood era*. Retrieved from http://curry.virginia.edu/uploads/resourceLibrary/25_kids_Today.pdf

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open, 1*, 1-31. doi: 10.1177/2332858415616358

Best, D. L., & Ornstein, P. A. (1986). Children's generation and communication of mnemonic organizational strategies. *Developmental Psychology, 22*, 845-853.

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review, 7*, 86-92.

Bjorklund, D. F., Ornstein, P. A., & Haig, J. R. (1977). Developmental differences in Organization and recall: Training in the use of organizational techniques. *Developmental Psychology, 3*, 175-183.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false

belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*, 647–663.Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach.* Hoboken, NJ: Wiley

Borkowski, J. G., Carr, M., & Pressley, M. (1987). 'Spontaneous' strategy use: Perspectives from metacognitive theory. *Intelligence, 11*, 61-75. doi:10.1016/0160-2896(87)90027-4

Brown, A. L. & Reeve, R. A. (1987). Bandwidths of competence: The role of supportive contexts in learning and development. In L. S. Liben & D. H. Feldman (Eds.), *Development and learning: Conflicts or congruence?* Hillsdale, NJ: Lawrence Erlbaum Associates.

Burchinal, M. (2008). Measures of classroom quality in pre-kindergarten and children's development of academic, language, and social skills. *Child Development, 79*, 732–749.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*, 166-176. doi: 10.1016/j.ecresq2009.10.1004

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem solving skills. In T. P. Carpenter, J.M., Moser, & T. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 9 – 24). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively Guided Instruction.* Porstmouth, NH: Heinemann.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (2000). *Cognitively guided instruction: A research-based teacher professional development program for elementary school mathematics*. (Report Number 003). Madison, WI: National Center for Improving Student Learning and Achievement in Mathematics and Science

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Education Research Journal, 26,* 499-531.

Carr, M. (2010). The importance of metacognition for conceptual change and strategy use in Mathematics. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (176-197). New York, NY: The Guildford Press.

Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology, 89,* 318-328.

Carr, M., Kurtz, B. E., Schneider, W., Turner, L. A., & Borkowski, J. G. (1989). Strategy acquisition and transfer among American and German children: Environmental influences on metacognitive development. *Developmental Psychology, 25,* 765-771.

Cavanaugh, J. C., & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development, 53*, 11-28. doi:10.2307/1129635

Cobb, P., Boufi, A., McClain, K. & Whitenack, J. (1998). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education, 28,* 258-277.

Coffman, J. L., Ornstein, P. A., McCall, L. E., & Curran, P. J. (2008). Linking teachers' memory-relevant language and the development of children's memory skills. *Developmental Psychology, 44,* 1640-1654. doi: 10.1037/a0013859

Coffman, J. L., Grammer, J. K., Hudson, K. N., Thomas, T. E., Vilwock, D., & Ornstein, P. A. (2018). Relating children's early elementary classroom experiences to later skilled remembering and study skills. *Journal of Cognition and Development.* Advance online publication. doi: 10.1080/15248372.2018.1470976

Crosnoe, R., Morrison, F., Burchinal, M., Pianta, R., Keating, D., Friedman, S. L., & Clarke-Stewart, K. A., & the NICHD Early Child Care Research Network. (2010). Instruction, teacher-student relations, and math achievement trajectories in elementary school. *Journal of Educational Psychology, 102,* 407-417. doi: 10.1037/a0017762

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38,* 352-362.

Dinsmore, D. L., Alexander, P., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. Educational Psychology Review, 20, 391-409. doi: 10.1007/s10648-008-9083-6

Duckworth, A. L. & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*, 939-944.

Dumontheil, K. T. (2012). Brain activity during a visuospatial working memory task predicts arithmetic performance 2 years later. *Cerebral Cortex, 22*, 1078-1085. https://doi.org/10.1093/cercor/bhr175

Duncan, G. J., Dowsett, C. J., Clasessens, A., Magnuson, K., Huston, A. C., Klebanov, P.,…Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428-1446. DOI: 10.1037/0012-1649.43.6.1428

Engel, M., Classens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis)alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis, 35*, 157-178. doi: 10.3102/0162373712461850

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27,* 403-434.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

Folds, T. H., Footo, M., Guttentag, R.E., & Ornstein, P. A. (1990). When children mean to remember: Issues of context specificity, strategy effectiveness, and intentionality in the development of memory. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 67-91). Hillsdale, NJ: Lawrence Erlbaum Associates.

Franke, M. L., Webb, N. M., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education, 60*, 380-392. doi: 10.1177/0022487109339906

Fuson, K. C. (1982). An analysis of the counting-on solution procedure in addition. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.*), Addition and subtraction: A cognitive perspective* (pp. 67-81). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology, 92*, 377–390.

