



University of
Massachusetts
Amherst

The Effects of Tools of the Mind on Math and Reading Scores In Kindergarten

Item Type	Dissertation (Open Access)
Authors	Mackay, Patricia Estrela
DOI	10.7275/2wyc-1g11
Download date	2026-03-05 12:10:43
Link to Item	https://hdl.handle.net/20.500.14394/39265

THE EFFECTS OF TOOLS OF THE MIND
ON MATH AND READING SCORES IN KINDERGARTEN

A Dissertation Presented

by

PATRICIA E. MACKAY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF EDUCATION

September 2013

Education

Children, Families, and Schools

© Copyright by Patricia E. Mackay 2013
All Rights Reserved

THE EFFECTS OF TOOLS OF THE MIND
ON MATH AND READING SCORES IN KINDERGARTEN

A Dissertation Presented

by

PATRICIA E. MACKAY

Approved as to style and content by:

J. Kevin Nugent, Chair

Jennifer T. Randall, Member

Matthew C. Davidson, Member

Christine B. McCormick, Dean
School of Education

DEDICATION

To my husband Steve Mackay and our boys, Michael and Matthew.

This accomplishment has not been without sacrifice for us all.
Thank you for your unwavering support.

ACKNOWLEDGMENTS

Over the past four years, I have received support and encouragement from many. I am truly thankful to Dr. Kevin Nugent, my adviser and dissertation chair, for your guidance and support throughout this process. I want to thank Dr. Matthew Davidson for your guidance and time as a Committee Member. My Dissertation would not have been possible without the help of Dr. Jennifer Randall. Your guidance has helped me see this idea from an initial proposal four years ago in Educational Research Methods to completion of this Dissertation. Many times, I doubted by ability to see this work to completion, and the guidance of my committee members helped me to accomplish this goal. I would also like to express gratitude to the district that participated in this research. Thank you to colleagues and administrators for your availability, your enthusiasm, and assistance with gathering necessary data. Finally, I would like to thank my family, parents, brothers, and extended family. These four years have been challenging and you always supported me. In particular, I thank my mom, for being a role model of educational excellence. You taught me to set high goals and to accomplish them.

ABSTRACT

THE EFFECTS OF TOOLS OF THE MIND ON MATH AND READING SCORES IN KINDERGARTEN

SEPTEMBER 2013

PATRICIA E. MACKAY, B.A., COLLEGE OF THE HOLY CROSS

M.S.W., BOSTON COLLEGE

Ed.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Emeritus J. Kevin Nugent

Although a limited body of research has supported the positive impact of the Tools of the Mind curriculum on the development of self-regulation, research supporting a direct relationship between Tools and academic achievement is extremely limited. The purpose of this study is to evaluate the effectiveness of the Tools of the Mind curriculum implementation in improving math and reading scores in Kindergarten by comparing scores obtained before and after Tools. This study also seeks to investigate the effects of SES on student achievement. Finally, this study seeks to identify contributions and challenges perceived by teachers during implementation. Participants included 93 students in the before Tools condition and 97 students after Tools. Students who had Tools scored statistically significantly lower on reading scores than students who did not have Tools. While students also scored lower on math after Tools, this difference was not significant. Differences were found in student scores based on SES. Qualitative results are based on interviews of six Kindergarten teachers, and revealed teachers' experiences with implementation. The findings of this study are intended to increase the understanding of the effectiveness of Tools and its implementation.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	v
ABSTRACT.....	vi
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER	
1. INTRODUCTION	1
Theoretical Framework.....	2
Constructivism.....	2
Sociocultural Perspective.....	3
Ecological Systems Theories	4
Dynamic Systems Theories.....	5
Self-Regulation	7
Tools of the Mind.....	11
Historical Review of Tools of the Mind.....	12
Tools of the Mind and Self-Regulation.....	15
Components of Tools	16
Alignment with Developmentally Appropriate Practice.....	23
Limitations of Tools Research	23
Summary.....	24
Statement of the Problem.....	25
Conceptual Framework.....	26
Purpose Statement.....	27
2. LITERATURE REVIEW	28
Executive Function and Self-Regulation	28
Self-Regulation/Executive Function and Academic Achievement	29
Executive Function and Math Achievement	53

Components of Tools of the Mind	77
Private Speech.....	78
Dramatic Play.....	96
Tools of the Mind Studies	107
Pilot Study.....	116
Summary	118
3. METHOD	119
Research Questions.....	119
Reading	119
Math	119
Qualitative.....	120
Participants and Sampling Procedures.....	120
Sample Size, Power, and Precision.....	122
Procedure	122
Research Design (Quantitative)	124
Reading Achievement.....	124
Math Achievement.....	125
Measures	125
Letter Naming Fluency	126
Measures of Academic Progress.....	126
Fidelity of Implementation Scale.....	127
Socioeconomic Status	128
Research Design (Qualitative)	129
Researcher Profile.....	129
Data Collection	129
Data Analysis.....	129
Trustworthiness.....	130
Summary.....	131
4. RESULTS.....	132

Quantitative Results	132
Results for Research Question #1	132
Results for Research Question #2	136
Results for Research Question #3	137
Results for Research Question #4	137
Results for Research Question #5	140
Results for Research Question #6	141
Qualitative Results	142
Results for Research Question #7	142
Challenges Associated with Tools Implementation	142
T1’s Perceptions of Challenges	143
T2’s Perceptions of Challenges	144
T3’s Perceptions of Challenges	146
T4’s Perceptions of Challenges	147
T5’s Perceptions of Challenges	148
Results for Research Question #8	150
Perceived Contributions of Tools	150
T1’s Perceptions of Contributions	150
T2’s Perceptions of Contributions	151
T3’s Perceptions of Contributions	151
T4’s Perceptions of Contributions	151
T5’s Perceptions of Contributions	151
Summary	152
5. DISCUSSION	153
Findings and Interpretations	153
Impact of Tools on Achievement	153
Impact of SES on Achievement	156
Impact of Fidelity on Achievement	156
Challenges Associated with Tools	157
Perceived Contributions of Tools	160
Strengths of the Study	162

Limitations	163
Design	163
Measures	164
Setting and Participants.....	166
Recommendations.....	167
Implications for Further Research	168
Conclusion	169
APPENDICES	
A. CONSENT FORM FOR PARTICIPATION IN A RESEARCH STUDY	171
B. POWER ANALYSES.....	174
C. TOOLS-K PROJECT FIDELITY OF IMPLEMENTATION SCALE	176
D. INTRACLASS CORRELATION COEFFICIENT	177
E. INTERVIEW QUESTIONS FOR TEACHERS QUESTIONS 6 & 7	178
F. ANCOVA OUTPUTS	179
G. ANOVA OUTPUTS	186
H. FIDELITY CORRELATIONS.....	189
REFERENCES	190

LIST OF TABLES

Table	Page
1.1 Tools Activities and Self-Regulation.....	22
3.1 Sample Demographics	123
4.1 Reading Achievement Scores by YEAR and SES	134
4.2 Math Achievement Scores by YEAR and SES	139
4.3 Summary of Qualitative Results	150
F.1 Full ANCOVA Output Table	182
F.2 ANCOVA Homogeneity of Variance Assumption: Levene’s Test	182
F.3 ANCOVA Normality Assumption: Shapiro-Wilk.....	183
F.4 ANCOVA Normality Assumption: Skewness & Kurtosis	184
F.5 ANCOVA Independence of the Covariate and the Independent Variable Assumption for SES.....	184
F.6 ANCOVA Independence of the Covariate and the Independent Variable Assumption for Tools Year.....	185
G.1 Full ANOVA Output Table.....	186
G.2 ANOVA Normality Assumption: Shapiro-Wilk	186
G.3 ANOVA Normality Assumption: Skewness & Kurtosis.....	187
G.4 ANOVA Homogeneity of Variance Assumption: Levene’s Test.....	187
H.1 Correlation between Fidelity and MAP Reading Scores	189
H.2 Correlation between Fidelity and MAP Math Scores	189

LIST OF FIGURES

Figure	Page
1.1 Conceptual Framework.....	27
4.1 MAP Reading Scores (means) Across Tools Conditions.....	135
4.2 MAP Reading Scores by SES and Year	137
4.3 MAP Math Scores by Tools Conditions.....	140
4.4 MAP Math Scores by Tools Conditions and SES.....	141
B.1 Power Analysis for .40 effect size.....	174
B.2 Power Analysis for .25 effect size.....	175
D.1 Intraclass Correlation Coefficient	177
F.1 ANCOVA Linearity Assumption Scatterplot-Overall.....	179
F.2 ANCOVA Linearity Assumption Scatterplot-by Tools Year 2009.....	180
F.3 ANCOVA Linearity Assumption Scatterplot-by Tools Year 2011	180
F.4 ANCOVA Linearity Assumption Scatterplot-by SES Free/Reduced Lunch	181
F.5 ANCOVA Linearity Assumption Scatterplot-by SES Not Free/Reduced Lunch	181
F.6 ANCOVA Homogeneity of Variances Scatterplot	183
G.1 ANOVA Homogeneity of Variances Scatterplot.....	188

CHAPTER 1

INTRODUCTION

Educational researchers have long debated the effectiveness of teaching methods and philosophies. Included in this debate has been the comparison between constructivist and highly guided methods of instruction (Kirschner, Sweller, & Clark, 2006). This paper will describe a constructivist approach to early childhood education, namely that of an innovative curriculum called *Tools of the Mind* (Bodrova & Leong, 1996). Included in this description will be a theoretical framework based on the scholarship of Russian psychologist Lev Vygotsky (1978), on which the curriculum was based. Because this curriculum is proposed to foster the development of self-regulation, an overview of self-regulation and its relation to achievement will follow. The goal of this study is to investigate the relationship between the development of self-regulation through participation in the Tools curriculum and academic achievement in kindergarten students. A comparison between classrooms using Tools of the Mind and classrooms using other more traditional, teacher-directed methods will facilitate this investigation and add empirical evidence in the debate about teacher philosophies.

This debate has coincided with changes in educational practices that have resulted in increasing academic demands in kindergarten (Au, 2005; Curwood, 2008). In the past, children in kindergarten were able to learn through play, exploration and their imaginations. Today, kindergarteners spend large amounts of time listening to teacher directed instruction and on testing of literacy and math skills resulting in what Miller and Almon (2009) call “crisis in the kindergarten” (p. 42). Some argue, however, that methods that rely only on implicit or incidental instruction might not be sufficient for all students to meet current expectations (Phillips, Clancy-Menchetti, & Lonigan, 2008).

These greater demands on children in kindergarten may not be conducive to developmentally appropriate classroom practices.

The National Association for the Education of Young Children (NAEYC) has defined “predominantly teacher-directed tasks, highly structured classes, large group work, paper/pencil tasks, rote learning, direct teaching of discrete skills, punishment, extrinsic rewards, and standardized assessments” as developmentally inappropriate (Bredekamp & Copple, 1997). Despite these practices being assessed as developmentally inappropriate, the push for early literacy and math skills by *No Child Left Behind* has caused schools to shift to these very practices (Jarrett & Waite-Stupiansky, 2009). Many schools use core programs that are teacher directed and scripted, leaving little room for learning through exploration and play.

Theoretical Framework

Constructivism

The lack of agreement about educational and learning philosophies is reflected in these changes in educational practices. In classrooms that contain highly guided instruction, students receive information from the teacher through lectures, demonstrations, seatwork, practice, drills, and testing (Gersten, Woodward, & Darch, 1986). The teacher is seen as the transmitter of knowledge. In contrast, constructivist instructional methods view meaning as created rather than acquired (Ertmer & Newby, 1993). Primarily based on the work of Dewey (1938) and Piaget (1952), constructivist education assumes that humans construct their own knowledge and that learning is an active process. Dewey (1938) contributed to the development of constructivism by rejecting authoritarian teaching techniques and suggesting that learning be grounded in

real experiences for inquiry. Piaget (1952) added that information should be presented to assist in problem-solving, so that experience, making mistakes, and finding solutions can enable assimilation and accommodation.

Contemporaries of Piaget and Dewey built upon these original ideas and added to the understanding of constructivism. Bruner (1960) proposed that learning should be authentic and not centered around teaching isolated skills. Bruner also believed that teaching activities should allow students to discover and construct knowledge. Another contemporary of constructivism was Lev Vygotsky (1978), a Russian psychologist. Vygotsky added the social dimension to the constructivist paradigm. Vygotsky believed that the social environment can help a child's cognitive development.

Sociocultural Perspective

Vygotsky (1978) believed that learning is largely mediated by social interaction of students and "More Knowledgeable Others," (MKO), such as teachers, parents, and even peers. Through the MKO and peers, collaboration and shared learning occurs, allowing for co-construction (Vygotsky, 1978). Vygotsky also coined the phenomenon known as "Zone of Proximal Development," or ZPD, which has been defined as "*the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers*" (Vygotsky, 1978, p. 86).

ZPD is based on the assumption that a child is only able to take the next step in their cognitive development if another person (MKO) supports them. This support has been called scaffolding, a term introduced by Wood, Bruner, and Ross (1976). Scaffolding includes the gradual withdrawal of support as the child increases his or her

capacity (Berk & Winsler, 1995). Rogoff (1990) expanded thinking on scaffolding and apprenticeship, and believed development to be an apprenticeship in which children learn the use of intellectual tools in structured activities with other adults and children. Elaborating on the concepts of scaffolding and apprenticeship, Rogoff (1995) introduced the concept of “guided participation.” Based on Vygotsky’s premise that development occurs through interactions with MKO’s who mediate intellectual activity, guided participation includes the direction provided by the MKO and observations and active involvement of the child in the activity. The concept of guided participation elevate the understanding of how children learn beyond how adults teach children or how reality is constructed by children.

These practices grounded in constructivist and socio-cultural theories have influenced the understanding of child development as well as the development of new educational programs. One such program is the Tools of the Mind curriculum (Bodrova & Leong, 1996). In contrast to the current trends of applying highly guided instruction in kindergarten classrooms, Tools of the Mind is an early childhood curriculum that is based on a constructivist view of learning. Based on the works of Vygotsky (1978), Tools operates under the assumptions that development and learning occur within the zone of proximal development and that learning is developed through the use of mediators, play, language, shared activities, and the social context. The practices and techniques utilized in Tools align with tenets of Ecological and Dynamic Systems Theories.

Ecological Systems Theories

The inclusion of the social context as a critical component to child development articulated in the Tools model aligns with the Ecological Systems Theory

(Bronfenbrenner, 1989). Bodrova and Leong (2007) acknowledged this theory and how it validated the importance of the social context. Shifting from the previous focus on the influence of nature in developmental studies, Bronfenbrenner's (1989) ecological model emphasized environmental subsystems that influence, or nurture, a child's development. These systems included family, peers, school, and neighborhood; links between home, school, and the neighborhood; settings that do not include the child, but affect the child, such as government or parent's workplace; dominant attitudes of the child's culture; and changes in persons or the environment over time (Bronfenbrenner, 1989). The model was adapted in 1995 to include basic biological contexts of the child. Adding a specific focus to the ecological factors that contribute to school transition, Rimm-Kaufman and Pianta, 2000 extended Ecological Systems Theory in their Ecological and Dynamic Model of Transition. This model added a specific emphasis on the importance of relationships, primarily at home and school, that are necessary for adjustment to school. This perspective will be an important one to consider in the evaluation of the Tools curriculum.

Dynamic Systems Theories

Technological advances in neuroscience in the new millennium resulted in the reconsideration of the influences of nature and nurture in developmental studies (Sameroff, 2010). In this dialectical process, "there is a *unity of opposites* in that development will not occur without both, and there is an *interpenetration of opposites* in that one's nature changes one's nurture and conversely one's nurture changes one's nature" (Sameroff, 2010, p. 9). This perspective, that nature and nurture cannot exist without each other, guided the development of a more contemporary, unified

understanding of development. According to Sameroff, four models are necessary for understanding human development.

The four models include personal change, contextual, regulation, and representational. The personal change model relates to understanding how children change over time. Contextual models, such as Bronfenbrenner's, focus on the impact of the various contextual influences on development. Regulation models express the dynamic process of experiences in development. Regulation models view the child as actively or dynamically participating in their development, rather than being passive recipients of experience. Based on biological processes, this model does not focus solely on self-regulation, but also includes other-regulation. "The balance between other-regulation and self-regulation shifts as the child is able to take on more and more responsibility for his or her own well-being" (Sameroff, 2010, p. 15). Contained in the regulation model is an understanding of other-regulation, called transactional other regulation, which is comparable to Vygotsky's (1978) zone of proximal development (Sameroff, 2010, p. 16). In cognitive development, each component of experience is a step mediated by others toward a higher level. Finally, in the representational model, experiences are encoded as cognitive representations, enabling the child to create order and find meaning in experiences. The combination of these models has resulted in a refined understanding of human development.

Earlier ecological models assumed that various environmental factors impacted the development of the child, but may have placed insufficient attention on the active role of the child and on biological factors (Thelan & Smith, 1998). As Smith and Thelan stated, "it is important to understand the processes by which the everyday activities of

children create developmental change – both the universal attainments and the individual pathways” (2003, p. 347). Later models, including Dynamic Systems Theory (Thelan & Smith, 1994), acknowledge the active, dynamic contributions that each child contributes to his or her own development. Dynamic Systems Theory proposed that a collaboration between multiple characteristics of a person and contexts produce behavior (Rose & Fisher, in press). The latter models also account for the biological factors involved in human development. Brain-based studies have contributed to a deeper understanding of the functions within the domain of neuroscience (Blair & Diamond, 2008; Shonkoff & Phillips, 2000). Studies in neuroscience have enabled an understanding that biological processes affect the capacity for self-regulation. Self-regulation is now understood as emerging out of a balance between biological and social influences.

Self-Regulation

For Vygotsky (1978), self-regulation included deliberate, intentional behaviors that children can acquire in their development of higher mental functions. Self-regulation signified a move from other- to self-regulation. In Vygotsky’s view, play, language, and in particular, dialogue were critical for the development of self-regulation. Additionally, Vygotsky (1978) and Luria (1979) made connections between brain and behavior. Contemporary theorists have added to the definition of self-regulation (Blair & Diamond, 2008; Pintrich, 2000; Schunk & Zimmerman, 2007). The process of self-regulation is defined by Pintrich (2000) as learning which involves setting goals by learners who “then attempt to monitor, regulate, and control their cognition, motivation, and behavior” (p. 453). Schunk and Zimmerman (2007) define self-regulation as “self-generated thoughts, feelings, and actions that are systematically designed to affect one’s learning of

knowledge and skill” (p. 8). Similarly, Blair and Diamond (2008) propose that balance between emotional arousal and cognitive regulation results in self-regulation. A closer examination of the neurobiology is necessary to understand these connections between brain and behavior.

Much debate has occurred regarding the specific areas of the brain responsible for executive functions. Scientists have based their arguments on scientific advances involved in assessing brain activity, such as functional neuroimaging (Gunnar & Quevado, 2007). These advances have enabled scientists to identify the pre-frontal cortex as primarily responsible executive function, as well as other areas of the brain (Blair & Diamond, 2008; Luria, 1979). Verbal, spatial, and object-processing regions of the brain are localized in the pre-frontal cortex. Emotion and working memory interact with the pre-frontal cortex, but distributed throughout the brain are integrations of cortical and sub-cortical systems (O’Hearn, Asato, Ordaz, & Luna, 2008).

Gunnar and Quevado (2007) suggested that cortico-limbic pathways provide that neural surface for emotion, emotional learning, motivation, and regulation. Gunnar and Quevado (2007) further explained that regulation of the sympathetic-adrenomedullary and hypothalamic-pituitary-adrenocortical systems comes together at the hypothalamus, an area of the brain involved in the integration of autonomic and endocrine functions with behavior. Posner and Rothbart (2007) provided a map of the brain relating specific forms of attention with areas of the brain. Alerting, for example, was associated with the thalamus, the posterior area, and the frontal area. Orienting was associated with the superior parietal lobe, the frontal eye field, the pulvinar, and the superior colliculus. Finally, Posner and Rothbart (2007) associated the prefrontal cortex and the anterior

cingulate gyrus with executive attention (p. 6). Other regions of the brain have also been associated with regulatory functions. According to Collette, Hodge, Salmon, & Van Der Linden (2006), the amygdala, for example, is involved in the processing of emotions, the storage of memories, and arousal; the limbic system is associated with emotion and memory; and the hypothalamus is associated with endocrine and autonomic functions. Academic development is a good example of the integration of these neural systems. Houdé, Rossi, Lubin, and Joliot (2010) assert that while the pre-frontal cortex and later the parietal cortex are engaged primarily by children to solve numerical tasks, children later engage the frontal, temporo-parietal, and occipito-temporal regions during reading. Thus, brain studies have contributed greatly to the understanding of the associations between executive functions, regulation, and learning.

The literature indicates that self-regulation and executive function are often used interchangeably. Singer & Bashir (1999) considered how aspects of self-regulation and executive function overlap, and yet function separately. According to Barkley, Murphy, and Fisher (2008), executive function refers to functions of the brain that work as the central processing center of activation, organization, integration, and management of other brain functions. Included in Barkley et al.'s (2008) definition of executive function are nonverbal working memory, internalization of speech (verbal working memory), self-regulation of emotions, and reconstitution (planning and generativity). Brown (2005) also included the executive functions of organizing, prioritizing, and activating for tasks, focusing, sustaining, and shifting attention to task, regulating alertness, sustaining effort and processing speed, managing frustrations and modulating emotions, utilizing working memory and accessing recall, and monitoring and self-regulation action (p. 20-58).

Zimmerman (1989) synthesized these aspects of self-regulation and executive function by stating that “students can be described as self-regulated to the degree that they are metacognitively, motivationally, and behaviorally active participants in their own learning process” (p. 329).