Ginsburg, H. P. (1982). The development of addition in the contexts of culture, social class, and race. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A developmental perspective,* (pp. 191-210). Hillsdale, NJ: Lawrence Erlbaum Associates

Grammer, J. K., Coffman, J. L., & Ornstein, P. A. (2013). The impact of teachers' memory relevant language on children's strategy use and knowledge. *Child Development, 84*, 1989-2002. doi: 10.1111/cdev.1210

Grammer, J. K., Coffman, J. L., Sidney, P. G., & Ornstein, P. A. (2016). Linking teacher instruction and student achievement in mathematics: The role of teacher language. *Journal of Cognition and Development, 17*, 468-485. doi: 10.1080/15248372.2015.1068777

Grammer, J. K., Purtell, K. M., Coffman, J. L., & Ornstein, P. A. (2011). Relations between children's metamemory and strategic performance: Time-varying covariates in early elementary school. *Journal of Experimental Child Psychology, 108,* 139–155. doi:10.1016/j.jecp.2010.08.001

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329-343.

Guttentag, R. E. (1984). The mental effort requirement of cumulative rehearsal: A developmental study. Journal of Experimental Child Psychology, 37, 92-106.

Harnishfeger, K. K., & Bjorklund, D. F. (1990). Children's strategies: A brief history, Contemporary views of cognitive development. In D. F. Bjorklund (Ed*.), Children's strategies: Contemporary views of cognitive development.* (pp.1-22). Hillsdale, NJ: Lawrence Erlbaum Associates.

Harnishfeger, K. K., Bjorklund, D. F., & Halleck, B. A. (1986, April). Developmental differences in the mental effort requirements for the use of an organizational strategy in free recall. Paper presented at Conference on Human Development, Nashville, TN.

Hudson, K. N., Coffman, J. L., & Ornstein, P.A. (2018). Addition in kindergarten: The role of mothers' and teachers' language. *Journal of Cognition and Development*, *19*, 65-86. https://doi.org/10.1080/15248372.2017.1415900

Jacobs, J. (1887). Experiments on "prehension." *Mind, 12,* 75–79.

Kail, R. V. (2003). Information processing and memory. In M. H. Bornstein, L. Davidson, C. L. M. Keyes, & K. A. Moore (Eds.), *Crosscurrents in contemporary psychology* (pp. 269–279). Mahwah, NJ: Erlbaum.

Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., Ruzek, E. A., & Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development, 84,* 1171–1190.

Klibanoff, R. S., Levine, S., Huttenlocher, J., Hedges, L. V., & Vasilyeva, M. (2006). Preschool children's mathematical knowledge: The effect of teacher 'math talk'. *Developmental Psychology, 42*, 59−69. doi: 10.1037/0012-1649.42.1.59

Kurtz, B. E., Schneider, W., Carr, M., Borkowski, J. G., & Rellinger, E. (1990). Strategy instruction and attributional beliefs in West Germany and the United States: Do teachers foster metacognitive development? *Contemporary Educational Psychology, 15*, 268-283.

Lange, G. (1978). Organization-related processes in children's recall. In P. A. Ornstein (Ed.), *Memory development in children* (pp. 101–128). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lehman, M., & Hasselhorn, M. (2007). Variable memory strategy use in children's adaptive intratask learning behavior: Developmental changes and working memory influences in free recall. *Child Development, 78*, 1068-1082. doi: 10.1111/j.1467-8624.2007.01053.x

Lehmann, M., & Hasselhorn, M. (2010). The dynamics of free recall and their relationship to rehearsal between 8-and 10-years of age. *Child Development. 81*, 1006– 1020. doi: 10.1111/j.1467-8624.2010.01448.x

Lehmann, M., & Hasselhorn, M. (2012). Rehearsal dynamics in elementary school children. *Journal of Experimental Child Psychology, 111*, 552–560. doi: 10.1016/j.jecp.2011.10.013

Liberty, C. & Ornstein, P. A. (1973). Age differences in organization and recall: The effects of training in categorization. *Journal of Experimental Child Psychology, 15*, 169-186.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic language and social skills. *Child Development, 3*, 732-749. doi: 10.1111/j.1467-8624.2008.01154.x

McCarthy, D. A. (1972). *Manual for the McCarthy scales of children's abilities.* New York: Psychological Corp.

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21*, 471–490.

McClelland, M. M. & Cameron, C. E. (2011). Self-regulation and academic achievement in elementary school children. In. R. M. Lerner, J. V. Lerner, E. P. Bowers, S. Lewin-Bizan, S. Gestdottir, & J. B. Urban (Eds.), Thriving in childhood and adolescence: The role of self-regulation processes. *New Directions for Child and Adolescent Development*, 133, 2-44.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*, 947-959.

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*, 359-378. https://doi.org/10.1016/S0959-4752(96)00021-7

Moely, B. E., Hart, S. S., Leal, L., Santulli, K. A., Rao, N., Johnson, T., & Hamilton, L.B. (1992). The teacher's role in facilitating memory and study strategy development in the elementary school classroom. *Child Development, 63*, 653-672.

Morrison, F. J., Ponitz, C. C., & McClelland, M. M. (2010). Self-regulation and academic achievement in the transition to school. In S.D. Calkins & M. Bell (Eds.), *Child development at the intersection of emotion and cognition.* (pp. 203–224):

Washington, DC: American Psychological Association.

Morrison, F. J., Smith, L., & Dow-Ehrensberger, M. (1995). Education and cognitive development: A natural experiment. *Developmental Psychology, 31*, 789-799.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards: About the Standards*. Retrieved from http://www.corestandards.org/about-the-standards

Ornstein, P.A., Baker-Ward, L., & Naus, M. J. (1988). The development of mnemonic skill. In F. E. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 31-49). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ornstein, P.A., Coffman, J. L., Grammer, J. K., McCall, L. E., & San Souci, P. P. (2010). Linking the classroom context and the development of children's memory skills. In J. Meece & J. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 42-59). New York: Routledge. doi: 10.1037/a0013859

Ornstein, P. A., & Corsale, K. (1979). Organizational factors in children's memory. In C. R. Puff (Ed.), *Memory structure and organization* (pp. 219-257). New York, NY: Academic Press.

Ornstein, P. A., Haden, C. A., & Elischberger, H. B. (2006). Children's memory development: Remembering the past and preparing for the future. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 143-161). New York, NY: Oxford University Press.

Ornstein, P. A., Medlin, R. G., Stone, B. P., & Naus, M. J. (1985). Retrieving for rehearsal: An analysis of active rehearsal in children's memory. *Developmental Psychology, 21,* 633-641.

Ornstein, P. A., & Naus, M. J. (1978). Rehearsal processes in children's memory. In P. A. Ornstein (Ed.), *Memory development in children*. (pp. 69-99). Hillsdale, NJ: Erlbaum.

Ornstein, P. A., Naus, M. J., & Liberty, C. (1975). Rehearsal and organizational processes in children's memory. *Child Development, 46,* 818-830. doi: 10.2307/1128385

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *The Classroom Assessment Scoring System Manual, K-3*. Baltimore, MD: Brookes.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, *41*, 219-225

Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, *23*, 141-158. doi:10.1037/t14094-000

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45,* 605-619.

Pressley, M., & Afflerbach, P. (1995). Verbal protocols of reading: The nature of constructively responsive reading. Hillsdale, NJ: Lawrence Erlbaum.

Pressley, M. P., & Van Meter, P. (1993). Memory strategies: Natural development and use following instruction. In R. Pasnak & M. L. Howe (Eds.), *Emerging themes in cognitive development* (pp. 128-165). New York, NY: Springer.

Ramirez, G., Gunderson, E. A., Levine, S. C., Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development, 14*, 187-202. doi: 10.1080/15248372.2012.664593

Rogoff, B. (1981). Schooling's influence on memory test performance. *Child Development, 52*, 260-267.

Rohwer, W. D. (1973). Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 8, pp. 1–57). New York: Academic Press.

Rothbart, M. K., & Posner, M. I. (2005). Genes and experience in the development of executive attention and effortful control. In L. A. Jensen & R. W. Larson (Eds.), *New horizons in developmental theory and research* (pp. 101–108). San Francisco, CA: Jossey-Bass.

Rudd, L. C., Lambert, M., Satterwhite, M., & Zaier, A. (2008). Mathematical language in early childhood settings: What really counts? *Early Childhood Education Journal, 36*, 75-80.

Schlagmüller, M., & Schneider, W. (2002). The development of organizational strategies in children: Evidence from a microgenetic longitudinal study. *Journal of Experimental Child Psychology, 81*, 298–319. DOI: 10.1006/jecp.2002.2655

Schneider, W. (1985) Developmental trends in the metamemory-memory behavior relationship: An integrative review. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance* (Vol. 1, pp. 57-109). Orlando, FL: Academic Press.