In the following literature review, several executive functions were found most frequently examined: working memory, shifting, and inhibition. Linked to the pre-frontal cortex (Luria, 1973), “these cognitive processes include the maintenance of working memory, in inhibition of proponent responding, and the appropriate shifting and sustaining of attention for the purposes of goal-directed action” (Blair, Zelazo, & Greenberg, 2005, p. 561). Working memory refers to the capacity for an individual to retrieve information from their long-term memory, manipulate, and adapt the information (Welsh, Nix, Blair, Bierman, & Nelson, 2010). Shifting refers to the cognitive flexibility required to transfer attention between different demands and operations of tasks (Diamond, 2006). Barkley (1997b) proposed that inhibition is comprised of the following three interrelated processes: 1) inhibition of a dominant response, 2) stopping of an ongoing response, and 3) interference control (distractibility).

In their hallmark book, *From Neurons to Neighborhoods*, Shonkoff and Phillips (2000) highlighted the period from birth to age five as highly sensitive for brain development. These authors articulated how early experiences affect brain development. Later, in 2011, The Center for the Developing Child at Harvard University published a working paper entitled, “Building the Brain’s ‘Air Traffic Control’ System: How Early Experiences Shape the Development of Executive Functions.” In this metaphor, just as an air traffic control system manages the busy airport and runways,

“executive function refers to a group of skills that helps us to focus on multiple streams of information at the same time, monitor errors, make decisions in light of available information, revise plans as necessary, and resist the urge to let frustration lead to hasty actions” (Shonkoff & Phillips, 2000, p. 3)

Members of this national council, including Shonkoff and Phillips, outlined that the three frequently highlighted executive functions listed above work together and are “building blocks for the development of both cognitive and social capacities” (p. 3).

Self-regulation has also been linked with school readiness (Bierman, Nix, Greenberf, Blair, & Domitrovich, 2008; Blair & Diamond, 2008; Raver, Li-Grining, Bub, Jones, Zhai, & Pressler, 2011; Snow, 2006) and as a protective factor aiding in the academic success of children in poverty (Raver, 2012; Sektnan, McClelland, Acock, & Morrison, 2012). This growing understanding of brain development has enhanced the ability of educators to recognize the important connections between executive functions and early school achievement, and social and emotional development. In the next session, a review of the history and components of Tools will facilitate an understanding of how the development of self-regulation is influenced by activities and techniques of the curriculum.

Tools of the Mind

In this section, an historical review of the development and implementation of Tools will provide the background on the development and implementation of the program. The second part of this section will include an examination of the program’s components and their relationship to self-regulation.

Historical Review of Tools of the Mind

Tools was developed at a time when nationally, evaluation of developmentally appropriate teaching techniques was taking place and when educational leaders were contemplating the impacts of high stakes testing. In response to these challenges and to support the cognitive development of young children, the collaboration between Russian psychologist Elena Bodrova and American Psychologist Deborah Leong resulted in the development of teaching tools to scaffold early learning and in an innovative teacher training technique (Bodrova & Leong, 2001). Bodrova and Leong (2007) reported that the development and implementation of Tools occurred in four phases, beginning in 1993. In its initial phase, the collaboration between Bodrova and Leong resulted in many complications. Because differences existed in expectations of young children's educational exposure and skills between the United States and Russia, and due to difficulties such as translation, the developers realized that Tools techniques needed to be created that were not only based on Vygotskian principles but also that could be adaptable for children in the United States. Within the second phase in 1996, large scale teacher training and implementation took place with 8 schools and 78 teachers representing preschool through second grades in a large urban district. Due to issues such as lack of a curriculum and training manuals inconsistent results, the full empirical investigation was not possible. However, informally, Bodrova & Leong (2007) reported promising findings. When teachers implemented the program with fidelity, very good progress was observed, even for students who were considered at-risk (p. 27).

The third phase of the program development narrowed the focus to kindergarten and a small pilot of preschool classrooms. This phase included an empirical study using

5 control (N=218) and 5 experimental (Tools, N=208) groups with an at-risk population of children. Demographics of teachers and students were matched for both groups. A pre-test, a writing sample, showed that both groups were similar in their development of early literacy. Kindergarten classrooms consisted of one teacher with an average class of 20 children. In preschool classrooms, the teacher student ratio was 2:18. The evaluation of the program took place during a six-month period (Jan-June, 1997), and utilized the computerized assessment system, the Early Literacy Advisor (ELA) that Tools had been refining. Use of this system allowed the researchers to align teaching techniques with assessment to assist teachers in understanding student progress and scaffolding expectations.

For the kindergarten classrooms, the intervention included use of three teaching techniques from the Tools program, namely scaffolded writing, writing learning plans, and sound analysis. Researchers estimated that these Tools activities occupied 10% of the instruction in the classroom per week. Each classroom was assigned a staff member (paraprofessional one day per week) to assist with implementation and data collection. Both experimental and control groups participated in a computerized phonics program, and both groups of classrooms had traditional literacy periods.

Assessments during the six-month period measured letter recognition, sound-to-symbol correspondence, words vs. pictures, instant words, and writing samples. Measures of self-regulation were only collected at the end of the assessment period in the spring. Researchers used S-Plus software for data analysis. Results showed better performance on all measures, including pre-literacy measures, and better rates of progress by children in the experimental group. Use of the Tools techniques did not result in

negative effects. Statistically significant differences were reported between experimental and control groups on the number of words written, complexity of written messages, correspondence between story and re-read of that story, use of writing conventions, spelling, and better phonemic encoding (p. 32). With respect to pre-reading, students using Tools techniques made statistically significant gains in sound to symbol correspondence, voice to print match, understanding of concepts of a sentence, and understanding of the symbolic function of printed word (p. 32). Statistically significant differences between the two groups were not found, however, for letter recognition, instant words, and words vs. pictures. Overall, a higher level of writing was observed in classrooms following Tools techniques. Similar results were observed in the preschool with students from classrooms using Tools techniques showing stronger growth in early literacy skills.

Results in this phase also indicated that students who participated in the Tools classrooms demonstrated better progress than peers in control groups. Additionally, Bodrova and Leong were able to control for fidelity, and found that teachers with the strongest results were the ones who had higher degrees of fidelity. An unexpected complication arose, however, when 30-60% of the children moved or were absent. The high level of absenteeism and mobility compromised the implications of the results. In an effort to align the curriculum with national and state standards, the project was moved to Mid-Continental Research for Education and Learning (McREL) in the final phase. The development of a curriculum appropriate for preschool and kindergarten was enabled during this phase, which focused research on two model classrooms. The development of tests and computerized assessments continued during this phase. This empirical

evaluation resulted in the implementation of the curriculum in a growing number of locations at the preschool and kindergarten levels.

Tools of the Mind and Self-Regulation

Self regulation is a central component of social development and is the cornerstone developmental milestone that Tools of the Mind supports. Bodova and Leong (2007) defined self-regulation as “the ability to act in a deliberate, planned manner in governing much of their own behavior (p. 127). By the end of kindergarten, children should be able to regulate their physical and emotional behaviors, as well as some of their cognitive behaviors. Because younger children are reactive to the environment, children who have developed self-regulation can inhibit their impulsive reactions to the environment and act in a thoughtful, planned way. Components of self-regulation that are targeted in this curriculum include inhibition of aggressive behaviors, focused attention, willfully ignoring distractions, delayed gratification, and controlling their emotions. Throughout the curriculum, activities that purposefully target the development of self-regulation are made available (Bodrova & Leong, 2007; Hyson, Copple, & Jones, 2006). These activities yield the acquisition of the “mental tools” that children need in order to learn in a deliberate fashion. Because a major premise of Tools is that children gain control of their behaviors, both internal and external by using mental tools (Bodrova & Leong, 1996), children participate in their learning, and are not just recipients of instruction by their teachers. In addition to building foundational skills in literacy and math, Tools activities also target the development of social-emotional school readiness.

Tools classrooms focus on helping children learn to self-regulate by using intentional make-believe play and games that target self regulation, by providing

opportunities for positive classroom interactions, and by specially constructed learning tasks. Blair (2002) has called for early childhood education programs to assist children in the development of self-regulation and school readiness by providing supportive environments. Tools of the Mind, then, may provide the development of literacy and mathematics skills necessary for required assessments, while at the same time, provide opportunities for social development, play, and exploration. Helping children learn how to develop emotional and behavioral self-control, a major objective of Tools, may increase academic achievement (Bodrova & Leong, 2007).

To better understand how self-regulation maybe related to achievement via Tools, an examination of the components of Tools is necessary. Specific components of Tools must also be defined and examined. These components include private speech, sociodramatic play, and mediators.

Components of Tools

The role of self-regulation/executive functions is explicitly focused in the curriculum, Tools of the Mind. Bodrova, Leong, and Akhutina (2011) explained that the development of intentional self-regulated behaviors is dependent on social interactions. The components and activities of Tools are aligned with Vygotskian as well as post-Vygotskian research, namely that of Elkonin (1963), Galperin (1992), and Venger (1988, as cited in Bodrova et al., 2011). Like Vygotsky, Luria (2002) believed that mental development occurred “in the process of objective activity and communication with adults” (p. 21). By further developing their executive functions, children gain control not just of their behavior but also of their cognition. In this way, Tools emphasizes the development of intentional, self-regulated behaviors.

This development is dependent on certain activities and philosophies utilized in the Tools curriculum, mainly use of mediators, sociodramatic play, private speech, and scaffolding. Vygotsky (1978) extended the idea of tools to the human mind as he viewed the process of development. Vygotskians believe that mental tools can be used, invented, and taught to others. Mental tools also take on two forms: external in the early stages of development, and internal in later stages when they exist in the mind without external support. Without these mental tools to help humans master their own behavior, we would be limited to reacting to the environment. Mental tools “enable humans to plan ahead, to create complex solutions to problems, and to work with others towards a common goal” (Bodrova & Leong, 2001, p. 17). Selected mental tools, specifically dramatized play, mediators, and other activities are described below.

According to Vygotsky (1978), “a mediator is something that stands as an intermediary between an environmental stimulus and an individual response to that stimulus” (as cited in Bodrova & Leong, 2001, p. 51). In Tools, mediators are created to prompt specific responses. In so doing, mediators can facilitate mental processes and social behaviors. Bodrova and Leong created activities that follow Vygotsky’s thinking by proposing that mediators function to help children solve problems and to make it possible for them to perform independently, as well as to transition from lower to higher mental processes. The processes that mediators facilitate include perception, attention, memory, and thinking. Because mediators are used to scaffold in order to help children perform without assistance, teachers need to plan the activity, the external mediator, and when to remove the mediator. Props, such as lips and ears, are examples of mediators. For example, during the activity “buddy reading” one child reads (tells the story, even if

they are making up the words) while the other listens. The reader holds paper lips, while the other holds paper ears. These mediators help keep the children in their roles. Because they switch, this activity fosters working together and taking turns. If used after children have developed internal strategies, however, external mediators lose their value and can be detrimental to learning.

Just as mental tools lead to the development of higher mental functions, play was also considered essential. Vygotsky (1962, 1978) considered play to be the source of development. Elkonin (2005) elaborated on the essential role of dramatization, citing historical thinking on the role of make-believe play. Vygotsky (1962) stated that play consisted of actions with objects, directed at a future social action (p. 439). In this way, play transfers meaning from one object to another. Consistent with Vygotskian thinking, Bruner (1972) believed that play assisted children in developing cognitive flexibility. Based on the works of other Vygotskians, Karpov (2005) described the associations between play and the increased ability to retell details from a story and enhancing problem-solving skills. Karpov (2005) explored how play contexts contributed to the development of cognitive self-control and regulation.

Bodrova and Leong (2005) have continued to assert the importance of play in Tools. Tools emphasizes structured dramatic make believe play as the leading activity of kindergarten-aged children (Bodrova, et al., 2011). These authors have contended that when children create pretend scenarios using props in symbolic ways, and engaging in conversations with their peers, they engage in what Bodrova and Leong (2005) called “productive play” (p. 37). Accordingly, children can learn to prioritize their goals and behaviors, as well as to learn to delay gratification. By doing so, children develop self-

regulation. These deliberate actions also demonstrate how children use language in their interactions with peers. The authors contended that children need this type of play to foster interpersonal skills and self-regulation, skills that are not often the result of teaching only letters and using flashcards. This thinking is supported by Bredekamp and Copple (1997), who emphasized that the development of social competencies is enabled by rich play contexts. An additional benefit of play includes creating opportunities for verbal interactions with peers and teachers (Dickinson & Tabors, 2001).

Tools incorporates play to foster the development of deliberate behaviors, such as planning for future activities and not just future location of the next activity within the classroom. Vygotskians believe that children who actively engage in mature dramatic play receive cognitive and social benefits (Bodrova & Leong, 2007). Children dramatize scenes from the books read in class. In the first six weeks of the program, children act out fairy tales, and later act out Magic Tree House books. This series of engaging chapter books was chosen as the series to be used during literacy activities. Before dramatizing, time is spent making props (mediators) to use in dramatization. Each child then chooses the role they will dramatize. Next, each child draws and writes their play plan. This plan helps keep them in their roles and reduces arguing over who gets to play which role. This same activity is available each day, so that children have turns playing each role. Play also facilitates what Bodrova and Leong (2007) call “cognitive de-centering” (p. 134). By allowing children to take the perspectives of others, play fosters building community in the classroom.

In dramatized play, Tools recommends that teachers intervene, but only to prompt staying in roles, using props, etc. If teachers intervene too much, play is no longer child-

directed, but rather teacher-directed. Activities that are only teacher directed lessen the opportunities for the development of self-regulation. In addition, when the teacher is intervening and directing, he or she is not able to observe each child and their zone of proximal development. It is through this thoughtful observation that teachers become aware of how much scaffolding to use. Teachers, then, play a critical role in scaffolding the instruction and assisting students in verbalizing their plans for play in order to facilitate development of cognitive processes (Yang, 2000). Scaffolding has been associated with increased learning and positive outcomes in young children (Bodrova & Leong, 2007; Henderson, Many, Wellborn, & Ward, 2002). Tools has identified the following teacher activities as fostering “higher levels of play:

1. Make sure children have enough time to play
2. Provide ideas for themes that extend children’s experiences and enrich the play
3. Choose appropriate props and toys
4. Help children plan their play
5. Monitor the progress of play
6. Coach individuals who may need help
7. Suggest or model how themes can be woven together
8. Model appropriate ways to solve disputes
9. Encourage children to mentor each other in play” (Bodrova & Leong, 2007, p. 146).

In addition to sociodramatic play, the development of private speech is emphasized in Tools. Vygotskian theories of development asserted that while engaged in tasks, children used private speech for regulation of the flow of their cognitive processes

(Luria, 1961;Vygostky, 1987). An important belief for Vygotsky (1987) was that children have the capacity to talk over meaningful experiences. Purpose can be expressed through dialogue. This dialogue has the capacity to move from regulation by other to self-regulation.

Like other Tools activities, private speech and sociodramatic play are dependent on the recognition by the teacher of the child's ZPD. The day to day learning and the development of school readiness are dependent on the scaffolding incorporated by the adults (Bodrova, et al., 2011). In these ways, teachers act as mediators, supporting the development of their students (Kozulin & Presseisen, 1995; Yang, 2000). Other activities are valuable from the Vygotskian perspective including games with rules, productive activities (drama and storytelling, block building, art and drawing), pre-academic, and motor activities (Bodrova & Leong, 2007). One example is the freeze game, where children listen to music then pose in a position shown on a card by the teacher. This activity fosters stopping (inhibiting) and purposeful behavior. Another example is the use of cooperative games to foster academic learning. These games give children the opportunity to work together. In addition to working together, early literacy and math are developed through processes that are unique to Tools, including play planning, scaffolded writing, and math games to foster problem solving (Bodrova & Leong, 2007).

These learning activities foster discovery and use of mental processes, as opposed to acquisition and recitation of specific facts. The peer group is essential, and dialogue is encouraged to solve problems together. Indeed, the Tools classroom is not a quiet classroom.

In contrast, closed-ended activities prevent children from constantly challenging

themselves and setting new goals, in addition to less need for self regulation with higher levels of teacher direction. Teachers must understand the regulatory cognitive processes that occur when a child is allowed to problem solve and set goals in line with his or her own developmental trajectory. This understanding may inform the manner in which teachers *direct* their classrooms. Table 1.1 summarizes the activities found within Tools classrooms and their relationship to self-regulation.

Table 1.1 Tools Activities and Self-Regulation

Type of Activity	Examples from the Curriculum
<i>Self-Regulation:</i> Effortful Control, Selective and Sustained Attention, and Emotion Regulation	Dramatic Play: children plan to act out a specific role from a story during make-believe play. They stay in that role until a later role is allowed. Working in Pairs: taking turns reading and listening
<i>Literacy:</i> Phonemic and Phonological Awareness Sound Symbol Correspondence	Sound games based on Elkonin Boxes Scaffolded Writing Read Aloud Activities
<i>Math:</i> Numbers and Operations Geometry	Practice in rote and meaningful counting (Numerals, Timeline Calendar, Number Line Hop Scotch) Using Venger Drawing and Block Building Activities
<i>Social-Emotional Development</i> Prosocial Skills Social Problem-Solving	Cooperative paired activities (Buddy Reading, Freeze Game, Story Discussions, Attribute Game) Practice social problem-solving during make believe play and discuss potential problems during Share the News

Recognizing the importance of executive functions for school success, a number of other programs have been developed, primarily at the preschool level, to focus on specific training of executive functions. Some have introduced school-based interventions that are carried out by experimenters and supported by teachers (Lennon, Li-Grining, Raver, & Press, 2011; Raver, 2004; Raver, Li-Grining, Bub, Jones, Zhai, & Pressler, 2011). These interventions included teacher trainings on effective management

of classroom behaviors, mental health consultation and stress reduction workshops, for teachers, and for children with behavioral issues, one to one activities that targeted the promotion of executive functions. Rueda, Rothbart, & McCandliss (2005) also introduced an intervention requiring one to one interactions with children for the specific training of executive functions. Some interventions have even been developed using innovative computerized programs to train the development of executive function (Klingberg, Forssberg, & Westerberg, 2002; Thorell, Lindquist, Nutley, Bohlin, & Klingberg, 2009). Distinct from these intervention programs, Tools offers a comprehensive early childhood Vygotskian-based curriculum (Bodrova, et al., 2011).

Alignment with Developmentally Appropriate Practice

Activities found within the Tools curriculum align with the many practices consistent with NAEYC's (2009) definition of developmentally appropriate early education. NAEYC recognized the importance of scaffolding, ZPD, play, safe and consistent relationships, shared learning with peers, and self-regulation, all happening within social and cultural contexts, and all consistently practiced in Tools classrooms. Tools has been extensively written about within NAEYC publications, and is recognized as an approved alternative to NAEYC accreditation for quality full-day kindergarten (doe.mass). Interest in Tools continues to grow, as evidenced by several research projects currently in process (SREE, 2012).

Limitations of Tools research

With reference to the research conducted during the developmental phases of Tools, some limitations should be considered. First, during the empirical phase, only three Tools techniques were introduced into the experimental classrooms. Because

current implementation of the program involves the program in its entirety, comparisons with Bodrova and Leong's (2007) study may not be practical. Second, both control and experimental groups received a typical literacy period and a computer-based phonics program. The only difference assumed between the groups was the intervention of Tools techniques. The conclusions made regarding the efficacy of Tools did not account for the possibility of an accumulated or combination effect of all of those interventions. Finally, because self-regulation was only measured during the post-test, conclusions about how Tools impacts self-regulation may have been compromised.

Summary

In summary, Tools attempts to accomplish the goals that Vygotsky and his contemporaries envisioned would be manifested through educational practices that enabled children to participate actively in their learning. The activities of this program consistently support this philosophy, as well as the development of self-regulation and executive functions. As Bodrova, Leong, and Akhutina (2011) concluded,

“They (the activities of Tools) do so by requiring children to (1) monitor and evaluate their own as well as their peers’ performance, engaging in prospective and reflective thinking—quintessential manifestations of executive function; (2) shift cognitive set, defined as the ability to flexibly maintain competing sets of rules or instructions that challenge working memory and the ability to hold multiple representations of an object or set of objects in mind and to switch between them; and (3) use language to structure their own and others’ behavior” (p. 9).

Despite supporting promoting the development of self-regulation, questions remain regarding the overall efficacy of the program. While the activities and techniques utilized in the Tools of the Mind curriculum support the goals of the development of self-regulation and executive function, the efficacy of the program as it pertains to academic achievement has not been adequately studied.

Statement of the Problem

The demands of high-stakes testing have resulted in kindergarten environments focusing more on direct instruction of explicit skills and less on social development. Self regulation is a central component of social development and is the cornerstone developmental milestone that Tools of the Mind supports. While Tools is gaining popularity, research on its efficacy is scarce. In particular, no research has been conducted which examines the effect of Tools on academic achievement. Much of the available data on Tools is not empirical, or may be biased, as much of research has been conducted by the developers of the program.

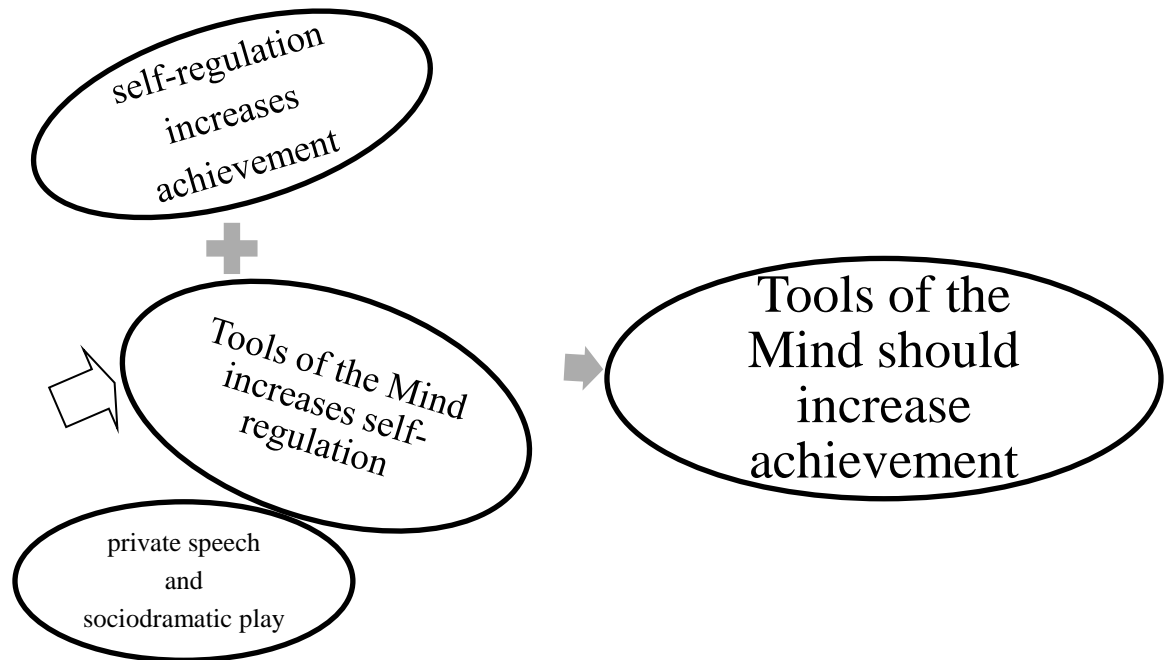
Current empirical evidence provides support that students in Tools classrooms showed higher levels of self-regulation and executive function (Barnett, Yung, and Yarosz, Thomas, Hornbeck, Stechuk, and Burns, 2008; Diamond, Barnett, Thomas, and Munroe, 2007). Executive function has also been linked with increasing academic achievement (Best, Miller, & Naglieri, 2011; Blair & Razza, 2007; McClelland, Cameron, Connor, Farris, Jewkes, & Morrison, 2007; and Ponitz, McClelland, Matthews, & Morrison, 2009). Additionally components of Tools, such as private speech Angina, Kommers, & Steehouder, 2011; Corkrum, Humphries, Mullane, & Theriault, 2008; Dougherty & White, 2008; Ferneyhough & Fradley, 2005; Lidstone, Meins, & Ferneyhough, 2011;

and Winsler, Manfra, & Diaz, 2007) and dramatic play (Elias & Berk, 2002; Fantuzzo, Sekino, & Cohen, 2004; and Nicolopoulou, Barbosa de Sá, Ilgaz, & Brockmeyer, 2010) have been associated with increasing self-regulation levels. While components of Tools have been associated with increased self-regulation, and while self-regulation and executive function have been associated with higher academic achievement, a gap exists with the lack of empirical research connecting achievement with Tools.

Conceptual Framework

The conceptual framework (see Figure 1.1 below) illustrates that self-regulation and executive function have been positively associated with achievement in both reading and math. The framework also illustrates that specific components of Tools have been associated with fostering the development of self-regulation and executive function. The framework, thus, illustrates the logic that because Tools fosters self-regulation, and self-regulation leads to better academic achievement, then Tools should have positive effects on academic achievement.

Figure 1.1 Conceptual framework.



Purpose Statement

The purpose of this study is to assess the effects of two types of instruction (self-regulation vs. teacher-directed) on the academic achievement of kindergarten students as measured by MAP scores. Due to its focus on self-regulation, children who experienced the Tools of the Mind curriculum are hypothesized to achieve higher MAP scores in mathematics and reading as compared to children who experience other traditional, teacher-directed curricula.

CHAPTER 2

LITERATURE REVIEW

In this literature review, empirical studies related to early childhood educational practices, self-regulation, and a specific curriculum, *Tools of the Mind* are reviewed. For my research, a focus was maintained on early childhood, with a specific focus on kindergarten if research was available. Much of the literature on self-regulation focused on early childhood which includes preschool and kindergarten. Empirical studies were drawn from professional journals from the last fifteen years (1998-2013). ERIC, Academic Search Premier, and Education Journals were the academic search engines used, with key words including: “self-regulation,” “self-control,” “executive function,” “kindergarten curriculum,” “private speech,” “dramatic play,” “Tools of the Mind.” Articles were selected based on their focus on the same age level as this inquiry and their relevance to academic achievement. For this research, articles that focused on specific disabilities related to self-regulation and cross-cultural aspects were not included, unless particularly relevant. Several articles that were specifically written about Tools of the Mind were also excluded, as they were not empirical and were authored by the creators of the curriculum, but were, however discussed in the introduction. The following review focuses on the cognitive and behavioral components self-regulation, the relationship between self-regulation and achievement, components of the Tools curriculum, and any available study on the efficacy of Tools.

Executive Function and Self-Regulation

This review begins with a study that demonstrated that executive functions were

possible in young children. Recognizing that executive functions are possible and important in young children is an important basis upon which the subsequent studies can build. Several studies have focused on the relationship between executive function and self-regulation and academic achievement. While some studies have examined the role of self-regulation on early literacy, including reading and math, others have centered their research on one subject, such as math.

Self-regulation/executive function and academic achievement

Davidson, Amso, Anderson, and Diamond (2006) provided examples of tasks associated with the development of executive functions and cognitive controls across a wide age range. Three abilities that demonstrated mature cognition included working memory, inhibition, and cognitive flexibility. This study attempted to make predictions about how these abilities develop and how they are related to each other. These abilities were tested using a battery of tasks, including the Simon task (Craft & Simon, 1970; Fitts and Seger, 1953; Hommel, 1995; Hommel, Proctor, & Vu, 2004; Lu & Proctor, 1995; Simon & Small, 1969; Simon, 1990; Simon & Berbaum, 1990, as cited in Davidson et al., 2006) and task-switching (Cepeda, Kramer, & Gonzalez de Sather, 2001; Cohen, Bixenman, Meiran, & Diamond, 2001; Crone, Bunge, Van der Molen, & Ridderinkhof, in press; Crone, Ridderinkhof, Worm, Somsen, & van der Molen, 2004; Reimers & Maylor, 2005; Zelazo, Craik, & Booth, 2004, as cited in Davidson et al., 2006). In the Simon task, a “non-spatial aspect of the stimulus (such as its color or identity) is relevant and its spatial location is irrelevant” (Davidson et al., 2006, p. 2038), increasing and decreasing working memory requirements. Task-switching is a task that attempts to measure cognitive flexibility and taxes both working memory and inhibition.

In this study, Davidson et al., (2006) hypothesized that young children would have more difficulty with inhibition tasks, whereas young adults would have more difficulty with tasks involving memory. Working memory and inhibition were also hypothesized to be independent of each other. With reference to task switching, researchers hypothesized that participants would perform better on task where inhibition was required all of the time, rather than only some times.

Participants included 325 people ranging in ages from four to forty-five years, with an even gender distribution. Most of the participants were Caucasian and from middle to upper-middle class families. All participants completed a total of four computerized tasks designed to influence working memory and inhibition control. Specific instructions regarding condition were given and participants were allowed practice. Tests were presented with Arrows, Pictures, Dots, and Abstract Shapes testing accuracy, reaction time, and percentage of anticipatory responses.

On the Arrows test, two conditions were presented. On the Congruent trials, an arrow pointed straight down and participants were to respond on the same side. On the Incongruent trials, the arrow pointed to the opposite side, requiring participants to respond on the side diagonal from the arrow. Accuracy, speed, and reduced anticipatory responses were higher as the age of the participants increased. For accuracy and anticipatory responses, these results were highly significant, but not for speed of responding: (accuracy: $F(1,312)=57.06$; $p<0.0001$; AR: $F(1,312)=35.73$, $p<0.0001$).

On the Pictures test, participants viewed pictures of either a butterfly or a frog. Two conditions were presented. If the participants saw a butterfly, either on the right or left, they were instructed to press the button on the left. If they saw a frog, participants

were instructed to press the button on the right, whether the frog appeared on the right or left. On this test, performance improved as the participant age increased, yielding highly significant results (accuracy: $F(1,222)=17.93, p<0.0001$; RT: $F(1,222)=35.36, p<0.0001$; anticipatory responses: $F(1,222)=10.8, p<0.0001$). When given a longer amount of time to respond, the youngest children did not perform better on accuracy.

On the Dots test, working memory and inhibition were strained, while on the other tests, strain was placed on either inhibition or working memory. Participants were presented two types of dots (striped black and white or solid grey). Half of the participants were instructed to make a response on the same side of the dot (striped) and on the opposite side of dot (solid grey). For the other half of the participants, the rules were reversed. Two conditions were presented: Congruent (responding on the same side) and Incongruent (responding on the opposite side). These tasks required remembering the rules as well as inhibiting a response to same side when instructed to respond on the opposite side. As the age of the participants increased, so did their ability to perform with increased accuracy and speed, and with decreased inhibitory responses. Results revealed that young children could perform well with a single task, but accuracy dropped with tests requiring mixed tasks. The accuracy of older children was also impacted, but to a lesser degree. Overall, participants performed better on tasks that were Congruent (spatially compatible) ($r(233)=2.09, p<0.04$; $t(217)=2.49, p<0.01$).

This study contributed to the understanding of executive functions, not only in adults but also with children. As results demonstrated, holding information in mind was possible for young children. Young children were also found capable of inhibiting responses, as long as the rules remained the same for tasks. The results of this study can

be transferable to classroom practices and expectations of young children. Because exercising inhibition was found to be harder for young children, focusing on the development of inhibition in the classroom could result in better outcomes for children in schools. These results may also inform educational practices, such as instructional and assessment methods. In order to understand how to make these changes in practice, breaking down the components of executive function is necessary.

The investigation of specific executive functions and how they relate to academic performance has been the topic of much inquiry. More specifically, performance on tasks involving working memory and inhibition have consistently related to performance in academic achievement.

McClelland, Cameron, Connor, Farris, Jewkes, and Morrison (2007) investigated whether behavioral regulation predicted achievement in emergent literacy, vocabulary, and math skills. A direct measure, called, Head-to-Toes Task (Cameron, McClelland, Jewkes, Farris, & Morrison, in press, as cited by McClelland et al., 2007) was used to measure behavioral regulation. Specific behaviors observed included inhibitory control, attention, and working memory. The children were measured on their ability to do the opposite of what was instructed verbally. The sample consisted of 310 preschool children from two geographic locations, Michigan and Oregon.

Participants from Michigan were predominantly middle-to upper-middle class socioeconomically and of diverse ethnic backgrounds near an urban area. Participants from Oregon were of a mixed socioeconomic background in a rural area. In Michigan, entering three and four year olds were recruited through fall preschool orientations and through mailings sent home in backpacks from six participating schools, one of which

was a Title 1 school. Overrecruitment occurred at the Title 1 school. Participants were enrolled in 42 classrooms, with a final participant number of 217 students. The mean age of the children in Michigan was 4.43 years at the fall testing. Of the participants, 76% were Caucasian, 9% were African American, 7% were East Indian or Asian, 6% were Middle Eastern, and 2% were Latino. In Oregon, participants were recruited from three preschools and three Head Start preschools. Of the 165 children invited to participate, 93 participated and were enrolled in 12 classrooms. The mean age in Oregon for fall testing was 4.58 years. In Oregon, 25% of the children were Latino (83% speaking primarily Spanish), 48% Caucasian, 19% Asian, and 7% of other ethnicities.

In the fall and spring, children were given emergent literacy, vocabulary, and math tests and a behavioral regulation assessment. Parents also filled out background questionnaires, providing information such as age, gender, prior child care experience, ethnicity, and parental level of education. To measure emergent literacy, vocabulary, and math skills, children were administered the Woodcock Johnson Psycho-Educational Battery III Tests of Achievement (Woodcock & Mather, 2001, as cited by McClelland et al., 2007). In analyses for achievement tests, W-scores were used based on a centered W-score of 500. To measure behavioral regulation, the Head-to-Toes Task was used. To complete the task, children are required to utilize inhibitory control, attention, and working memory. For each of the 10 items, a possible score of 0, 1, or 2 was obtained. A 0 was incorrect; 1 was a self-correct, defined as “any motion toward the incorrect response but where the child then stopped and responded correctly” (p. 951); and 2 points were given for correct responses without faltering. The sum of the scores was calculated, with a range of 0 to 20 for fall and spring. Other measures of behavioral control,

specifically, the Social Skills Rating System (SSRS; Gresham & Elliot, 1990, as cited by McClelland et al., 2007) and the Child Behavior Rating Scale (CBRS; Bronson, Tivnan, & Seppanen, 1995, as cited by McClelland et al., 2007). Both teacher ratings, were used to compare with the Head-to-Toes Task to obtain preliminary validity for the latter. Both were used at the Oregon site, and only CBRS was used in Michigan.

A number of background variables were significantly correlated with behavioral regulation scores. For example, children who had stronger behavioral regulation were older ($r = .28, p < .001$) and had parents with more years of education ($r = .20, p < .001$). In the spring, girls appeared to have stronger behavioral regulation than boys ($r = -.14, p < .05$). Significant correlations for behavioral regulation and all three academic achievement areas were obtained, with the strongest correlation emerging between fall math and fall behavioral regulation ($r = .47, p < .001$) and the weakest correlation emerging between fall behavioral regulation and spring literacy ($r = .18, p < .001$). Results of this study indicated that children with higher behavioral regulation had better achievement in emergent academic skills in both fall and spring, supporting the use of the Head-to-Toes Task to predict achievement.

The results of this study are important as they support the idea of teaching children about behavioral regulation to better focus their attention and develop mechanisms to inhibit certain behaviors to improve their academic achievement. Children who are able to focus their attention and inhibit certain behaviors are better able to follow directions and complete tasks, which may lead to improved achievement. Of particular relevance is the special topic related to transition from preschool to kindergarten addressed in this study. Continued work in helping students with this

transition is suggested by the results of this study.

While the results are encouraging, there were some limitations with the study. The number of participants in Oregon was less than half that of Michigan. In addition, the samples from both regions may not have been adequately similar to exclude confounding factors, such as their differing socioeconomic levels. For example, one sample was from an area on the outskirts of an urban center, which may have increased the experiential knowledge of that group. Another limitation involves the lack of predictability of future academic success. The children were not followed through the end of kindergarten, nor were separate samples of kindergarteners investigated. Finally, while children with better behavioral control appear to be more successful, the study is not causal. Readers must not assume that better behavioral control causes greater achievement.

Expanding the study of McClelland et al., (2007), Ponitz, McClelland, Matthews, and Morrison (2009) examined a new assessment of behavioral regulation, teacher rated classroom functioning, and kindergarten outcomes. Behavioral regulation consisted of attentional focusing, working memory, and inhibitory control. Specifically, the authors were interested in finding out whether behavioral regulation upon entry into kindergarten predicted achievement in math, literacy, and vocabulary for end of kindergarten measures. The new assessment, Head, Toes, Knees, Shoulders (HTKS) was used to measure behavioral regulation in a sample of 343 kindergarten students. Participants were recruited from two sites, Oregon and Michigan. In the Michigan site, recruitment resulted in participation from approximately 38% of the district's entering preschoolers. In the Oregon site, participants were recruited from the three accredited preschools and

Head Start, resulting in participation of 58%.

Data were collected from parents (attention focusing and inhibitory control); from teachers (reports on children's classroom behaviors and interpersonal skills); and from children (administration of behavioral regulation assessment and achievement measures). Normed questionnaires were used to gather data from parents. Exploratory factor analysis was conducted at each site, with the two largest factors identified as behavioral regulation and interpersonal skills, with respect to teacher ratings. Child achievement data was obtained through normed achievement tests in mathematics, literacy and vocabulary using the Woodcock-Johnson Psychoeducational Battery III Tests of Achievement. The authors utilized the HTKS task to measure behavioral regulation of the participants. This task consisted of responding to commands, such as "touch your head" and "touch your toes." Low scores indicated lower behavioral regulation; higher behavioral regulation was indicated by high scores. Results indicated that higher levels of achievement were reached for students in the spring who displayed higher levels of behavioral regulation in the fall. In addition, the authors were able to show adequate cross-examiner consistency for the HTKS (66%).

While results indicated that higher levels of achievement were reached for students who displayed higher levels of behavioral regulation ($r = .25, p < .01$ for attentional focusing; $r = .20, p < .01$ for inhibitory control), some limitations emerge in this study. Several threats to internal validity can be found. At both sites, families dropped out of the study (24% at the Michigan site and 50% at the Oregon site). Differences may have existed between those that remained and those that dropped out. In addition, in both cases, the participants were chosen from a group of families who were

already participating in other longitudinal studies related to achievement in early grades. Their participation could be indicative of a difference with the regular population of families enrolled in school. The sample itself may have been sufficiently different to confound comparisons. For instance, more members of ethnic minorities were found in the Oregon site. Other threats to validity include some of the assessments having been translated. The authors took precautions by utilizing a native Spanish speaker to translate the assessment and then retranslate for the investigators to obtain results. While noble, translating a test into a language it was not normed in can adversely affect the construct validity.

Nonetheless, the study does have strengths, including the extent to which training of raters took place and levels of interrater reliability obtained. The strong relationship between self-regulation and mathematics success might lead educators to develop interventions and programs to foster growth in social-emotional domains in order to attain greater academic success. Future research may seek a sample that is independent of other studies and comparison groups that are more similar in background variables, such as SES in order to improve internal and external validity.

Continuing the examination of the relationship between emotional and behavior regulation in preschool and kindergarten achievement, Howse, Calkins, Anstopoulos, Keane, and Sheton (2003) hypothesized that the predictor variable of emotion regulation would be mediated by behavioral self-regulation. The authors hypothesized that this mediation would result in the observations of higher levels of achievement in children who demonstrated higher levels of regulation. Other variables considered included maternal education levels, IQ, and socioeconomic levels.

Available from a racially and economically diverse sample of an ongoing longitudinal study, the current sample consisted of 125 families identified from scores on the concurrent study's behavior scales. The mean age of the children was 4.5 years at the time of the preschool assessment. Participants were 47% male and 53% female, 37% African-American, and 63% European-American. Socioeconomic (SES) levels varied from lower- to upper-middle class.

Assessment took place during one preschool testing point and twice during kindergarten. During the preschool assessment, observations were conducted in a laboratory setting using activities designed to elicit feelings of anger and frustration to provide data about emotion regulation. Parents completed a checklist to assess emotion regulation as well. In the first follow-up assessment during kindergarten, approximately one year after the preschool assessment, laboratory observations were repeated and IQ tests administered. Parents again filled out a questionnaire. At the final kindergarten assessment, achievement tests were conducted and teacher ratings of self-regulation were collected.

The associations between maternal education, which was used as an SES measure, gender, and age were examined during preliminary analyses. Maternal education was significantly correlated with child IQ, $r = .38, p < .001$, with literacy achievement, $r = .22, p = .04$, and with listening comprehension, $r = .32, p < .002$, and was marginally significant with math achievement, $r = .19, p = .07$. Gender differences were not found on IQ or other achievement measures.

All correlations between IQ and achievement measures were significantly correlated, with correlations ranging from $r = .40 - .64, p < .01$. Behavioral self-

regulation was impeded by emotion regulation. Conversely, children who achieved the highest in literacy, math, and listening comprehension also exhibited greater regulation in the classroom. Hierarchical regression analyses were conducted to test the mediating relationship between emotion regulation and achievement. First, results were able to show that emotion regulation and behavioral self-regulation were related, $r = .41$, $p = .001$. Next, each area of achievement was tested with regression equations. Regarding literacy, results revealed a relation between emotion regulation and literacy achievement with maternal education and IQ held constant. Literacy achievement with maternal education and IQ already in the equation was also found to be predicted by behavioral self-regulation. Emotion regulation no longer predicted literacy achievement scores with emotion regulation entered last, leading to the suggestion that behavioral regulation mediated the relationship between emotion regulation and literacy achievement. Math achievement scores followed similar patterns.

This study is significant because it is able to provide evidence that behavioral self-regulation is important for the development of academic achievement. Moreover, this study was able to articulate the mediating effects of behavioral self-regulation when emotion regulation might be compromised. Despite these contributions, this study was not able to provide guidance on whether children can be taught to be more self-regulated. In addition, some ratings used may have been subjective. For example, if parents have poor emotion regulation, the ratings on their own children may have been affected by their own levels of regulation. Regardless of these shortcomings, the results of this study have implications for practice and future research. These results may influence the development of interventions for children who show difficulty early with emotion

regulation as they may be at risk for academic problems in the future. Future studies may seek to examine specific aspects within the classroom that may foster the development of self-regulation. Additional research is needed to examine how frustration impacts attention as well as to examine the relationship between different types of regulation.

Interest in the relationship between self-regulation and academic development has caused continued inquiry about the different types of regulation. Specifically, the interrelations of executive function, effortful control, and false belief understanding, and their individual impacts on emerging math and literacy development in kindergarten were studied by Blair and Razza (2007). In addition, the authors were interested in children from low socioeconomic backgrounds, as these children were deemed at increased risk for school failure. The sample consisted of 170 children who were predominantly White from rural and nonurban locations, and also attended Head Start. Participants were recruited through a letter sent home to families through Head Start. Participants ranged in ages from 3.9 to 6.11 years at the time of testing, and consisted of 80 girls and 90 boys.

Data were collected while the participants were in preschool and repeated when the participants were in kindergarten. Data were collected through administration of psychological tests, questionnaires to parents and teachers, repeated measures of executive function, and verbal and non-verbal assessments. Using multiple regression, the authors examined the effect of executive function, effortful control, and false belief understanding on math and literacy development in kindergarten students. While each of the independent variables was found to be moderately correlated, each also accounted for unique variance in early mathematics and literacy ability. For example, a moderate correlation was found between the inhibitory control aspect of executive function

measured in preschool and in kindergarten, $r = .39, p < .05$.

In this study, the authors were able to articulate the relationships between effortful control, executive function, and false belief understanding to academic ability finding moderate correlations. Due to a lack of research examining the impact of self-regulation on early school success, this study contributed to and expanded the existing knowledge of self-regulation. In addition, this study attempted to examine the interrelatedness of several variables and their impact on early academic success. Focusing on multiple factors that affect academic success, the authors were able to control for some extraneous variables.