Schneider, W., Borkowski, J. G., Kurtz, B. E., & Kerwin, K. (1986). Metamemory and motivation: A comparison of strategy use and performance in German and American children. *Journal of Cross Cultural Psychology, 17*, 315-336

Schneider, W., Kron, V., Hünnerkopf, M., & Krajewski, K. (2004). The development of young children's memory strategies: First findings from the Würzburg Longitudinal Memory Study. *Journal of Experimental Child Psychology, 88*, 193-209. doi:0.1016/j.jecp.2004.02.004

Schneider, W., & Löffler, E. (2016). The development of metacognitive knowledge in children and adolescents. In J. Dunlosky & S. K. Tauber (Eds.), *Oxford library of psychology. The Oxford handbook of metamemory* (pp. 491-518). New York, NY, US: Oxford University Press.

Schneider, W. & Pressley, M. (1997). *Memory Development between Two and Twenty*. Mahwah, NJ: Lawrence Erlbaum Associates

Schneider, W. & Sodian, B. (1997). Memory strategy development: Lessons from longitudinal research. *Developmental Review, 17*, 442-461.

Schneider, W., & Weinert, F. E. (1995). Memory development during early and middle childhood: Findings from the Munich Longitudinal Study (LOGIC). In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 263-279). Mahwah, NH: Lawrence Erlbaum Associates.

Schunk, D. H., & Zimmerman, B. J. (1998). *Self-regulated learning: From teaching to self-reflective practice.* New York, NY, US: Guilford Publications.

Scribner, S., & Cole, M. (1978). Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review, 48*, 448-461.

Sharp, D., Cole, M., & Lave, C. (1979). Education and cognitive development: The evidence from experimental research. *Monographs of the Society for Research in Child Development, 4*, 1-112.

Siegler, R. S. (1987). The perils of average data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*, 250-264

Siegler, R. S. (1989). Mechanisms of cognitive development. *Annual Review of Psychology, 40,* 353-379.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York, NY: Oxford University Press.

Siegler, R. S. (2003). Implications of cognitive science research for mathematics education. In K. J. Kilpatrick, W. G., Martin, & D. E., Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 119-233). Reston, VA: National Council of Teachers of Mathematics

Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist, 46*, 606-620. doi:10.1037/0003-066X.46.6.606

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Robinson, M. (1982). The development of numerical understandings. In H. W. Reese & L. P. Lipsett (Eds.), *Advances in child development and behavior:* (Vol. 16 pp. 242-213). New York: Academic Press.

Siegler, R. S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229-293). Hillsdale, NJ: Lawrence Erlbaum. Associates.

Smedt, B. D., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquiere, 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,* 186-201. doi: 10.1016/j.jecp.2009.01.004

Sodian, B., & Schneider, W. (1999). Memory strategy development: Gradual increase, sudden insight, or roller-coaster? In F. E. Weinert & W. Schneider (Eds.), *Individual development from 3 to 12: Findings from the Munich Longitudinal Study* (pp. 61-77). Cambridge, UK: Cambridge University Press.

Throndsen, I. (2011). Self-regulated learning of basic arithmetic skills: A longitudinal study. *British Journal of Educational Psychology, 81*, 558-578. doi: 10.1348/2044-8279.002008

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, (1), 3 – 14. doi:10.1007/s11409-006-6893-0

Vygotsky, L. (1978). *Mind in society.* Cambridge, UK: Cambridge University Press.

Wagner, D. A. (1978). Memories of Morocco: The influence of age, schooling, and environment on memory. *Cognitive Development, 45,* 389-396.

Webb, N. M., Franke, M. L., Ing, M., Chan, A., De, T., Freund, D., & Battey, D. (2008). The role of teacher instructional practices in student collaboration. *Contemporary Educational Psychology, 33,* 360-381.

Webb, N. M., Franke, M. L., Ing, M., Wong, J., Fernandez, C. H., Shin, N., & Turrou, A. C. (2013). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research, 63,* 79-93. doi: 10.1016/j.ijer.2013.02.001.

Weinert, F. E., & Schneider, W. (Eds.). (1999). *Individual development from 3 to 12: Findings from the Munich Longitudinal Study*. Cambridge: Cambridge University Press.

Wellman, H. M. (1977). Preschoolers' understanding of memory-relevant variables. *Child Development*, *48*, 1720-1723. doi:10.2307/1128544

Wellman, H. M. (1978). Knowledge of the interaction of memory variables: A developmental study of metamemory. *Developmental Psychology, 14*, 24-29. doi:10.1037/0012-1649.14.1.24

Wood, D., Bruner, J. S., Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17,* 89-100. doi: 10.1111/j.1469-7610.1976.tb00381.x

Woodcock, R.W., McGrew, K., & Mather, N. (2001). *The Woodcock-Johnson Tests of Achievement: Third Edition.* Itasca, IL: Riverside Publishing.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81,* 329-339.