Strong correlations were, however, not observed in any of the measures. The sample, while of adequate size, focused only on low-income children. Generalizing these results to kindergarten children at large may not be appropriate, in that children of higher socioeconomic status may have different results, different academic readiness skills upon entering kindergarten, that may make them appear better regulated and better achieving. Further, the sample was not randomly selected, making generalizability even more problematic. Of particular concern is that on some variables, the child's performance was measured in a standardized method, while on other variables, the child's performance was rated by parents and teachers. Mixing these two methods may confound the results, as parents and teachers are susceptible to subjectivity, reducing the validity of the results. Nonetheless, this study points to the need to develop self-regulatory skills to improve potential of success in emergent academics. As this study described, the need for these skills may be even more important for already at risk children, such as those from low-income backgrounds.

Self-regulatory skills have been also described as either school skills or learning related skills. McClelland, Acock, and Morrison (2006) examined the impact of skills acquired in kindergarten on later elementary school performance. Focusing specifically on the role of learning-related skills, this study compared the achievement of children between kindergarten and 6th grade. Learning-related skills included the skills necessary for academic achievement. Including the executive functions of attention, behavioral self-regulation, and social competence, McClelland et al., (2006), defined learning-related skills as those that “describe behaviors, such as self-control, staying on task, organizing work materials, working independently, listening and following directions, and participating appropriately in groups” (p. 472).

This study set out to examine if the learning-related skills of kindergarten predicted reading and math skills between kindergarten and grade 6. In addition, this study also set out to compare the reading and math skills of children who displayed learning-related skills and those who did not. Learning-related skills, as rated by teachers, were expected to predict reading and math performance growth. Researchers also expected children who did not display learning-related skills to lag behind peer on reading and math between kindergarten and grade 6. This achievement gap was expected to widen.

The sample for this study included 538 children in Greensboro, NC. Of the 538, 51% were Caucasian, 49% were African-American and 51% were male. At the beginning of kindergarten, the average age was 65 months. Mothers achieved an average 13.61 years of education. Due to attrition, the final sample size was 260. Missing data resulted in the use Full Information Maximum Likelihood (FIML) estimation in *Mplus*.

All analyses, thus, had a sample size of 538. Descriptive statistics for the actual sample were similar to those using FIML.

Several methods were used to collect pertinent data. Parents completed a background questionnaire that provided demographic information. General intelligence was measured using a short version of the Stanford-Binet Intelligence Scale-Revised (Thorndike, Hagen, & Sattler, 1986, as cited by McClelland et al, 2006), as a control variable in the study was children's IQ. Learning-related skills were assessed using the Cooper-Farran Behavior Rating Scales (CFBRS; Cooper & Farran, 1991, as cited by McClelland et al., 2006), using only the teacher-rated, work-related skills subscale. Reading and math were measured using subscales of the Peabody Individual Achievement Test-Revised (PIAT-R; Markwardt, 1989) for grades K-2. Between grades 3-6, reading and math skills were measured using the North Carolina End-Of-Grade Tests (North Carolina Department of Public Instruction , NCDPI).

Results of the latent growth curve analyses using *Mplus* indicated that children's reading and math scores between K-6 were significantly related to kindergarten learning-related skills, with correlations between learning-related skills and reading ranging from .38 to .50, $p < .05$, and from .41 to .49, $p < .05$, for math. Because two separate measures were used (K-2 and 3-6), separate correlations were performed. Results indicated significantly correlated measures between K-2 and 3-6, with correlations ranging from .59 to .67, $p < .05$.

Once IQ, age, ethnicity, and maternal education levels were controlled for, initial reading levels at K (standardized coefficient = .17, $p < .001$) and growth in reading (standardized coefficient = .35, $p < .001$) between K and grade 2 were significantly

predicted by learning-related skills. Similarly, math levels between K and grade 2 (standardized coefficient = .17, $p < .001$) and growth in math skills (standard coefficient = .19, $p < .05$) were predicted significantly by learning-related skills.

Children who did not display learning-related skills at the beginning of kindergarten were found to fall increasingly behind between grades K-2. These children had lower IQ's and their mothers had lower attained lower levels of education. Between grades 3-6, however, these children did not appear to grow significantly behind, despite their overall lower performance in math and reading. Thus, the achievement gap significantly widened only between K-2.

This study added to the growing body of research regarding the impact of executive function skills on academic achievement. The results of this study were significant in that they suggest that children who are low stay low in early elementary years. In addition, these results suggested how consistent the relationship between kindergarten learning-related skills and achievement throughout elementary school can be. Another strength of this study is related to the age of the participants. Most of the other studies examined in this literature review are limited to the preschool years, and this study focuses on kindergarten through grade 6.

Despite these contributions, this study has limitations. Two different measures were used as the study was dependent on state testing for grades 3-6. The two measures in addition to the high attrition rate could have accounted for the finding of lack of achievement gap in grades 3-6. Although the research design included measures to correct for attrition of participants, interpretation of these results should be conducted with caution. Because only teacher ratings were obtained at the beginning of the year

regarding work-related skills, these ratings could have been subjective rather than actual measures of child behaviors. Furthermore, because background data was not collected on teachers, researchers could not determine if the ratings reflected the cultural expectations of teachers, which could have been of the dominant culture. In addition, one of the variables not controlled for was special education status. The lack of finding of achievement gap between grades 3-6 could have been influenced by services provided to students through special education for children who had performed poorly and qualified for services after grade 2.

Despite these limitations, practical implications can be derived from this study. As the results of this study suggest that children who performed low stayed low, educators can focus their attention early to interventions to support students academically as well as with developing competencies related to learning skills. Future studies should control for variables such as special education status to determine if the achievement gap truly ceases to be significant after second grade. In addition, future studies should seek collect data that are the same across grade levels making comparisons more effective.

Continuing with a focus on self-regulation, Rimm-Kaufman, Curby, Grimm, Nathanson, and Brock (2009) examined how self-regulation and classroom quality are related to adaptive behaviors in kindergarten classrooms. Self-regulation includes emotion management, attention and focus, and inhibition of certain behaviors. Gender, preschool experience, and family demographic information were also considered for analysis as they may contribute to risk factors in children. Classroom quality, which included emotional support, classroom management, and instructional support, was also examined as a potential moderator for adaptive behaviors. Adaptive behaviors were

defined by the authors as “children’s ability to persist at work, stay on task, attend to learning goals, and participate actively in learning” (p. 958). In this study, the authors hypothesize that later adaptive classroom behaviors would be predicted by self-regulation upon school entry, and that higher classroom quality would be related to better adaptive behaviors.

Participants were recruited from four rural districts from one mid-Atlantic state prior to entrance to kindergarten. Participants were predominantly poor from working class families. From the 333 children that were signed up by their parents, a sample of 172 children was randomly selected (4 or 5 from each classroom). From parent questionnaires, information was obtained regarding gender, income, parental marital status and educational attainment. Chi-square analyses did not show differences in the children that were selected and those that were not. Of the 172 participants, 80 were girls and 92 were boys with a mean age of 5.41 years. The participants were mostly Caucasian (144), African American (23), or other (5). Most of the families (39) reported incomes between \$15,000 and \$29, 999. Most of parents had a high school education and 130 mothers reported being married. Most of the children (103) did not attend preschool.

Data were collected using parent questionnaires to obtain background information, Teacher questionnaires about the behaviors and work habits of the students, and direct assessments of each child’s self regulation during the first five weeks of school by research assistants who were blind to the purpose of the study. From the Preschool Self-Regulation Assessment (Smith-Donald et al., 2007, as cited in Rimm-Kaufman et al., 2009), four subtests (Balance Beam, Pencil Tap, Toy Sort, and Gift Wrap) were used to assess self regulation. In the Balance Beam task, higher scores reflected higher self

regulation; for the Pencil Tap task, low scores (0-16) reflected higher self regulation; in the Toy Sort test, higher scores (up to 120 s) reflected higher self regulation; and in the Gift Wrap test, higher scores (up to 60 s) reflected higher self regulation. Confirmatory factor analysis generated the fall self regulation score. Measures of classroom quality were analyzed using ten dimensions of classroom quality. The means levels of the ten dimensions were calculated for each teacher across observation times. Four dimensions were related to the teacher providing emotional support ($\alpha = .93$); three dimensions were related to classroom management ($\alpha = .87$); and the remaining three dimensions were related to instructional support for learning ($\alpha = .94$). Teacher ratings on positive work habits and cognitive and behavioral control were used to measure adaptive classroom behaviors. Analyses were conducted using hierarchical linear modeling.

Teacher's report of children's behavioral and cognitive self control as well as work habits upon entering kindergarten was associated with children's self regulation. Most significant was the role of classroom management in predicting adaptive kindergarten behavior. Not surprisingly, classrooms that offered structured management practices contained children with higher levels of behavioral control ($t = 3.00, p < .01$), cognitive self control ($t = 3.76, p < .001$), and positive work habits ($t = 2.81, p < .01$).

Strengths of this study include the extensive piloting that took place in relation to the measures of child self regulation. In addition, training for research assistants was available, leading to very high intercoder reliability (interclass correlation [ICC] = .99). One great contribution of this study is how it offers a multifaceted view of classroom quality. However, some limitations are observed in this study. Due to only selecting 4 or 5 students from each classroom, the study is not able to account for other disruptive

students or lack of self regulation in the classroom. Many poorly regulated students in one classroom may affect the outcomes of the students that were included in the study, as well as the overall quality of the classroom. In addition, this study focused only on a rural area. Children from suburban and urban areas may present differently. Therefore, generalizing the results of this study to kindergarteners across the United States is not possible.

Future studies may seek a larger sample size and higher number chosen from each classroom. Researchers may also seek comparison groups that are both alike and different to control for differences in type of background, such as rural and urban. By having similar groups to be compared, stronger generalizations may be achieved. Nonetheless, this study does contribute to the understanding of classroom quality. It may also guide certain interventions improving the quality of classrooms to help students be more regulated, which in turn may lead to better achievement.

With a similar focus as Brock et al., (2009), Li-Grining, Votruba-Drzal, Maldonado-Carreño, and Kelly (2010) examined whether academic performance through fifth grade during elementary school was influenced by early learning approaches. These early approaches, which the authors named “early approaches to learning” (Li-Grining et al., 2010, p. 1062), included self-regulation behaviors such as persistence, emotion regulation, and attentiveness.

The data obtained for this study was retrieved from data that were collected in a longitudinal study of the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K). This sample was nationally representative of over 20, 000 kindergarteners and contained strong academic achievement data. Data collection began when this cohort

was in Kindergarten during the 1998-1999 academic year, and continued with five additional waves (spring of kindergarten, fall and spring of first grade, spring of third grade, and spring of fifth grade). Participants included students that of whom data were collected during all six phases. Based on these criteria, a total of 9,790 children were eligible to participate in the first wave sample. This first analysis produced achievement scores to be used as a covariate in the next wave of analysis with 10,666 participants. Information from two sets of scaled from the ECLS-K was used to capture early approaches to learning, namely, teacher and parent rating of social behaviors and self-regulations measures. Parent and teacher reports provided data regarding child, family, and school characteristics.

Two main effects models were devised and estimated in Hierarchical Linear Models (HLM). The first model tested variability in initial levels of achievement and trajectories of achievement. The second growth model examined the associations between early approaches to learning and achievement trajectories (waves 2-6). Variables that were controlled for included initial academic skills and child, family, and school characteristics. To analyze interactions, race/ethnicity, gender, and SES interactions were included in the first set, while an interaction term between wave 1 reading and early approaches to learning to predict reading trajectories.

Results of initial analyses (based on Chi-square tests) revealed great variability in the trajectories and initial levels of math and reading. Results suggested that 0.38 and 0.56 of an additional point in math and reading each month, respectively, was linked with each unit increase in early approaches to learning. Children who had better approaches to learning scored .56 and .52 of a standard deviation better in math and reading,

respectively, by the end of fifth grade, than their peers who had scored lower on approaches to learning. Results did not reveal associations between early approaches to learning and poverty status or parent's education. However, results did suggest that girls with better approaches to learning fare better in math by fifth grade, while boys fare better in reading by fifth grade. It is possible that these early approaches to learning serve as protective factors.

Certain limitations were found with this study. Because omitted data could not be controlled, causal conclusions should not be made from this study. While very comprehensive, the ECLS-K did not include data on classroom methods, child IQ, and other variables, which may have been at work and confounded the conclusions made. Despite these limitations, the study contributed greatly to the extant body of research that has suggested the relationship between self-regulation and achievement. Additionally, this study may have uncovered gender differences that point to the mediating, protective role that such learning approaches may have for different subjects. Future research may attempt to include variables not included in the data set used for this study to be able to increase the reliability of conclusions. Additional research is needed on the mediating role of self-regulation not just in relation to school achievement, but as a protective factor.

Switching from self-regulation to executive function, Best, Miller, and Naglieri (2011) assessed the relationship between executive functions and academic achievement in a comprehensive study of 2036 participants ranging in age from five to seventeen years of age. In this study, three tasks of executive function, nine academic tests from the Woodcock Johnson Tests of Achievement, revised (WJ-R), and performance aspects were

used to assess academic achievement. To assess executive functions, the standardized Cognitive Assessment System (CAS; Naglieri & Das, 1997a, as cited by Best et al., 2011) was used. In this study, researchers also attempted to examine how executive function develops over time.

Out of the sample of 2036 participants, 580 were under the age of seven. The average age was 9.4 years. Participants represented a broad area in the United States, including variables of race, gender, and parental education. The sample was representative of the population of the United States.

Three subtests of the CAS required that participants make a plan of action, apply that plan, and monitor the effectiveness of the plan related to the task. Subtests included Matching Numbers, Planned Codes (corresponding letters and codes), and Planned Connections (sequencing of numbers and letters). A multivariate analyses of variance (MANOVA) was used to determine whether improvements in executive function performance continued into adolescence. Further analyses were performed to evaluate correlations between completion time and accuracy on the tasks, and whether and how age was related to each to further inform how executive functions and achievement are associated.

Results indicated that differences in performance existed based on age for the younger age group (Pillai's $F(6,1810)=94.90, p<.001, n^2_p=.24$) and the older age group (Pillai's $F(12,3201)=61.16, p<.001, n^2_p=.19$). Significant age differences were observed. Significant improvements were also noted in performance from ages 5 to 6 to 7 for all tasks. Improvements continued through age 15 on Matching Numbers and Planned Connections, as well as through age 17 on Planned Codes. These data supported the

hypothesis that executive functions appear to develop through adolescence. Age differences were found on Matching Numbers and Planned Codes for the younger age range (Pillai's $F(10,1806)=63.08, p<.001, \eta^2_p=.26$) as well as for the older age range (Pillai's $F(20, 4484)=42.75, p<.001, \eta^2_p=.16$). For children between five and six, and between 6 and 7, significant improvements were observed in completion time and accuracy, but not on Planned Codes. As age increased, the magnitude of the age-related differences dropped, supporting the hypothesis that improvement slows down in adolescence. Results also indicated that the more accuracy a student showed, the quicker he or she worked, but this dynamic depended on age and task. With respect to math and reading achievement, similar trends were observed on tasks of executive function. In math, executive function appeared to be more closely related to problem solving than calculation.

Collectively, these findings suggest that focusing on executive function could have an impact on academic achievement in both reading and math in young children. This study contributes to the growing body of research and current understanding of the relationship between executive function and academic achievement. In particular, the strengths of this study include having a substantial sample size representative of the population of the United States, making these results more generalizable. The type of measures used also contribute to the strengths of this study, namely the use of CAS and the WJ-R. Both tests are standardized measures, and the WJ-R enabled the researchers to assess the relationship between executive function and achievement. One relative weakness is related to the possibility of other developmental processes, such as metacognition, occurring at the same time, and their particular role in impacting

achievement.

The studies reviewed above clearly articulated a relationship between self-regulation and academic achievement as measured by student outcomes in literacy or math in either preschool or kindergarten. The results of these studies have increased the understanding of this important association, and should be considered by educators in their planning of educational programs to assist students in the development of self-regulation. The following studies supported this conclusion, but focused specifically on the impact of self-regulation and executive function on later math achievement.

Executive function and math achievement

Brock, Rimm-Kaufman, Nathanson, and Grimm (2009) examined the relationship between executive function and kindergarten achievement. In this study, the authors identified two categories of executive function, namely, hot and cold. Hot executive function was associated with the coordination of emotional processing, while cool executive function was associated with the coordination of cognitive processing. The purpose of this study was to examine how hot and cool executive function affect achievement, learning-related behaviors, and how learning-related behaviors impact achievement at the kindergarten level.

Participants included 173 students randomly selected from 333 who signed permission to participate while enrolling for kindergarten. The sample included 90 boys and 83 girls from seven elementary schools in rural locations in the Southeast. Students were mostly Caucasian-American (73%) and African-American (17%). Most students (61%) had not attended any type of formal preschool. Thirty-six teachers, with an average number of 18 years of teaching experience, also participated.

Family demographic information was collected through questionnaires. Executive function (Preschool Self-Regulation Assessment, PSRA; Cameron & Morrison, 2007; Smith-Donald, Raver, Hayes, & Richardson, 2007, as cited by Best et al., 2011) and achievement (Woodcock-Johnson III Tests of Achievement) tasks were administered during Fall and Spring sessions. Four tasks were used to measure executive function. Cool executive function was measured using the Balance Beam task and the Pencil Tap task (Blair & Razza, 2007; Diamond & Taylor, 1996, as cited by Best et al., 2011). Hot executive function was measured using a Toy Sort task and a Gift Wrap task (Carlson, 2005, as cited by Best et al., 2011). Children's learning-related behaviors and engagement were assessed by teachers and research assistants using the Observed Engagement in Learning Scale (Rimm-Kaufman, 2005, as cited by Best et al., 2011). Learning-related behavior ratings included items pertaining to self-direction, hyperactivity and distractibility, working independently, and self-control. Engagement ratings included items on five classroom behaviors, including self-reliance, attention, disruption, compliance, and engagement.

To analyze results, hot and cool executive function scores were initially created. These scores were then used as predictors in hierarchical linear model analyses. A moderate correlation ($r = .50, p < .01$) was found between hot and cool factors. Upon entering school, greater hot executive function presentation was associated with less family risk, $r = -.16, p < .05$, and higher cognitive ability scores, $r = .17, p < .05$. All academic and behavioral outcomes were found to be mildly positively correlated with hot executive function, $r = .15-.35$, with the exception of Fall reading performance. Demonstration of greater cool executive function upon entering school was also

associated with less family risk, $r = -.31, p < .01$, and higher cognitive ability, $r = .42, p < .01$. All academic and behavioral outcomes were found to be moderately correlated with cool executive function, $r = .29 - .46$. Moderate positive correlations were also found between behavioral and academic outcomes, $r = .26 - .41$).

With reference to classroom-level variance, two of the four outcomes were significant with intraclass correlations (applied problems, letter word, learning-related behaviors, and observed self-regulation). Because these students were nested in classrooms, further hierarchical analyses were conducted. Out of these analyses, several predictors emerged as significant for achievement. Higher Spring math scores were associated with the Fall predictors of cool executive function, $t = 3.09, p < .01, d = .21$; cognitive ability, $t = 2.96, p < .01, d = .20$; and Fall math scores, $t = 6.95, p < .01, d = .45$. For predicting Spring reading scores, only prior achievement and cognitive ability were significant in Fall. No achievement outcomes were predicted by hot executive function. Small associations found through analysis of effect size were found between cool executive function and math achievement.

Higher teacher ratings regarding learning-related behaviors and higher engagement were given to children with higher cool executive function, higher scores of cognitive ability, who attended preschool, and to girls. These results highlighted the significance of cool executive functions on behavioral outcomes. Results of an analysis of the impact of learning-related behaviors on the relationship between executive function and achievement indicated that the following factors remained significant predictors of math in Spring math achievement measures: cool executive function, cognitive ability, and prior math performance. However, learning-related behavior and

engagement were not found to be significant predictors of Spring math achievement, leading Brock et al., (2009) to conclude that “behaviors exhibited in the classroom do not account for the relationship between executive function and math achievement” (p. 345).

This study contributed to the understanding of the unique association between cool executive function and math performance, although the goal of the study was to investigate the relationship between executive function and reading and math. Cool executive functions did not predict gains in reading achievement. The authors speculated that the largely reading-focused programs found in kindergarten perhaps allow students to make gains in reading regardless of their executive function levels upon school entry. Hot executive function was not associated with achievement. This result may be related to the expectation by teachers that this developmental stage includes behavioral dysregulation and that teachers arrange their classrooms and have management styles to compensate for the emotionally-based processing in their students. The results of this study demonstrate that better classroom behaviors are critical for classroom learning to take place and add to the understanding of school readiness.

Despite these contributions, some limitations were found in this study. The rural population from which this sample was obtained may have contributed to the levels of non-significance found in status of previous preschool and level of family risk. For example, the average level of cognitive ability could have been attributed to family risk factors, but not shown significance for gains in achievement. Outcomes on executive function tasks may have also been related to the motor ability of the children. Motor skills were tapped during the tasks, although the children did not need prior early literacy or numeracy. The study also focused on short time frame. While studying students in

kindergarten provided information pertinent to this study, prediction of later achievement was not addressed. Additionally, the study did not provide insight into the specific classroom activities that may support executive function development. Nonetheless, these results offer support for including the fostering of executive function in early childhood programs.

With a similar focus as Brock et al., (2009), Duncan, Dowsett, Claessens, Magnuson, Huston, and Klebanov (2007) examined the relationships between school readiness (defined by school-entry academic, attention, and socio-emotional skills) and later reading and math achievement. This study is considered unprecedented as it included data from six large-scale longitudinal studies, a wide range of school readiness indicators, predictors, multiple dimensions of academic achievement, rigorous analytic methods, and gender and socioeconomic variables. These components allowed for a broad examination of the association of early skills and later achievement through a meta-analysis of the six longitudinal studies with specific attention to the impact of early attention, academic, and socio-emotional skills on later achievement.

Six data sets, one from each longitudinal study, contributed measures of children's academic, attention, and socio-emotional skills at ages 5-6. Teacher reports, test scores, and early grade retention were used to measure achievement outcomes, with some achievement measures occurring as late as early adolescence in some data sets. Attention and socio-emotional behaviors were measured using parent, teacher, and observation reports. Analysis of data sets began with an estimation of similar sets of regression models across the studies.

With standardized coefficients ranging from .05 to .53, results of regression

analyses indicated that later reading and math achievement were almost always significantly predicted by school entry reading and math skills. For more than half of the coefficients, attention and attention problems resulted in statistically significant coefficients, but coefficients for socio-emotional behaviors were rarely significant. Results indicated similar patterns of significance across gender and socioeconomic variables. Of all of the school-entry skill categories, the three that appeared the most predictive of later reading and math achievement were entry level reading/language, math, and attention. Interestingly, “rudimentary math skills appeared to matter the most with an average standard coefficient of .33” (Duncan et al., 2007, p. 1437). Reading and attention standard coefficients were significantly lower than math (.13 and .07, respectively). From these analyses, behavior problems and social skills did not seem to be associated with later achievement.

Limitations were found within this study. Because socio-emotional measures have lower validity, bias cannot be ruled out on those measures. The authors acknowledged that the causal impacts of early math may be overstated. While problem behavior was not associated with later achievement, students with problem behaviors should not be ignored, as their behaviors may have deleterious effects on themselves and others. The results of this study do not inform educators about the types of programs that might be most useful for the development of academic achievement. Despite these limitations, the results of this study make a strong case for paying attention to early math skills. This emphasis is contrary to the emphasis placed on reading in most early childhood education programs (Brock et al., 2009). This study was also able to distinguish that attention, but not problem behavior or social skills were found to predict

achievement. These results are in contrast to those obtained by Brock et al., (2009) regarding the predictive role of behavior. Future studies might investigate the concurrent and ongoing development of reading and socio-emotional skills as these skills might not be present just upon school entry. An implication of this study might be the examination of potential interventions to improve the development of early math skills.

Furthering the understanding of executive function on early academic skills, Welsh, Nix, Blair, Bierman, and Nelson (2010) examined the associations between working memory and attention control (domain-general cognitive processes) and growth in reading and math achievement (domain-specific skills) across the kindergarten year. For the purposes of this study, executive functions included working memory and attention control.

The sample for this study included 164 children enrolled in Head Start classrooms in three counties in Pennsylvania. Of the 164, 14% were Latino, 30% were African-American, 56% were European-American, and 57% were female. The average age at the onset of Head Start was 4.49 years. With 68% of the sample families living below the poverty line, 40% of the students lived in households with two parents, 43% lived in households with single mothers, and 17% lived with family or in foster care. With reference to maternal education, 33% of the mothers had not completed high school, 46% had either a high school diploma or its equivalent, 19% had some technical training, and 2% had graduated from college. Participation was sought by identifying those students at the end of their Head Start year that would be eligible for Kindergarten registration. Data collected for another research project was used as the Head Start/prekindergarten data.

Data were collected through assessment of several measures. Reading and math

skills were assessed using three measures: The Print Knowledge, Blending, and Elision scales of the Test of Preschool Early Literacy (Lonigan, Wagner, Torgesen, & Rashotte, 2007, as cited by Welsh et al., 2010). The Print Knowledge subtest assessed letter, picture, or word naming. The Blending subtest assessed phonological processing. In the Elision subtest, children were asked to deconstruct compound words. These three scores yielded a composite score to represent emergent literacy skills. Reading achievement at the end of kindergarten was assessed with four measures. From the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001, as cited by Welsh et al., 2010), the Letter-Word ID and Story Recall subtests were used. These subtests assessed decoding skills and memory of details from stories. Two subtests of the Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999, as cited in Welsh et al., 2010) were used: The Sight Word Efficiency Scale and the Phonemic Decoding Efficiency scale, a timed test of nonsense words. A standardized, averaged composite score of the four subtests yielded a score of reading achievement at the end of kindergarten. Emergent math skills were assessed using the Applied Problems scale of the Woodcock-Johnson III Tests of Achievement. This test assessed the understanding of numbers and quantity.

General cognitive abilities were assessed using three tests: The Backward Word Span (Davis & Pratt, 1996, as cited by Welsh et al., 2010), the Peg Tapping Test (Diamond & Taylor, 1996, as cited by Welsh et al., 2010), and the Dimensional Change Card Sort (Frye, Zelazo, & Palfai, 1995, as cited in Welsh et al., 2010). In the Backward Word Span, children listened to a list of words and then repeated the words in reverse order. This task required children to both store and manipulate information in working

memory. On the Peg Tapping task, children were asked to tap on a wooden dowel once if the researcher tapped twice, and vice versa. This task measured inhibitory control, as children had to inhibit the urge to copy what the researcher did. On the Dimensional Change Card Sort, children had to sort cards of red or blue rabbits and boats based on color or shape on one dimension, and then the other. This task allowed the researcher to rate the child's skills in shifting attention. Correlations among the three measures and domain-general cognitive abilities ranged from .26 ($p < .001$) to .35 ($p < .001$).

Language, considered a covariate, was tested using the Expressive One-Word Picture Vocabulary Test (Brownell, 2000, as cited by Welsh et al., 2010) and two subtests of the Test of Language Development-Revised (Hamill & Newcomer, 1997, as cited by Welsh et al., 2010). Syntax comprehension was assessed using the Grammatical Understanding subtest and syntax expression was assessed using the Sentence Imitation subtest. A composite score for language at the beginning and end of prekindergarten was obtained from the three scales.

Among all cognitive variables, results indicated highly significantly correlations. In addition, results indicated that both domain-specific and domain-general cognitive skills were considerably stable from beginning to end of kindergarten. Zero order correlations were computed to assess whether cognitive skill composites predicted reading and math achievement at end of kindergarten measures. All correlations were statistically significant at the $p < .001$ level: kindergarten reading achievement was predicted by prekindergarten emergent literacy skills ($r = .40$ to $.48$); prekindergarten emergent numeracy skills predicted kindergarten math achievement ($r = .53$ to $.54$). Executive function skills in prekindergarten also predicted reading achievement in

kindergarten ($r = .39$ to $.58$).

To demonstrate how growth in domain-general cognitive skills in prekindergarten would affect growth in emergent literacy and math skills, a series of path models were estimated. At the beginning and end of prekindergarten, initial levels of executive functions (domain-general) predicted growth in emergent literacy skills ($\beta = .29$). Kindergarten reading achievement was predicted by growth in emergent literacy skills during Head Start ($\beta = .25$). Kindergarten reading achievement was also influenced by growth in executive functions during the prekindergarten year ($\beta = .36$).

For math, growth in emergent math skills during prekindergarten was predicted by initial levels of executive function ($\beta = .20$). A reciprocal relationship was also found, where growth in executive function during prekindergarten was predicted by initial levels of emergent numeracy skills ($\beta = .21$). Kindergarten math achievement was significantly influenced by initial levels of emergent literacy skills and growth in emergent numeracy skills during the prekindergarten year ($\beta = .19$ and $.25$, respectively). Kindergarten math achievement was also highly influenced by growth in executive function during the prekindergarten year ($\beta = .35$).

This study not only replicated McClelland et al.'s (2007) study that suggested that the development of literacy and numeracy skills are positively related to early executive function skills, but also extended other related research (Blair & Razza, 2006; Bull & Scerif, 2001). Specifically, this study demonstrated how growth in working memory and attention control during prekindergarten affect achievement in reading and math in kindergarten. Results of this study also suggested that “working memory and attention control provide an important foundation for domain-specific academic learning” (Welsh

et al., 2010, p. 49). Adding to the strength of this study was the ability to assess cognitive skills at three points. This longitudinal approach enabled the researchers to explore the role of executive functions into the following academic year.

Despite its strengths, this study also has limitations. Only three tasks were included in the measurement of working memory and attention control. The authors noted that children at this age have difficulty with longer sessions and tasks. On the positive side, the tasks that were chosen had been used previously with this age group and had been associated with the executive functions of working memory and attention control. Challenges in assessing executive functions in young children across different dimensions were also confronted. As with any correlational study, causation cannot be assumed with these results. Due to the high poverty rates of the sample, other factors, beside language ability, could have contributed to the results. These factors may include developmental delays, cognitive limitations, or mental health issues. Future research should continue with longitudinal designs, to enable the testing of the sustainability of the positive effects of executive function on emerging academic skills. Future studies may also include comparison groups of other risk factors or no risk factors.

Maintaining a focus on math, Bull and Scerif (2001) studied three executive functions (inhibition, switching, and working memory) to assess how they predict math ability. Other studies examined the same research question, but focused on significantly older students and were, thus, not included in this review (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; St. Clair-Thompson & Gathercole, 2006). The sample consisted of ninety-three children ranging between six and nine years, and was comprised of 50 boys and 43 girls. Participants represented a variety of schools in Scotland.

Participants in this study were asked to complete three executive function tasks (the Wisconsin Card Sorting Task (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993, as cited by Bull & Scerif, 2001), Stroop Task (Salthouse & Meinz, 1995, as cited by Bull & Scerif, 2001), and Counting span (Case, Kurland, & Goldberg, 1982, as cited by Bull & Scerif, 2001), and dual task (Baddeley, Della Salla, Papagno, & Spinnler, 1997, as cited by Bull & Scerif, 2001). The WCST tested a child's ability to conceptualize a task, maintain a set of stimuli, and to switch sorting criteria. The Stroop task provided stimulus materials, colors or numbers, and the researcher recorded timed responses. Three conditions were present (baseline, crosses to count, and Incongruent/Congruent). Interference was also measured. Finally, the dual task performance assessed the storage of verbal and spatial information. Digit span and visual tracking were tested.

To understand which of the measures of executive function were related to math, correlational analyses were performed. Results indicated that several measures of executive function correlated significantly with math ability. Higher working memory span was significantly related to higher math ability ($r = .43, p < .01$). Children who scored higher on math ability also had lower interference of irrelevant information ($r = -.46, p < .01$). Further analyses revealed that “working memory span, perseveration, and inhibition efficiency were each found to predict a significant amount of variance in mathematics ability (19%, 19%, and 21%, respectively)” (Bull & Scerif, 2001, p.282-283). Because these tasks shared requirements, predicting the amount of variance of each task was difficult.

This study contributed to the current understanding of the role of executive

function in the development of math abilities. Results indicated that children with lower math abilities had the greatest difficulty switching learned strategies. This study has implications for educational practices. Knowing that children need assistance to learn to inhibit previous or extraneous information and to shift may help teachers help students improve working memory, and ultimately, to improve their math performance.

Understanding that executive function does not represent one task, but rather a diverse set of executive functions, will help educators themselves think about curriculum, instruction, and targeted interventions.

Continuing the study of how executive function predicts math ability (Bull & Scerif, 2001), Bull, Epsy, and Wiebe (2008) studied whether academic achievement at age seven could be predicted by preschool performance on short-term memory, working memory, and executive functions. This study examined the relationship between early reading and math skills and executive functioning skills to assess the potential screening of preschoolers who might later develop difficulties in math.

Participants in this study included 54 girls and 40 boys in preschool in England, with an average age of 4.5 years. A battery of cognitive measures as well as math and reading outcomes were used to test participants as they entered school (P1), at the end of first grade (P2), and at the end of third grade (P3). Basic counting, phonics, and reading skills were assessed. Math was assessed using the Performance Indicators for Primary School (PIPS; Tymms, 1999, as cited by Bull et al., 2001) using a computerized program. Central executive functions were tested using the Shape School (Epsy, 1997, as cited by Bull et al, 2008), a story book with colorful shapes and images as the characters. The Story Book examined naming speed (Condition A), and inhibition (Condition B),

whereby children had to recall names and feelings of the characters. In a third condition, switching, children were asked to name figures with hats and colors of the characters that had colors, and then to shift between hats and colors. Another task involved the children viewing a design of blocks and then being asked to replicate that design with blocks (Tower of London; Korkman, Kirk, & Kemp, 1998, as cited by Bull et al., 2008), following specific instructions. This task measured children's ability to inhibit the urge to skip some instructions and build the end product. The Corsi blocks were used to assess short-term and working memory. The task involved pointing to blocks in the order presented, then backwards. Finally, Digit span was a task that assessed the forward-backward ability to repeat a sequence of numbers. Reversing required storage and organizing of the information.

Results showed one gender difference, where girls scored significantly higher on the Tower of London task than boys ($t(102)=2.43, p=.02$). Growth curve modeling and hierarchical linear modeling were performed to attempt to uncover the relationship between the predictor variables and achievement. Retention of verbal information was significantly related to later math and reading. A one-digit increase in verbal span at the P1 test was associated with a 2.01 point increase on the PIPS math and a 3.05 point increase on the PIPS reading. On Corsi blocks, a 2.39 point increase in PIPS math was seen for each additional chunk on non-verbal information retained at the P1 test. Shape School was also predictive of later math and reading. On the Inhibition condition, one unit above the mean at P1 resulted in a 4.91 increase in PIPS math and a 5.28 point increase in PIPS reading. Succeeding on more complex trials of the TOL was also associated with an increase of .53 in PIPS math and .61 in PIPS reading.

To evaluate the relationships between the predictor variables and PIPS math scores, correlational analyses were performed. All of the P1 cognitive measures were significantly correlated with PIPS math, with the strongest correlations observed in Digit Span Backward ($r = .52, p < .001$) and TOL ($r = .46, p < .001$). None of the Span tests was significantly correlated with math achievement by the end of P3. All of the central executive measures were, however, correlated with math ability, especially regarding visual-spatial working memory as tested on the Corsi-Backwards scan ($r = .39, p < .001$). For reading, at the P1 assessment, all predictor variables were significantly correlated, as in math, At the end of P3, only Digit and Corsi span, which test short-term memory, were significant for reading.

Regression analyses revealed a significant overlap in variance of reading and math achievement (lowest=36%, highest=56%). At P3, all significant associations between executive functions and outcome measures were removed, as executive functions accounted for similar variance in both math and reading (35.7%). Remaining significant predictors of math included Corsi Backwards span (5.5%, $p < .05$) and of reading, Corsi and Digit spans (4.4% and 3.0%, respectively, $p < .05$).

Several of the results did not have significance in this study. Shifting, for example, was not predictive of math or reading in the long run, but it did predict achievement at each testing point. Future studies may need to increase the age range of the sample, as the effects of shifting may be more suitable for measurement in older students. Despite this weakness, this study highlights the importance of visual-spatial working memory and its role in math achievement. As the study demonstrated, lack of executive function skills may predict future math difficulties. Teaching specific skills

targeting reading comprehension and math word problem-solving, which influence a child's ability to inhibit (as found by Blair & Razza, 2007) may be a way to influence student progress in reading and math. Based on the results of this study, preschools that assist children with the development of these executive functions may provide their students at an academic advantage that appears to be sustained at least until age seven. Implications related to the teaching and modeling of specific math tasks, such as estimation and use of number lines, should be considered.

Comparable with the focus of Bull et al., (2008), Clark, Pritchard, and Woodward (2010) examined executive function abilities of preschoolers to assess how they predict early math achievement. Another goal of this study was to assess whether potential associations found between preschool executive function and math endure when general cognitive ability and reading ability are controlled.

Participants were recruited randomly using a hospital database of age appropriate students necessary for the study. This selection resulted in a final total of 113 participants (55% male and 3.5 multiple birth). This study was conducted in New Zealand, and demographic information of the participants and families was consistent with the region, reflecting a broad range of maternal education (19.8% left school before age 16, 63.5% completed secondary school, and 16.7% attained a university degree) and a broad range of income levels. Of the 113 participants, 4.2% were Maori, 87.4% were New Zealand-European, 6.3% were from other European backgrounds, and 2.1% were Pacific Islanders. All families were able to speak English. Due to attrition, a total of 104 participants completed both 4-and 6-year old testing.

Upon turning four, a comprehensive developmental assessment was conducted for

each child, including measures of cognitive ability and executive function. Interviews of parents were conducted to capture information about family functions, social background, and child health and development. At age six, the same procedures were followed with the addition of standardized tests of educational performance (Woodcock-Johnson III, Math Fluency and Passage Comprehension). Executive function measures included a task used to assess complex executive planning (Tower of Hanoi; Simon, 1975; Welsh, 1991, as cited by Clark et al., 2010), a task to assess shifting/cognitive flexibility (Flexible Item Selection Task, FIST; Jaques & Zelazo, 2001, as cited by Clark et al., 2010), and a task to assess inhibitory control and set shifting (Shape School). At the preschool measure, teachers completed the Behavior Rating Inventory of Executive Function-Preschool (BRIEF; Gioia, Epsy, & Isquith, 2003, as cited by Clark et al., 2010). At the 6-year old measure, teachers completed a questionnaire to rate early classroom performance and behavior.

Descriptive statistics used to analyze independent samples included *t*-tests and chi-square tests. Relationships between measures of executive function were examined using Pearson correlation coefficients and analysis of variance (ANOVA). With respect to executive function at age four, results indicated that increased efficiency on executive function of inhibitory control, set shifting, and metacognitive planning was associated with a math achievement advantage of 5-10 points at age six. Teachers also rated children better in math when they had better preschool performance on executive function. On the task of complex executive planning, the majority of the children (55%) performed in the average range, while 23% performed below average, and 22% performed above average. On the task assessing cognitive flexibility/shifting, 75% of the

children demonstrated clear understanding of the task. Poorer efficiency was observed on the Shape School measure, which tapped into inhibitory control. On the BRIEF-P, a portion of the children (5-14%) scored above the clinical cutoff while 10.6% demonstrated significant difficulty with executive function.

With respect to measures collected at age six, just over one half of the students scored in the average range of Math Fluency. Teachers rated 4.8% of the students as having advanced math abilities, 11% below average, and 2% delayed. The majority (76%) of the students were achieving at or above expected levels in reading. Associations between executive function in preschool and later math achievement were examined, and found to be in the moderate range ($r = .18-.48$). One phase may have been too difficult (Shape School) and did not result in correlations with either academic or executive function measures.

Children who did not perform well on the standardized achievement measures at age six also failed initial levels of Tower of Hanoi, the task used to assess complex executive planning, at age four. Conversely, children who achieved higher on Tower of Hanoi at age four attained a five point increase on the WJ-III Math Fluency Subtest, $F(2,101) = 3.04, p < .05, n^2 = .06$. Similarly, children who scored higher on the FIST, which measured shifting, at age four also scored significantly higher on the WJ-III Math Fluency Subtest, $F(2,99) = 8.49, p < .001, n^2 = .14$, at age six. Children who scored higher on the Shape School task at age four also scored higher on the WJ-III Math Fluency Subtest at age six, $F(2,96) = 7.04, p < .001, n^2 = .13$. Together, these results indicate the positive influence of executive function on early math.

Similar significant relationships were observed between clinic-based measures of

executive function at age four and later teacher-rated math performance. This observation was true on the tasks that assessed shifting and inhibitory control. However, children who were perceived by their teachers at age four as having difficulty with executive function scored one standard deviation lower than their peers on the WJ-III Math Fluency Subtest. Similarly, rated by their teachers at age six as having below average math scores were also identified at age four as having executive function difficulties. Items on the BRIEF-P tapping working memory, planning, and inhibitory control were found to be more closely related to later teacher-rated math performance than were items tapping emotional control. These results point to the associations that executive function, specifically planning and monitoring goal-directed activity, switching attention set, and attention have on later math performance.

An examination of the associations between executive function performance measures resulted in measures that reflected a single construct, and were found to account for 48% of the total variance across all measures of tasks. Hierarchical linear regression models were used to examine the association between the composite executive function factor score. Covariates included SES, gender, IQ, and reading level. Results indicated that on the WJ-III Math Fluency Subtest, performance (after accounting for SES) was predicted by overall executive function ability at age four, $t(98) = 6.51, p < .01$, and explaining approximately 30% of the variance in math. Executive function also contributed 7% of the variance when a model included reading achievement, and an additional 3% in a model that included both reading and IQ. Gender was not significant for either association with Math Fluency or with interactions with executive function.

Results further indicated that on the teacher-rated math performance measure,

children's overall executive function at age four significantly correlated with later math achievement, $t(94) = 4.67, p < .001$, with this model explaining approximately 20% of the variance. After consideration of reading, the association remained ($\beta = .26, p < .05$) with an additional 6% of variance in this model. Unlike in the Math Fluency Subtest, when reading and IQ were considered, the executive function composite no longer predicted later teacher-rated math achievement, $t(96) = 1.30, p = .20$. Male children were rated higher by teachers (opposite of results found in Brock et al., 2009) but no gender interaction was found with either executive function or reading achievement. Variance analyses suggested that executive function was as predictive as reading or IQ on the WJ-III Math Fluency Subtest.

The results of this study pointed to the usefulness of considering the relationship between executive function and math performance. This consideration might be helpful when trying to identify students who have difficulty with math. A contribution of this study is the suggestion that executive function might be best understood as a construct, rather than as isolated functions. Another strength of this study was its rich design in that it utilized multiple measures to attempt to make associations between executive function and math. Despite these strengths, some limitations were noted. One weakness was that the results did not provide full support for the constructivist method of developing problem-solving. Some children may need more explicit teaching due to their delayed self-discovery, as evidenced by the association between the early delays in executive function and later difficulty with math. In addition, the authors recognized the difficulty in obtaining independent measures of executive function in young children. Finally, teacher rating lacked standardization and were thus open to bias.

In spite of these limitations, this study contributed to the growing scholarship of executive function, especially in relation to math. Future studies may plan to focus on the exploration of specific neurological regions and how they relate to early math skills. A continuation of this study would also be useful to be able to follow this sample and later provide longitudinal data on longer-term effects of executive function on math performance.

Not only assessing a relationship between executive function and performance, the next study examined the usefulness of an innovative intervention to promote the development of executive function. Röthlisberger, Neuenschwander, Cimeli, Michel, and Roebbers (2012) evaluated the impact of an intervention designed to promote executive function in a small group setting for children in prekindergarten and kindergarten. The executive functions that were examined in this study included working memory, interference control, and cognitive flexibility. The sample included a total of 144 students who had been selected from play-oriented preschool contexts and kindergarten classrooms. The sample included students from 22 school areas with varying degrees of socioeconomic backgrounds. The children were randomly assigned into the intervention or control group. Due to lack of complete measurement points, only 135 students remained in the sample. In the prekindergarten grade, ages ranged from 54 to 67 months, while in the kindergarten grade, ages ranged from 66 to 81 months. No differences in age or gender were noted between the control and intervention groups, and both groups had more boys than girls. Each group contained both prekindergarten (5 year old) and kindergarten (6 year old) students.

The intervention group received a six-week program to promote primarily the

three targeted executive functions. Because of the complexity of executive function tasks, other executive functions were likely tapped as well. A total of 19 tasks, including dimensional card sort, Stroop, and trail-making tasks were implemented twice per week with the experimenter and the remaining three days with trained teachers. Each session lasted 30 minutes. The intervention took place over the course of six weeks. Students in the control group did not participate in any activities outside of their typical curriculum.

Pre- and post-tests, namely the Simple Flanker Task, the Mixed Flanker Task, and the Complex Span Task, were conducted individually with students. The Complex Span Task (Laneman & Carpenter, 1983; Pickering & Gathercole, 2001, as cited by Röthlisberger et al., 2012) assessed working memory and included the dependent variable of total number of correctly recalled trials. The Simple Flanker Task (Roebbers & Kauer, 2009, as cited by Röthlisberger et al., 2012) assessed interference control and included the dependent variables of overall reaction time in milliseconds, overall accuracy in percentage, and conflict score in milliseconds. The Mixed Flanker Task (Diamond, Barnett, Thomas, & Munroe, 2007, as cited by Röthlisberger et al., 2012) assessed flexibility and included the dependent variables of overall accuracy in percentage and reaction time in milliseconds.

Statistical analysis did not reveal sex differences, leading the researchers to collapse the data across gender. Because a small but significant group difference was found for the pre-test ($p = .04$), pre-test measures were considered covariates. A between-group analysis of covariance was conducted for pre- and kindergarten separately. Regarding intervention effects in 5-year-olds, on the working memory task, the pre-test level was significantly related to the post-test performance, $F(1, 68) = 12.5$,

$p < .01$, $d = .89$. After controlling for pre-test level, a significant effect of group on post-test performance was found, $F(1, 67) = 3.0$, $p < .05$, $d = .42$. These results indicated that students in the intervention group performed better on the working memory task than the students in the control group. With reference to the Interference Control Task, the pre-test level was found to be significantly related performance of overall reaction times on the post-test, $F(1, 66) = 10.1$, $p < .01$, $d = .78$. Unlike the working memory task, no significant group effect was found. On the task examining flexibility, the pre-test level was significantly related to the post-test shifting reaction times performance, $F(1, 67) = 12.8$, $p < .01$, $d = .88$, as well as shifting accuracy, $F(1, 66) = 8.4$, $p < .01$, $d = .59$. A significant group effect of shifting accuracy on post-test performance was also found after controlling for pre-test level, $F(1, 66) = 5.8$, $p < .01$, $d = .59$. These findings indicate that children in the intervention group significantly outperformed children in the control group on shifting accuracy.

With respect to intervention effects in the 6-year-old group, for the working memory task, the pre-test level was found to be significantly related to performance on the post-test, $F(1, 61) = 44.3$, $p < .001$, $d = 1.71$. A significant group effect was not found. For the Interference Control Task, the pre-test level was significantly related to post-test overall reaction times, $F(1, 59) = 48.3$, $p < .001$, $d = 1.81$, as well as conflict reaction times, $F(1, 60) = 13.7$, $p < .001$, $d = .97$. On accuracy, a significant group effect was found, $F(1, 61) = 2.8$, $p < .05$, $d = .43$. These results indicate that students in the intervention group were more accurate in their overall performance on the Interference Control Task than students in the control group. On the task examining flexibility, pre-test levels were found to be significantly related to shifting reaction times in post-test

performance, $F(1,60) = 25.0, p < .001, d = 1.29$. A significant group effect was not found.

Teachers reported an overall satisfaction with the intervention, its training, and in the developmental appropriateness of tasks. Despite this satisfaction, teachers were concerned with the number of task, the difficulty of implementation while simultaneously supervising a larger group of students, and with difficulty of integration into the curriculum. Rating on a scale of 1 to 7, with 7 being “very well”, six trained experimenters rated the intervention as “good” in terms of how well it could be integrated into the classroom settings.

This study is significant in that it is the first intervention documented to experimentally induce increases in executive function in young children, specifically with the group effect found for interference control in the kindergarten group. Because the sample consisted of typically developing students, the findings from this study are more generalizable. Moreover, these results support previous suggestions to include specific teaching of executive function in early childhood (Diamond, Barnett, Thomas, and Munroe, 2007).

Despite these strengths, this study has its limitations as well. Because the design emphasized accuracy but not speed, reaction time group effects may not have been found. Lack of this and other training effects may have been due to differing levels of brain maturation before the intervention took place. The study did not control for differences in parental practices or schooling experiences, and either or both could have accounted for some variance. The authors acknowledged that children’s performance may have been influenced by the Hawthorn effect, as the children may have been performing in

response to receiving attention. These results were limited to the pre- and kindergarten years, and the study did not include follow-up measure into future grades. Nonetheless, these results should cause the educational community to embrace teaching young children skills of executive functions explicitly and not rely on their implicit acquisition. Future studies should consider including a heterogeneous sample of risk levels to make broader comparisons. In addition, future studies should consider a longer study period to follow students into higher grades to make broader inferences about the effect of interventions.

The studies reviewed in this section all examined the relationship between self-regulation or executive function and academic achievement. While some studies focused primarily on the specific role that executive function plays in the fostering of later academic skills and performance, all of the studies highlight the importance of paying attention to executive function in early schooling. Combined, the results of these studies suggest a positive relationship between self-regulation or executive function and academic achievement. Researchers have also examined other factors to assess their influence on the development of self-regulation or executive function. These factors include private speech and dramatic play, two significant components of the Tools of the Mind curriculum.

Components of Tools of the Mind

While research on the actual curriculum is scant, two components of Tools that have been studied in relation to their influence on the development of self-regulation are private speech and dramatic play. These components are grounded in Vygotskian (1978) theory and are essential to the theorized gains children can make in the development of

self-regulation through the Tools curriculum.

Private Speech

Private speech has not only been a hallmark component of Tools of the Mind, but also been studied as a primary method of developing self regulation. Krafft and Berk (1998) examined the use of private speech in preschools and studied its specific impact on open-ended activities and make-believe play. In this study, two preschool settings were observed: Montessori, which stressed more closed-ended table activities and traditional, which offered more play activities. The consideration of educational settings is instrumental to the investigation. According to Vygotskian theory, both private speech and made-believe play allow for development to progress and to foster self regulation. The authors hypothesized that children in the traditional setting would use more private speech as they would be provided with more opportunities for play and that open-ended tasks would result in more private speech than closed-ended tasks. In addition, the authors hypothesized that the more external regulation that took place (by the teacher), the less private speech would be observed.

The sample in this study consisted of 59 children aged three to five. There were 20 three-year olds, 19 four year-olds, and 20 five-year olds. Among them, there were 24 boys and 35 girls. Thirty were from Montessori, and 29 were from traditional preschool. Most were from families of middle SES. Within the sample, 81% of the participants were Caucasian, one child was African American, and two children were Asian American. Both groups were from cities in the Midwest. Data gathered at the outset revealed that children from both groups were comparable in SES, but the children from Montessori performed slightly higher in vocabulary measures, $M_s = 12.7$ and 10.1 , F

(1,53) = 9.7, $p < .01$, indicating an advantage over traditional group. To control for this factor, verbal ability was used as a covariate in data analyses.

Data were collected through individual observations of each child. Data collected related to type of play (functional, constructive, fantasy, unoccupied/onlooker, and transition), goal of the activity (open-ended, closed-ended), adult involvement (direct involvement, watcher/helper, uninvolved), peer involvement (solitary, parallel, associative, cooperative) and private speech/social speech (affect expression, word play and repetition, fantasy play speech, describing one's own activity and self-guidance, inaudible muttering, and other). Intercoder agreement averaged at .86, achieving high intercoder reliability for all observational categories, making the results more reliable. Coders were also blind to the purpose of the study and observations occurred in random order, further increasing reliability.

The results of this study indicate that private speech was fairly prevalent, occurring on the average in nearly 27-34% of sampling intervals. Fantasy play speech occurred most frequently, $M = 7.9$ and word/play repetition happened the least, $M = 2.9$. No gender differences were observed. As hypothesized, more open-ended tasks were observed at the traditional preschool than at Montessori, and consequently, children in the traditional preschool ($M = 12.6$) displayed three times as much fantasy play as did the children in Montessori ($M = 3.5$). In the traditional preschool, greater availability of open-ended tasks, fantasy play, reduced teacher direction, and great activity involvement contributed to the overall result of twice as much private speech being emitted than at Montessori. Children in the traditional preschool also displayed more describing own activity/self-guidance, the task most indicative of self regulation. The authors suggest

that closed-ended activities prevent children from constantly challenging themselves and setting new goals, in addition to less need for self regulation with higher levels of teacher direction. One implication of this study is to increase understanding of the regulatory cognitive processes that occur when a child is allowed to problem solve and set goals in line with his or her own developmental trajectory, which may inform the manner in which teachers *direct* their classrooms.

Although informative, results obtained in this study may not generalize outside of the classrooms studied. The mere parental choice, which can be quite purposeful in selection of educational settings, may indicate that children who attend Montessori versus traditional preschools may differ in ways not identified in this study. This difference may compromise the results as the two groups may not have been equal at the beginning of the study. Nonetheless, this study, unlike similar research, occurred in natural settings as opposed to laboratories. This study aligns well with the theory the Tools is based on and supports the use of private speech and make-believe play as self regulators in Tools classrooms. Future studies may further investigate other components of Vygotskian theory by including the supporting role of scaffolding in the development of play and private speech to promote self-regulation.

Similar to the self-regulatory function of private speech studied by Kraft and Berk (1998), Fernyhough and Fradley (2005), examined the function of private speech while completing executive tasks. Based on the Vygotskian (1934/1987) premise that children will use private speech more as a task becomes more difficult, Fernyhough and Fradley (2005) sustained the notion that private speech is ineffective if the task is either too easy or too difficult. If the task is not within the child's ZPD, private speech is not likely to

occur. A second premise this study was based on is that using private speech contributes to the development of self-regulation (Vygotsky, 1934/1987). This study investigated whether performance improved with the use of private speech due to the assumed increase in self-regulation.

The participants in this study consisted of 46 children between five and six years old. Twenty-one were female and 25 were male, with an average age of 71.2 months. The sample was derived from two primary schools in the English Midlands, with both schools' demographics being similar. All of the children were White, with the exception of one of mixed race. Two separate testing sessions took place one week apart. During each testing session, the children had to complete four trials of the Tower of London (ToL; Bull, Epsy, & Senn, 2004, as cited by Ferneyhough & Fradley, 2005) task.

Sessions were videotaped and viewed for coding of activity that was task-relevant, of social and private speech, and of performance. Utterances during the task were classified as either social or private. Social utterances were coded for eye contact, behavior, content markers, and temporal contiguity. Private utterances were categorized into three levels. Task-irrelevant speech or affect expressions were considered Level 1 (PS1); task-relevant externalized private speech was considered Level 2 (PS2); and "task-relevant external manifestations of inner speech" (p. 110) were considered Level 3 (PS3). Inter-rater reliability for the four types of speech (social, PS1, 2, and 3) was $\kappa = 0.80$. Teachers were also asked to rate each child for talkativeness on a scale of 0-10. In addition to the ToL task, receptive verbal ability was assessed after the second trial of ToL using the British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Pintile, 1982, as cited by Ferneyhough & Fradley, 2005). Task performance was measured using

moves-to-solution (*MTS*) and time-to-solution (*Time*). Total Move Values and Total Times were also scored.

In the rate of social and private speech, no significant differences were observed. At least one utterance of private speech was produced by 89% of children over both sessions. Very low incidences of task irrelevant private speech were observed, suggesting to the authors that the difficulty in the task lead to increased use of task-relevant private speech. All of the correlations between teacher ratings of talkativeness and the different types of speech were positive, none was significant ($r's_{[44]} = .13-.15$), with the exception of task-relevant externalized private speech ($r_{[44]} = .37, p < .05$). Task-relevant external manifestations of inner speech were not correlated between sessions, leading the authors to suggest that children were no longer using this form of private speech in Session 2 if they used it in Session 1.

Results of the ToL task did not reveal overall practice effects that would have been caused by repetition of same task at Session 2, $F_{[1,44]} = 2.75$ in the 2 (Session) x 4 (Task Difficulty) repeated measures ANOVA on the *MTS* scores. A significant main effect was found of Task Difficulty, $F_{[3, 132]} = 81.39, p < .001$. As task difficulty increase, the mean *MTS* increased linearly, $F_{[1,44]} = 237.65, p < .001$. Similarly, a 2 (Session) x 4 (*Time*) repeated measures ANOVA revealed significantly faster solving in Session 2, $F_{[1,44]} = 42.76, p < .001$. With reference to private speech and task difficulty, and to test the hypothesis that a quadratic relationship existed between self-regulatory private speech and task difficulty, a 2 (Task Difficulty) xx 4 (Speech Type) ANOVA was conducted, using the rate of speech as the dependent variable. No main effects were noted of Session, Task Difficulty, and Speech Type. Additionally, there were no interactions.

These results suggested that when tasks were of moderate difficulty, all forms of speech peaked. To test whether private speech is most relevant when tasks are within a child's ZPD, for each of the four difficulty levels, separate correlation coefficients between private speech and *Time* were computed. No significant correlations were found. The lack of significance in these latter correlations does not support the relationship between ZPD and use of private speech. However, when each task-item was broken down and analyzed separately, some support was found. When the task was easy (2 moves), children used mostly silence/success, while for the most complex tasks (5 moves), children used relevant/failure. When the task was of moderate difficulty, (3-4 moves), children used relevant private speech and were the most successful.

The results of this study contributed to the theoretical understanding of the Vygotskian principles employed in private speech, namely that all children appear to use some form of regulatory private speech. Self-regulatory private speech, rather than social speech, was correlated with simultaneous task performance. Some limitations were present, however. In the non-contrived setting of the typical classroom, children may not have been able to be as task relevant with their private speech had they not been 1:1 with the evaluator. Larger scope longitudinal studies might provide additional information about the relationship between private speech and task difficulty. Such longitudinal research may also improve understanding of why some children do not demonstrate private speech and whether some children demonstrate external self-regulatory private speech as task become more difficult. This last pattern does not necessarily align with the Vygotskian body of knowledge regarding private speech, but would enhance this knowledge.

Consistent with the focus on self-regulation, Winsler, Manfra, and Diaz (2007) examined the behaviors of children during a task performance and the influence of private speech among five year old children. In this study, Winsler et al. (2007) examined two groups of children, one with behavior problems and one without, and observed them during a task after manipulating their speech. The authors also explored which conditions were most helpful to children with the use of private speech. Participants were drawn from an earlier longitudinal study involving two similar groups. The sample consisted of 72 five year old children in an urban area of Northern California. All of the participants attended some form of preschool, and at the time of the study, 68% of the sample was already in kindergarten. While considerable variation existed in socioeconomic background, there were no significant group differences on any demographic variables at the time of recruitment. In both groups, males were represented more than females.

Both groups were given a motor task, the speech-action coordination “hammer” task (Balamore and Wozniak, 1984; Luria, 1961, as cited by Winsler et al., 2007). The first part of the task involved tapping a sequence of colored pegs upon instructions, followed by prompts to either speak the steps as they were completing them and then not to speak the steps out loud. The second part of the task involved hitting the peg a number of times. Children were rated on how long of a sequence they could complete correctly, as well as whether they followed the verbal instructions to use private speech and to stop private speech.

Analysis of spontaneous use of private speech, the results revealed that 79% of the children used some private speech during the phase of the sequencing task when no

instructions to do so were given. Results indicated that 21% of the children remained silent the entire time. The authors found that 93% of the children who had been identified as behavioral spoke to themselves at least once, in comparison to 70% of the typical children, ΦR effect size = .28. Similarly, in the trials that were completed with speech without instruction, the average proportion was .56 (behavioral = .69, typical = .47). These results indicate that children with externalizing behaviors used private speech during a motor task more than children who did not have those behaviors. Similar differences were not found for the counting task.

Moreover, when told to use speech children in both groups performed better than when not instructed to do so, whether or not they used it, $F(1,69) = 4.16, p < .05$. Children were also found to talk more when they were instructed to do so than when they were not instructed, revealing a significant effect, $F(1,69) = 7.19, p < .01$. These data support the encouragement of private speech as a strategy to complete tasks with better control. As the authors point out, many children with behavioral problems may already speak out or to themselves, and this behavior may lead teachers to intervene to make the classroom quieter. However, as suggested in the study, using private speech actually helps to maintain control of the task. Therefore, teachers should understand the importance of such a strategy and not insist on silence. This implication may greatly inform classroom management practices and the need for early intervention, especially with at-risk students. Furthermore, this study supports the use of private speech as a self-regulator as suggested by Tools.

Despite the above implication, this study did have some limitations. The sample was drawn from an existing study, leaving open the possibility that parents (thus family

environments) who chose to allow their children to participate may be different than those who did not allow participation. The participants also represented many types of preschool and childcare experiences. While many demographics were controlled, the type of preschool environment and years in preschool were not controlled. Preschools and childcare centers offer varying degrees of qualification of teachers, quality of instruction, structure, etc., opening up the possibility that these children were not the same in this regard. In addition, 21% of the children with behavior problems were in a single parent household, compared to only 5% of the control group. This difference could pose additional questions as to how the two groups differed prior to the task data collection. On the other hand, such variation in childcare experiences may be representative of the general population, and because positive results were found for the use of private speech in both groups, these results may be generalized beyond that community in California. In other words, regardless of background, using private speech may be helpful as a method of self-regulation during completion a task such as sequencing.

With a specific focus on children with Attention Deficit Hyperactivity Disorder (ADHD), Corkum, Humhries, Mullane, and Theriault (2008) studied the role of private speech in problem-solving and inhibition tasks. Based on the ADHD research conducted by Barkley (1997), as cited in Corkum et al., (2008), four executive functions necessary for self-regulation of behavior include working memory, the ability to synthesize information for problem-solving, self-regulation of emotional states, and internalization of speech (p. 98). This study focused on the last function, private speech, and its role during tasks that require inhibition and attention.

The sample in this study included thirty-two children between the ages of 6 and 11. Sixteen children (one girl and fifteen boys) were diagnosed with ADHD and comprised the ADHD group, while thirteen boys and three girls comprised the control group. None of the children in the sample had IQ ratings lower than 80 or other serious developmental and biological impairments. Children in the control group were matched with children in the ADHD group on age and grade.

Multiple types of measures were used to obtain necessary data. Parents filled out demographic questionnaires to allow the researchers to compute each family's socioeconomic status, as well as to ensure appropriate matching. The Conner's Parent Rating Scale-Revised (CPRS-R; Conners, 1997, as cited by Corkum et al., 2008) was used to confirm the diagnosis of ADHD. Intelligence was measured using the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991, as cited by Corkum et al., 2008). Specifically, the Vocabulary and Block Design subtests were used to estimate an intelligence quotient. Problem-solving tasks included the Picture Arrangement and Object Assembly subtests of the WISC-III. These tasks were used to attempt to elicit private speech and were not included in the quotient of intelligence. To measure sustained attention and inhibition, the Conner's Continuous Performance Test II (CPT-II; Conners, 2000, as cited by Corkum et al., 2008) was used. The CPT allowed the researchers to create a variable that assesses the relationship of private speech and performance. To code private speech, frequency counts were used to measure private speech, social speech, and uncodable speech after viewing videotaped sessions of the CPT tasks. On the day of the tasks, children with ADHD were required to be medication-free for previous 24 hours.

Demographically, the two groups did not differ on age but did differ on IQ ($t(30) = -2.53, p < .02$) and SES, $t(30) = 2.91, p < .007$. While the ADHD group had significantly lower IQ scores ($M = 98.19, SD = 15.07$), the mean estimated IQ scores for both groups fell within the average to high average range. Confirming the presence of ADHD symptoms, the ADHD group had significantly higher T -scores on subscales of the CPRS-R. The amount of private speech did not appear to be related to language ability. ANCOVA was used to compare the two groups on their accuracy on the problem-solving task and their T -scores for mean reaction times on the CPT-II. In this measure, higher T -scores indicated lower performance. Estimated IQ and SES were covariates. Results indicated that the two groups did not differ significantly on the problem-solving task, $F(1,28) = 2.13, p = .16$. Children with ADHD performed poorer on the CPT-II ($F(1,28) = 16.98, p < .001$).

The three types of private speech were examined using two separate MANCOVAs, with group (ADHD and Controls) as the between-subjects factor and the three private speech dependent variables. Results indicated that for both the problem-solving task and the CPT-II, several of the private speech variables were significantly correlated. Multivariate analysis revealed that more task-relevant external private speech was used by the ADHD group ($M = 8.06, SD = 8.67$) than the control group ($M = .31, SD = .60$), $F(1,29) = 5.68, p < .02$, and more task-relevant internal private speech ($M = 19.13, SD = 9.99$) than the control group ($M = 11.13, SD = 8.53$), $F(1,29) = 5.97, p = .02$. These results indicate that children with ADHD could be using this strategy more as a way to self-regulate on tasks that require more effortful control.

This study demonstrated the importance of the use of private speech to enhance

self-regulation, with particular emphasis on inhibition. This research may be beneficial to inform educators that children, in particular children with ADHD, may use external private speech as a way to regulate. Having this understanding, educators may view external private speech in a more positive way, rather than an example of disruptive behavior. Interventions to provide direct instruction on the functions of private speech should be considered.

While this study may have practical implications, limitations were present. The small sample size of this study limits the ability to generalize the findings outside of this sample. While some factors were controlled during the matching process, some factors were not, such as estimated IQ, parental income and education levels. Within the design, however, SES and IQ were covariates in an attempt to control for these variables. A final limitation of this study is related to the coding of private speech. According to the method prescribed in this design, private speech was measured while observing tongue and lip movements. Internal private speech can also occur without these observable movements, which this study was not able to account for. Additional research is needed to increase the capacity to study private speech. Despite these limitations, this study contributed to an understanding of the importance of private speech, especially when these verbalizations might be interpreted as negative behaviors.

While the last study reviewed suggested that educators re-interpret children's expressions of private speech, the next study suggested an assessment of the association between private speech and problem-solving. Daugherty and White (2008) examined the relationship between private speech and creativity with at-risk children. This study also investigated if that relationship varied across task and context. The sample for this study

consisted of 32 preschool children from the Southeast enrolled in a child care center that served Head Start and public preschool programs. Of the 32 participants, 18 were from Head Start and 14 from preschool. The average age of the participants was 4.6 years. Twenty-six children were identified as African American and six as Caucasian. Out of the sample, 24 were identified as at-risk, which was related to family income and public assistance.

To measure creativity, the Torrance creativity test (TCAM; Torrance, 1971,1981, as cited by Daugherty & White, 2008) was used. Four activities that comprise the TCAM were designed to measure creative thinking, and children were able to respond either verbally or non-verbally. Children that demonstrate alternative ways than the ones taught to solve the problem were viewed as having higher levels of creativity. Activities 1, 3, and 4 were related to creative measures of fluency and originality, while Activity 2 was related to imagination. Private speech was coded into seven categories [fantasy play, self-direction (task oriented), emotional release, humming and singing, inaudible muttering, random and isolated, and reading aloud].

Data collection occurred in two settings, the classroom open play area and the structured task situation (logical-mathematical activities). Children's private speech was observed and tape recorded. To increase reliability, scoring was conducted by Scholastic Testing Service and produced standardized scores of creativity ($M = 111.1$), imagination ($M = 94.2$), and originality ($M = 98.9$). Correlation statistical procedures identified several significant relationships. Originality measures ($r = .56, p < .01$) and fluency ($r = .53, p < .01$) increased as self-direction increased. The latter correlation showed that the number of idea responses on the TCAM was related to the quantity of private speech

employed during the math activities. In both contexts, fluency and originality were associated with the grand total of private speech as well as total math private speech. These findings suggest that as more private speech was used, more ideas were generated on the tasks. In contrast, the amount of private speech in open play was not related to either measure. This finding may indicate that in the free play activities, children were not involved in mature play that required problem-solving and creativity.

The findings of this study were similar to those studied in children of higher SES (Daugherty & Logan, 1996, as cited in Daugherty & White, 2008). This similarity supports the premise that private speech may assist researchers in identifying creativity in diverse samples of children. Strengths of this study included the use of the TCAM to measure creativity. The TCAM was used previously with disadvantaged children and had not been found to be biased against race or gender (Torrance, 1971, as cited in Daugherty & White, 2008). In addition, recorders of private speech received training and interobserver reliability was very high (98-100%). Despite these strengths, this study includes some limitations. Generalizability of results was limited by the small sample size. Other factors could have contributed to creativity as well. This study did not include cognitive measures, and aptitude of children may have been a confounding factor. Regardless, this study was able to show that private speech increased the number of ideas a student had, thereby influencing creativity and problem-solving.

Lidstone, Meins, and Fernyhough (2011) were also able to show how private speech increased executive function. This study examined individual differences in the private speech of children across four tasks, namely Tower of London (ToL), digit span, and two measures of spatial IQ. Additionally, this study sought to examine the

consistency of the private speech of children across two time points, and contexts (classrooms versus laboratory).

The sample consisted of 25 typically developing children between the ages of 8 and 10. The sample was derived from three mainstream schools in England, two of which were in areas considered moderately disadvantaged, and the third in an area considered moderately affluent. Thirteen of the 25 participants were girls. The mean age of the participants was 9.4 years.

Tasks were presented over the course of two sessions with a mean time between sessions of 11 months. During the first assessment, eight problems of the ToL task, a digit span task, and two subtests of the British Ability Scale (Recall of Designs and Pattern Construction; Elliott, Smith, & McCullough, 1996, as cited by Lidstone et al., 2011) were all completed in the laboratory. During the second assessment, 12 ToL problems and 10 minutes of numeracy schoolwork of that day were completed, again in the laboratory. A remaining 10 minutes of numeracy schoolwork was completed in the classroom, resulting in three observations. The observers were specifically looking for the planning between the directions and the participants' verbal responses.

One digit span task was completed during each assessment period. Each consisted of a dual task condition (tapping their foot while completing the task or repeating the word "Monday" while completing the task). The digit span task involved repeating a sequence of digits orally in order as presented. Private speech was coded during between the start of the trial and the presentation of the question. On the *Recall of Designs* task, participants had to reproduce an abstract drawing from memory. On the *Pattern Reconstruction* task, participants recreated a two-dimensional block design to

match a picture. Private speech was coded during the period of the trial and ending when the participant announced he or she was finished or timed out. All of the private speech was coded from video recordings of the assessment sessions.

Results indicated that more private speech was used during the ToL and digit span task than the two spatial tasks, $Zs > 3.70, p < .001$. More private speech was also noted at Time 2 on the ToL than Time 1, $Z = 4.11, p < .001$. For numeracy, the rate of private speech did not vary across contexts. Positive correlations were found between the rate of private speech on the ToL at Times 1 and 2, $\rho(25) = .54, p = .01$, as well as between internalization scores on ToL at Times 1 and 2, $\rho(21) = .57, p = .02$. A positive correlation was also found between numeracy completed in the classroom and completed in the laboratory, $\rho(25) = .53, p = .01$, as well as between internalization scores on numeracy completed in the classroom and in the laboratory, $\rho(22) = .58, p = .01$. These results show strong consistency of the rate of private speech production across tasks, time points, and contexts.

This study was significant in that it was the first to find a relationship between private speech production on a memory task and on problem-solving and executive function tasks. This study also demonstrated that the results obtained in artificial laboratory settings are consistent with results in the classroom. In addition, results indicated that children use similar levels of private speech across different tasks. While this study included many strengths and implications for practical applications in early childhood education, some limitations were observed. The internalization scale used produced skewed distribution, and thus, limited statistical analysis to non-parametric measures. The sample size was quite small and demographically limited. Although one

of the schools differed in SES, comparisons regarding SES were not made. Regardless of these shortcomings, having this understanding of how children internalize private speech is important for educational interventions.

A final study related to private speech found associations between private speech, self-regulation, and task performance. Agina, Kommers, and Steehouder (2011) examined children's speech use, self-regulation, and task performance in relation to the absence of external regulators, such as a teacher. This study contrived two non-human conditions for providing instruction to students. The first condition was the Verbal-Gesture-Silent (VGS) and the second was the Silent-Gesture-Verbal (SGV). Feedback, known as Knowledge Response (KR), had three conditions, verbal KR with verbal encouragement, visualization-presentation of KR in the absence of verbal encouragement while gesturing, and no KR. This study aimed to learn how children perform given the conditions of VGS and SGV, and hypothesized that the VGS condition would result in children making more private speech than social speech, as well as showing higher self-regulation, task-performance, and satisfaction.

The sample for this study included 40 children with an average age of 5.4 years from a preschool in Libya. Each condition was assigned 20 of the participants, with equal gender distribution. All of the participants spoke Libyan, and none was identified as being diagnosed with ADHD, autism, or hearing and vision issues. The stimulus was provided by an animated computerized program that did not require any previous training. Tasks were based on the concurrent daily preschool activities. Three types of instruction were contained in the game (verbal, gesture, or silent). In the verbal condition, children would hear that their answer was either correct or incorrect. In the

gesture condition, children would see a happy or sad face as feedback. No voice or visual feedback was provided in the silent condition. Speech utterances were observed and coded as private or social. Self-regulation was measured by the computer's score of manifested self-regulation based on the simple/complex tasks. Task performance was scored as either correct or incorrect. Satisfaction was measured through a chat with the characters from the computer program.

Sessions were video and audio recorded. Analysis of variance was used to determine significant condition effects. Results indicated that a significant condition effect, $F(2,37) = 25.64, p < .01, \eta^2 = .80$, existed suggesting that the VGS condition yielded higher performance than the SGV condition. In addition, the VGS condition resulted in more private speech. During the verbal unit, "23% of private speech with no utterances of social speech for the VGS condition" was found, and ".08% of private speech with 16% of social speech for the SGV condition" (p. 1123) was found. Higher degrees of self-regulation, task performance, and satisfaction were also observed with the VGS condition.

Results from this study suggested that children performed better, used more private speech, and had higher levels of self-regulation when the encouragement they received started verbally, followed by gestures, and ending in silence than when the encouragement began with silence, followed by gestures, and ending with verbal. When the encouragement began with silence, children continued to talk to themselves, but this speech was viewed as social speech and not task related. The type of encouragement used during performance can affect children's speech use. This study adds to the existing body of literature on private speech, as well as to the importance of how teachers scaffold

the development of private speech and self-regulation.

Some limitations were noted in this study. The relatively small sample size limits the ability of this study to be generalized. While the child's satisfaction was an important variable to measure, satisfaction as a self-report was obtained from subjective data. The design of this study and its computerized nature should be considered strengths. Most of the existing data collected on private speech has been collected via observations and psychometrics. This study included technology that can obtain data on private speech in the laboratory, allowing for some researchers to conduct empirical research that may not otherwise be possible. At the same time, laboratory environments are not always conducive generalization due to their artificial nature. Future studies should focus on the degree of encouragement by teachers and whether some levels of encouragement actually inhibit student progress.

The articles in this section have all related to the influence of private speech on self-regulation and executive function. In addition to self-regulation, private speech was also related to creativity and problem-solving (Daugherty & White, 2008). Task performance was also related to private speech (Agina et al., 2011). This relationship is consistent with studies in the previous section that related self-regulation and executive function to academic performance. If self-regulation and executive function positively impact academic achievement, and if private speech positively influences the development of self-regulation and executive function, then private speech should impact task performance. The next section will include a review of literature on the effects of dramatic play, a second component of Tools, on academic outcomes.

Dramatic Play

According to Vygotsky (1930-1935, 1978), sociodramatic play increases a child's ability to internalize social norms and increase self-regulation. Several studies have attempted to demonstrate this relationship by investigating children, educators, child development centers, and their curricula.

Comparing impulsive children and non-impulsive children, Elias and Berk (2002) searched for differences in quantity and maturity levels of sociodramatic play. This study also attempted to find a relationship between play and subsequent self-regulation for both impulsive and non-impulsive children. Other relationships considered included verbal ability, age, and gender with respect to play, self-regulation, and temperament.

Fifty-one children were observed over two time periods, early Fall and late Winter/early Spring. Participants came from two traditional day care programs in a city in the Midwest, and included 24 three-year olds and 27 four-year olds from four classrooms. Classrooms were similar in class size, teacher-student ratio (1:9/10), gender, and curricula. All participants were Caucasian, English-speaking, and middle- to upper-middle class.

Observations occurred at both centers focusing on the play centers of housekeeping and blocks. Children were instructed to stay at the chosen center during that activity's time. To measure play and self-regulation, the Smilansky Scale (Smilansky and Shefatya, 1990, as cited by Elias & Berk, 2002) was used, coding complex sociodramatic (CSD) play and solitary play. Maturity of play was assessed by observing five elements of play, namely imitative role play, make believe with objects, make believe with regard for situations and actions, interaction, and verbal communication (Elias & Berk, 2002, p. 223-224). Persistence in play episodes was also

coded. In 30-second observation intervals, each element was recorded as present or not present. This procedure resulted in a total frequency score for sociodramatic play and a persistence score.

Self-regulation was measured during observations of the clean-up and circle time periods. The clean-up period allowed observers to see how the children demonstrated the ability to follow the clean-up routine, while the circle time period allowed the observers to measure how the children displayed attention. Teacher ratings of children's play and self-regulation as well as parent assessments of temperament were included. Assessment of temperament included ratings of impulsivity and aggression. Verbal ability was measured using the Vocabulary subtest of the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R; Wechsler, 1989, as cited by Elias & Berk, 2002). Including verbal ability scales was important due to the possibility of variance in performance in dramatic play being a function of verbal ability.

Based on t-test measures, differences were not found between the two day care centers on any play, self-regulation, or verbal ability measure, making both settings comparable. Results showed that verbal ability was significantly related to total sociodramatic play frequency, $r=.35, p<.01$, and to CDS play frequency and persistence, $r=.35, p<.01$ and $r=.42, p<.01$. Verbal ability was not related to self-regulation measures. Relative to age and gender, two-way ANOVAs were performed and resulted in age being related to observed self-regulation. Three-year olds engaged in lower levels of self-regulation than four-year olds, $M=3.8$ and $4.7, F(1,46)=16.1, p<.01$, and $M=4.3$ and $4.9, F(1,46)=13.0, p<.01$. While statistical analysis did not reveal gender differences, boys were rated as more impulsive than girls by teachers, $M=3.8$ and $2.0, F(1,46)=8.7, p<.01$.

Circle time attentiveness was significantly decreased when higher impulsivity ratings were observed at both testing times ($r = -.39, p < .01$, and $r = -.51, p < .01$).

Correlation analyses did not reveal relationships between CSD play and self-regulation development. However, CSD play frequency and persistence with clean-up were significantly correlated with future clean-up performance at time 2 only, $r = .32$ and $.33, p < .05$. Results also indicated that solitary play was negatively correlated with future clean-up performance, $r = -.33, < .05$. On impulsivity measures, children who displayed high impulsivity also displayed lower levels of self-regulation at clean-up time 1, $M = 3.9$ and $4.6, F(1,21) = 10.1, p < .01$ and less regulation on teacher ratings, $M = 2.5$ and $4.1, F(1,21) = 13.8, p < .01$. These differences did not appear at the second testing time, supporting the role of mature play in the development of self-regulation.

Results of this study have practical implications for the development and implementation of early childhood curricula that foster CSD play, especially to target children who are highly impulsive. The design used in this study enabled the researchers to measure levels of play at different points in time. The inclusion of verbal ability as a factor was an important one, as some children may not participate in collaborative play because their verbal ability is compromised. In addition to the Vygotskian variables of pretend play and self-regulation, other variables, however, may have influenced the results of this study. For example, the prior ability for self-regulation was not measured. With a relatively small sample from two very similar settings, the results of this study may not generalize to other settings. Future studies should include larger and more diverse samples. In addition, future studies should include a closer investigation of the specific components of play that increase self-regulation.

Continuing the investigation of the relationship between play and self-regulation, the next study included a distinction between cognitive and emotional regulation. Fantuzzo, Sekino, and Cohen (2004) examined the role of interactive peer play in the development of cognitive and emotion regulation, as well as language, during the preschool years. This study also compared assessments of peer play conducted at the beginning and end of one year.

Participants included preschool-aged children from a large Head Start program in a large urban city in the Northeast. Out of the population of the almost 5,000 children enrolled in this program, two samples were drawn. The first sample of 242 children was the focus of the study of the relationship between play and other competencies. Children in this sample had a mean age of 59.2 months and were 50.2% male. During the following year, the second sample, that included 746 children, was the focus of the comparison of play at the beginning and end of the year. Children in the second sample had a mean age of 55.2 months and were 47% male. The majority of the children in both samples were African-American (73%), Hispanic (14%), Caucasian (8%), and Asian-American (5%).

To measure interactive peer play, the Penn Interactive Peer Play Scale (PIPPS-T; Fantuzzo, Coolahan, Mendez, McDermott, & Sutton-Smith, 1998, as cited by Fantuzzo et al., 2004) was used. This scale described peer play interactions and helped researchers distinguish children with positive peer interactions from other children. Three dimensions were reported: Play Interaction, Play Disruption, and Play Disconnection. Classroom self-regulation was measured by Emotion Regulation and Autonomy on the California Child Q-Sort (CCQ; Block & Block, 1969, as cited by Fantuzzo et al., 2004).

The Emotion Regulation criterion described flexibility and emotion modulation. The Autonomy criterion described the initiative, agency, choice, and self-determination of each child. To measure emotional and behavioral adjustment, researchers used the Adjustment Scales for Preschool Intervention (ASPI; Lutz, 1999; Lutz, Fantuzzo, & McDermott, 2002, as cited by Fantuzzo et al., 2004). This scale consisted of teacher ratings of emotional and behavioral adjustment. Receptive vocabulary was assessed using the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997, as cited by Fantuzzo et al., 2004). Classroom learning competencies were assessed using the Child Observation Record (COR; High/Scope Educational Research Foundation, 1992; Schweinhart, McNair, Barnes & Larner, 1993, as cited by Fantuzzo et al., 2004), which was used to observe thirty classroom competencies.

Data were analyzed using canonical variance analyses. To determine the relationship between PIPPS-T and other competencies, bivariate and canonical analyses were conducted. In addition, canonical variance analyses between self-regulation and PIPPS-T were conducted. The relationship between PIPPS-T and receptive vocabulary was also analyzed.

Results showed a significant bivariate relationship between receptive vocabulary and play interaction ($r = .27$). Between receptive vocabulary and play Disconnection, a significant negative bivariate relationship was found, with no significant relationship found between receptive vocabulary and Play Disruption.

With reference to classroom relationships, classroom peer play behaviors and dimensions of self-regulation were significantly correlated, Wilks' Lambda = .70, $F(6, 434) = 14.11, p < .0001$. A positive relationship between Play Interaction and Autonomy

($R_c = .43$) and a negative relationship between Play Disruption and Emotional Regulation ($R_c = -.37$) were found. Higher levels of emotional regulation were found in children who displayed lower levels of Play Disruption.

Of the measures across time, three significant variate pairs emerged: Social Agitation, Social Isolation, and Social Resistance. Play Disruption was associated with aggression, ($R_c = .64$). Disengagement in play on the PIPPS-T was associated with Play Disconnection, with a strong negative relationship between Isolation and Play Interaction ($R_c = -.91$). Play Disconnection was also associated with Opposition ($R_c = .05$). A strong significant relationship was found between the Play Interaction dimension of the PIPPS-T and classroom competence measures of the COR ($R_c = .64$). In addition, peer play interactions were positively related to other school readiness competencies.

These results supported existing scholarship (Bodrova & Leong, 2007), that children who display more competent self-regulation are the children who also display higher levels of play interaction. Strengths of this study include the types of measures used, all with high levels of validity and reliability. This study is one of the few studies that have made connections to other readiness competencies. On the negative side, teacher reports, which are subjective, were the only measure of classroom constructs. Overall, this study contributes to the knowledge of the role play plays in the development of social competencies and self-regulation in early childhood education.

Similar to the study conducted by Fantuzzo et al., (2004), Nicolopoulou, Barbosa de Sá, Ilgaz, and Brockmeyer (2010) found that social pretend play promoted the development of social competencies in addition to promoting the development of cognition and language, as proposed by Vygotsky (1978). In addition to Vygotsky's ideas

on social pretend play, Nicolopoulou et al. added Vivian Paley's (1990) storytelling/story acting practices to review their impact on cognition, narrative, and social skills.

Over a two-year period, storytelling and story acting activities were observed in seven full-day child care centers. Participants from diverse ethnic backgrounds and aged 3-5 years comprised the sample of 147 children. The sample was 48% White, 22% Hispanic, and 19% African-American. In addition, 63% of the participants came from families identified as low-income. The location of this study was in a city in the Northeast of the United States. Two groups were established for the study. Eighty-one students were in the experimental, storytelling/acting group, while 66 students were in the control group.

Observations were conducted twice per week. Storytelling, although supported by teachers, was child-directed. Pre-and post- tests measured expressive vocabulary, narrative skills, emergent literacy abilities, and social competence. Children in the experimental group composed a total of 553 stories over the two year study period.

Data analysis included a "mixed repeated analysis of variance (ANOVA), Condition x Semester x Age x Year" (Nicolopoulou et al., 2010, p. 48). Overall, results indicated that storytelling promoted narrative, emergent literacy, and social competence skills. Of significance was the interaction between Condition and Semester. Significance was found at the $p < .05$ level on interactions with total score narrative comprehension, print and word awareness, and inhibition, and at the $p < .001$ level on single picture, assertion, and disruption. Mean differences were significant for the following variables: print and word awareness means: experimental = 4.95 to 7.38; control = 5.30 to 6.27; inhibition means: experimental = 7.91 to 10.35; control = 6.12 to 5.50; and decreased

disruption means: experimental: 1.81 to 1.53; control: 1.26-2.02. The decreased disruption means showed that in the control group, with no storytelling/acting, disruption did not decrease, but rather it increased.

The choice of participants was a source of strength for this study. Unlike the study conducted by Paley (1990, as cited by Nicolopoulou et al., 2010), that focused mainly on middle class Caucasian children, this study increased the diversity of the sample to be more representative of the population of the United States, and thus increasing the ability to generalize these findings. The study also yielded an unexpected result. During the observations, the researchers noted that children who appeared to have behavioral difficulties appeared to participate more in storytelling/acting. This activity of play appeared to increase social competence not only of children who already had some social competence, but of children who did not. This finding is relevant as it demonstrates how this activity can increase social competence and increase the inclusion of behaviorally at risk students in classroom activities. More research, however, is needed to corroborate these findings.

Taking a different approach, Hanline, Milton, & Phelps (2008) investigated different forms of symbolic representation involved in sociodramatic play to examine if they predicted reading and math in the early elementary grades. In this study, the researchers included children with and without disabilities. Participants for this study were drawn from another longitudinal study that included six play observations over a three-year period. Of this original group of 117 children, 51 (28 male) participated in this study. Twenty-two of the children were identified as receiving special education services for a variety of disabilities. The study was conducted at an inclusive child care program

that included preschool and kindergarten programs.

To measure developmental skills in the domains of personal-social, adaptive, motor, communication, and cognitive, the Batelle Developmental Inventory (BDI; Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984, as cited by Hanline et al., 2008) was used. Data on sociodramatic play were collected by researched assistants by video-recording play sessions twice per year. During the play sessions, the adult did would respond to the child but would not initiate interactions. The videotapes were then coded to be analyzed. For follow-up measures, children who had been observed during preschool were assessed after their transition to kindergarten. Two measures of academic skills were used: the Test of Early Mathematics Ability (TEMA-2; Ginsburg & Baroody, 1990, as cited by Hanline et al., 2008), and the Test of Early Reading Ability (TERA-2; Reid, Hresko, & Hammill, 1989, as cited by Hanline et al., 2008).

Sociodramatic play was broken down into three aspects for coding (Rogers, 1988, as cited by Hanline et al., 2008). “Symbolic agent refers to what or to whom the play is directed; symbolic complexity refers to the number and inter-relatedness of schémas used in play; and symbolic substitution refers to the level of concreteness and/or abstractness of children's props” (Hanline, 2008, p. 22). These three aspects served as the dependent variables.

To assess if aspects of sociodramatic play predicted reading and math abilities, hierarchical linear modeling was used. Each child’s highest score on the aspects of sociodramatic play were used, as they were deemed to represent the greatest ability of the child at preschool. In the first model, each child’s regression coefficient was used as that child’s growth rate, as measured by increase or decrease in scores per month. The

intercept for each child was an estimation of each child's reading or math score at eight years old. In the second level of analysis, growth rate and estimated score at age 8 were used as dependent variables to be able to investigate the correlates of scores at different ages. The effect of the three aspects of symbolic play on TEMA and TERA scores and growth rates were calculated. Symbolic complexity was not found to be significant for either math or reading. Play-symbol agent was found to have positive effects in math achievement ($r = 7.037, p = .001$), but negative effects in reading ($r = -3.596, p = .047$). Symbolic substitution was found to have positive effects on both math ($r = 7.037, p = .011$) and reading ($r = 6.985, p = .022$). This final result is of importance as it highlights a specific aspect of play as predictive of math and reading achievement.

Despite this contribution, this study has limitations. Poor generalizability resulted from the small sample size and one location from which participants were drawn. In addition, little discussion occurred regarding the opposite effects on reading and math by the play-symbol aspect. However, the strong relationship found between symbolic substitution and reading and math make this study significant. This finding is consistent with Vygotsky's (1978) theories of development. Specifically, these results align with the assertion that children move from concrete to abstract thinking, and that sociodramatic play may provide a transition between the concrete and the abstract. In this way, play moves the child into higher mental functions. These results are consistent with findings that play promotes cognitive development (Fantuzzo et al., 2004 and Nicolopoulou et al., 2010)

In the previous two sections, studies related to the relationship between self-regulation or executive function and academic achievement were reviewed. Their

combined results show strong support for the positive influence that self-regulation and executive function can have on literacy and math skills. In a related way, the studies reviewing the effects of private speech and dramatic play on the development of self-regulation were also able to show positive relationships. As two essential components of the Tools curriculum, private speech and dramatic play are critical during implementation of the program for the development of self-regulation. In the next section, the review shifts to the few studies available that have focused on the assessment of the curriculum. Because private speech and dramatic play have been found to increase levels of self-regulation, and because self-regulation has been associated with academic achievement, the question remains if the program, Tool of the Mind, will demonstrate positive effects on self-regulation and/or academic achievement.

Tools of the Mind Studies

Despite recent and growing interest in empirical assessment of Tools, very little research has actually been conducted related to its impact on self-regulation and even less so on its impact on academic achievement. Four studies were located, beginning with the original empirical study conducted during the development of the Tools curriculum by Bodrova and Leong (2007). However, because evidence could not be found that this study was peer-reviewed, it was not included in this literature review, but rather in the introduction of this paper. Two studies included research on both the impact of Tools on self-regulation and achievement. Finally, one qualitative study reported on the experience of teachers within the learning and implementation of the program.

With interest in this new approach to early childhood education, Diamond, Barnett, Thomas, and Munro (2007) examined the impact of Tools of the Mind on the

development of executive function. Core executive functions in preschoolers were analyzed, including “inhibitory control, which involves resisting habits, temptations, or distractions; working memory, which involves mentally holding and using information; and cognitive flexibility, which involves adjusting to change” (p. 1387). This study was conducted in an urban school district, and included 18 classrooms initially, and added three in the following year. Teachers and students were randomly assigned (stratified) to either the Tools or standard curricula. Resources and training were allocated evenly across all classrooms and teachers. Participants included 147 preschoolers (62 in standard curriculum and 85 in Tools). The average age of the participants was 5.1 years. Of the total number of participants, 78% of the children came from homes with a yearly income of less than \$25,000.

Diamond et al. (2007) estimated that Tools teachers spent approximately 80% of their day engaging children in activities that promote executive function skills, including using private speech, dramatic play, and aids to increase memory and attention. Executive functions were measured using the Dots (Davidson, Amso, Anderson, & Diamond, 2006) task and a Flanker task, which test all three areas of executive function. In the Dots task, a red heart or flower appeared on the right or left, and the children were instructed to press the same side (congruent) or the opposite side (incongruent), followed by a mixed trial. In the Flanker task, children were asked to focus on an object (triangle) and either the same object appeared (congruent) or a different one appeared (incongruent). These measures were administered at the end of the year in year two of implementation of Tools.

Results of this study did not find gender differences, and found that older children

did better than younger children on some tasks. Statistically significant differences were found throughout the study between Tools classrooms and classrooms with the control curriculum, although on the Dots-Congruent measure, both groups were comparable. On the Dots-Incongruent measure, more trials were answered correctly by children in Tools than the control ($t[142] = 2.62, p < .005, \beta = .28$). On the Standard Flanker task, again, more children responded correctly on more trial from Tools classrooms than from the control group ($t[142] = 2.84, p < .004, \beta = .31$). These last two tasks mentioned taxed the executive function of inhibition. The remaining two tasks, namely the Dots-Matrix and the Reverse Flanker, required more cognitive control as they placed more demands on all three areas of executive function. In the Dots-Mixed task, only 29% of the children from the control group were able to pass the pretest, which was required to continue with the task. A statistically significant difference was found between children passing the pretest (Wald = 5.00, $p < .02$). On the Reverse Flanker task, the children in Tools classrooms performed significantly better than the control ($t [120] = 3.87, p < .0001, \beta = .39$). These results suggest that the more difficult a task became cognitively, the greater the performance difference was between Tools and control. Academic measures were only obtained on students in Tools in year two, and therefore could not be compared with the control group.

Some limitations can be identified with this study. While measures were extensive, there were no before Tools measures. A comparison cannot be made, therefore, between before and after intervention, nor can means be compared at the outset to determine if the groups were similar enough for comparison at the end. The lack of pretest measures weakens internal validity. Another limitation is the oversight on the part

of the research team to obtain academic measures from all the groups (control, year 1, and year 2). Because of this lack of data, this study cannot examine the impact of Tools on academic achievement. Despite its limitations, this study was, however, able to show a positive relationship between Tools and executive functioning. The year of implementation for most classrooms was year two, increasing the fidelity to the program. With greater fidelity, more can be generalized about the impact of the program and less can be speculated about the teacher's familiarity with the program. Proper matching on background variables increases the chance that children performed better due to Tools, although pretests would have strengthened this conclusion. Finally, both programs were new to the teachers, although the control group was more traditional and standard. However, both groups had to learn a new program, making them similar in another way. Implications of this study include informing principles of delivery of education. The more teachers utilize activities that encourage the development of executive functions, the more in control a classroom may be. Similarly, this study demonstrated that a program like Tools can be effective in the development of executive functioning, can be taught, and can be cost-effective. Future studies may attempt to connect the impact of improving executive functioning with improvement in achievement and should include pretest measures to be able to improve internal validity.

In an attempt to continue the empirical evaluation of Tools and make associations with academic achievement, Barnett, Yung, and Yarosz, Thomas, Hornbeck, Stechuk, and Burns (2008) assessed the educational effects of the Tools of the Mind curriculum. Sponsored by the National Institute for Early Education Research (NIEER), this study compared the outcomes of two curricula, Tools of the Mind and an existing traditional

curriculum used in the district, on social behavior, language and literacy growth.

Participants were 3 and 4 year-old preschool children in a New Jersey community of whom 80% of the district qualified for free and reduced lunch and of whom 70% spoke a language other than English at home as their primary language. Seven classrooms were made available by the district for the implementation of Tools and eleven classrooms for the control curriculum. Both teachers and students were randomly assigned to either curriculum yielding 88 children participating in Tools and 122 in the control. Of these participants, 47% were females, 53% were males, and 93% identified themselves as Hispanic or Latino. At entry to the program, parent questionnaires were given to obtain demographic information, which revealed no differences in the family backgrounds of the two groups.

Both classroom and child measures were utilized. Child measures included six different instruments, including the Woodcock-Johnson Applied Math Problems and Letter-Word Identification Tests (WJ-R), Get Ready to Read (GRTR; Whitehurst & Lonigan, 2001, as cited by Barnett et al., 2008), Wechsler Preschool Primary Scale of Intelligence Animal Pegs subtest (WPPSI), Peabody Picture Vocabulary Test-III (PPVT-III), Expressive One-Word Picture Vocabulary Test (EOWPVT), and the Woodcock-Munoz-Revised (WM-R; Woodcock & Munoz-Sandoval, 1996, as cited by Barnett et al., 2008) and Oral Language Proficiency Test (Ballard & Tighe, 1999, as cited by Barnett et al., 2008) at the beginning and end of the year. The last two tests were administered only to Spanish-speaking participants. Finally, at the end of the year only, the Social Skills Rating System (SSRS; Gresham & Elliot, 1990, as cited by Barnett et al., 2008) was completed by teachers on all participants. Classroom assessments to examine various

aspects of the classroom environment consisted of four observation instruments: Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008, as cited by Barnett et al., 2008), Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford, & Cryer, 1998, as cited by Barnett et al., 2008), Supports for Early Literacy Assessment (SELA; Smith, Davidson, Weisenfeld, & Katsaros, 2001, as cited by Barnett et al., 2008), and Preschool Classroom Implementation (PCI; Frede, 1989, as cited by Barnett et al., 2008) rating scale. These measures provided information about the classroom, such as emotional climate, classroom management, instruction style, sensitivity of the teacher, and engagement.

Analyses of the child measures revealed statistically significant effects of curriculum on behavior using the SSRS (effect size of $-.47$, $p < .05$), indicating that Tools yielded less common behavior problems than the control. Results also suggest that Tools was more effective than the control in promoting language development (PPVT-III, effect size of $.22$, $p < .05$; and OLPT, effect size of $.35$, $p < .05$). Analyses of the classroom revealed that Tools classrooms scored substantially higher, with differences observed in quality of literacy environment and instruction, frequency of scaffolding techniques, and teacher sensitivity.

In this study, the authors were able to demonstrate the positive effects on Tools on the experiences of children in the classroom and social development. Of most relevance based on the results is the impact of Tools on social behavior. These results support the premise that self-regulation can be taught.

In addition to being only one of two studies specifically examining the effects of Tools, strength was also found with the design of this study. Analyses were chosen to

control for fidelity and multiple comparisons. A particular strength of this study is that both students and teachers were randomly assigned into treatments. Two-tailed *t*-tests revealed that both groups were not significantly different at the beginning of the year. Due to the high number of families speaking Spanish, Spanish speaking testers were made available as well. Both groups were compensated and coached equally, reducing the threat to internal validity related to resentment by the control group. Other strengths included training of the data collectors, reaching 100% agreement.

Despite much strength, this study does have some limitations. This study was not able to establish a highly confident effect on cognitive development. In addition, the Social Skills Rating System was only completed at the end of the year, and could have improved validity if it were also completed earlier in the year. While a strength was to provide Spanish speaking evaluators, children who were identified as Spanish speaking at the beginning of the year were tested again in Spanish at the end of the year. These children may have lost much of their “school” Spanish by the end of the year and more reliable results may have been obtained by repeating the language assessment that determined which language to administer in that was conducted at the beginning of the year. The population was also very homogenous, almost entirely Hispanic and low-income. Future research may focus on a more heterogeneous population to improve external validity. In addition, future studies may seek to find the relationship between Tools and academic achievement.

Finally, a limitation may be that the study took place in the first year of implementation of Tools. Tools is not just a program; it is an entire curriculum that challenges the traditional, teacher directed teaching style of many existing teachers. In

the first year, teachers may be caught up on creating the materials and learning as they make mistakes. Thus, true fidelity to the program may not be achieved until subsequent years of implementation. Future studies may find different results if Tools is examined in later years of implementation.

Unlike the previous two Tools studies that focus on child outcomes, a recent qualitative pilot study captured the experience of teachers learning and teaching the Tools of the Mind curriculum during the early phases of implementation (Imholz & Petrosino, 2012). The participants, all women from a school in New Jersey, included three preschool and two kindergarten teachers with varying years of experience teaching. The school community was predominantly White (83%), Hispanic (6.3%), African-American (3.1%), Asian (4.9%), and Native-American (0.16%), with a median household income of \$105,710.

Five teachers were selected to participate in this study, and were subsequently observed and interviewed over the course of one year following the Tools training. Data were collected during the 2009-2010 academic year, which was the second year of implementation of Tools. This year coincided with the district's transition from half-day to full-day kindergarten.

From the observations and interviews, this study uncovered several challenges and concerns experienced by teachers during training and implementation. These challenges were categorized into subcategories, including training and ethical issues, as well as classroom challenges. While high praise was offered for the actual training sessions, all five teachers struggled to incorporate the conceptual changes involved with the new curriculum. Teachers reflected that the training was more centered on the

experience of the child in the classroom and ignored the experiences of teachers learning the new curriculum. Imholz and Petrosino stated

“the greatest hurdle in the transition is conveying the curriculum to the teachers who are, in many instances, taking on the tasks of becoming a Tools teacher by decree from above and who bring with them a host of their own ideas about how to teach early childhood education” (p. 187).

Balancing learning the curriculum and reflecting on student progress was also described as a challenge by teachers. Teachers also articulated a desire to have more time to meet with other teachers to discuss strategies and concerns. One teacher questioned whether this program prepared her students for the next grade. Other issues uncovered in this study included teachers find keeping data of student progress challenging, as well as scheduling all of the prescribed Tools activities. Omitting some activities, such as dramatization, was one practice to deal with the challenge of time. This omission, however, demonstrated a lack of understanding of the important role dramatic play has on the development of self-regulation.

While challenges were uncovered, teachers were also able to recognize important contributions made by the program. The program had enhanced the overall classroom management skills and repertoire of teaching techniques in most of the teachers. Two teachers articulated interest in collaborating with first grade teachers to work toward smooth transitions for students. Participants also commented on observing an overall increase in the ability of their students to work cooperatively and to engage in verbal expression. Despite these contributions, teachers were not able to articulate gains in academic achievement.

This study contributes to the growing interest and knowledge of Tools. Unlike other studies that focus on the experience of the child or one particular aspect of the program, this study recognized the need to include the experiences of teachers in the training and implementation of the program. Understanding the experiences of the teachers may need to be given greater attention when studying the efficacy of this program, as evidenced by the challenges these teachers articulated. Conversely, these challenges may have also been associated with or confounded by the change to a full-day from half-day program.

Pilot Study

The scarcity of scholarly research about Tools ignited my interest in conducting a pilot study prior to this current study to examine preliminary statistical analysis on the relationship between Tools of the Mind and reading and math achievement in kindergarten. While a considerable body of research supports the positive correlation between self-regulation and achievement, and because the smaller body of research supports the relationship between Tools and the development of self-regulation, the question remained as to whether Tools would have a positive effect on achievement.

A total of 60 Kindergarten students from an elementary school in the Northeast participated in this study. The average age of students at the beginning of Kindergarten was 5.4 years and 6.3 years at the end of Kindergarten. Students in this sample were predominantly Caucasian (77.7), and some were Hispanic (13.5), African American (4.5%), and other (4.3%, from Africa, Asia). Approximately 40% of the students came from households that spoke a language other than English. Approximately 57.2% of the students were from low-income families, as defined by free and reduced lunch status.

This pilot study used available Measures of Academic Progress (MAP) testing data from spring testing during the years 2009 and 2011. Data from 2010 were not used, as this year was the first year of implementation of the Tools program. Thirty students were randomly selected from each year using a computer-generated list of random numbers, resulting in 60 participants. The full eligible population was not used, in an attempt to a random selection element in the study. Institutional consent to conduct research using available data was obtained from the superintendent.

The MAP is a diagnostic and computerized adaptive assessment. MAP scores were used to measure differences in student achievement in math and reading across two types of curriculum, traditional (2009) and Tools of the Mind (2011). In the primary analysis of data, means of math and reading scores were compared to examine whether a significant difference in achievement is observed using Analysis of Variance (ANOVA). Because Tools of the Mind is the independent variable, the design of this pilot study was a single-factor between group design using two One-Way ANOVAs. One ANOVA was completed for reading and one for math. Results of the ANOVA revealed that in reading, achievement differed significantly across both instructional methods, $F(1,58) = 6.145, p = .016$. The difference between both means is significant since the p -value obtained is less than the significance level of .05. In math, achievement did not differ significantly across both instructional methods, $F(1,58) = 2.580, p = .114$.

The results were quite perplexing, as gains in achievement were expected. Aside from the small sample size, other factors may have been present during the implementation of Tools that could not be explained by quantitative measures. These results indicated that additional research was necessary in order to assess the potential

effects of Tools on achievement.

Summary

This literature review has highlighted the positive impact that self-regulation can have on early achievement in reading and math. Distinctions have not been noted, however, regarding the different types of regulation (emotional, behavioral, and cognitive). Specific learning tasks, based on Vygotskian theory of development, have been related to the development of self-regulation, such as private speech and dramatic play. Despite this relationship, the scarcity of peer-reviewed studies that have tested the components of Vygotsky's theories has resulted in limited documentation of their specific effectiveness in achievement. While two studies (Barnett, et al., 2008, and Diamond, et al., 2007) have contributed to the existing literature about the efficacy of Tools of the Mind with respect to the development of self-regulation, neither study was able to establish a relationship between Tools and academic achievement. The goal of the present study is to extend the scholarship on Tools and attempt to make conclusions about the effectiveness of Tools on reading and math achievement.

CHAPTER 3

METHOD

The purpose of this study is to assess the efficacy of the early childhood curriculum, Tools of the Mind (Tools). Of specific interest is the effect of the Tools curriculum on the reading and math achievement of kindergarten students. This study continues and extends the analysis conducted during the pilot study I conducted in 2012. In the current study, not only will main effects of Tools be analyzed, but also the effects of additional variables will be analyzed. The design of the current study addresses the potential effects of covariates, including fidelity and socioeconomic levels. Using both quantitative and qualitative methods, this study will provide a more comprehensive picture of the effectiveness of Tools implementation.

Research Questions

The following research questions structure the basis for this study:

Reading

1. What is the relationship between the Tools of the Mind curriculum and achievement on measures of reading?
2. Is there a significant difference in reading achievement across socioeconomic levels?
3. What is the relationship between fidelity to Tools curriculum and achievement on measures of reading?

Math

4. What is the relationship between the Tools of the Mind curriculum and achievement on measures of math?

5. Is there a significant difference in math achievement across socioeconomic levels?
6. What is the relationship between fidelity to Tools curriculum and achievement on measures of math?

Qualitative

7. What are challenges that teachers experience with implementation of Tools?
8. What contributions do teachers perceive implementing Tools has made?

Participants and sampling procedures

Participants were selected from a public elementary school located in a small and urban town in New England. This school was selected because it was one out of three schools in the area using the Tools of the Mind curriculum. This school educates students from preschool through second grade. The district cutoff for enrollment is August 31, meaning that students have to turn five on or before that date in order to be eligible to register for kindergarten. Students in both groups were similar demographically (doe.mass.edu). In 2008-2009, students were predominantly Caucasian (78.6%), and some were Hispanic (12.7%), African-American (4.7%), and other (4%). A total of 7% of students were identified as not speaking English as their first language. During this year, a total of 427 students were enrolled of which 152 were in kindergarten. Approximately 44% of the students were considered low-income based on their free and reduced lunch qualification. The churn rate, indicating the number of students who newly enter or exit during the school year, during this year was 17.3%. During this year, the school participated in a large, national reading grant initiative (Armbruster, Lehr, Osborn, Adler, & National Institute for Literacy, 2001). This initiative sought to have every child

reading fluently by the end of third grade. Funding from this program provided several interventionists to provide intensive reading supports to students.

In 2010-2011, students were predominantly Caucasian (77.5%), and some were Hispanic (11.7%), African-American (5%), and other (10%). More families indicated “mixed race” in 10-11, accounting for the increase in “other.” A total of 7 % of students were identified as not speaking English as their first language. During this year, a total of 426 students were enrolled of which 137 were in kindergarten. Approximately 52% of the students were considered low-income based on their free and reduced lunch qualification. The churn rate during this year was 17.6%. Although the school contains a preschool program, these numbers do not reflect preschool enrollment. During this academic year, the position of a specific behavior interventionist was added. Additionally, the previous national grant had terminated and specific reading interventionists were no longer available.

Permission was obtained to gather and analyze data from both years from the district as well as from the building principal. The selection of five teachers who taught in kindergarten under both conditions (Tools & non-Tools) was carefully planned to enable comparison between both groups. Including other teachers would not ensure controls for teacher effects. In addition to gathering achievement data pertaining to the students who had been in kindergarten during both years with these teachers, the need for qualitative interviews emerged. These five kindergarten teachers were asked to participate. While gathering secondary data did not require Institutional Review Board permission regarding human subjects, conducting qualitative interviews did. Following institutional permission, individual meetings were held with prospective participants to

solicit participation. University Institutional Review Board protocol for informed consent was followed. *Appendix A* includes a sample participant consent form.

The majority of teachers in this district were between the ages of 49-56 earning an average salary of \$60,260 during 08-09 and \$62,586 in 10-11. Of the five teachers who participated in this study, one had additional licensure in special education. The number of years of total teaching experience varied from 2 to 33 and of teaching at this school from 2 to 13 years. All of the teachers were Caucasian-American.

Sample size, power, and precision

An a priori statistical power analysis was conducted using XLSTAT (XLSTAT version 2013.2, a Microsoft Excel[®] add-in) software program in order to verify the needed sample size. Effect sizes in this statistic reflect Cohen's *d* (Cohen, 1988). Given an alpha level of .05, a power level of .80, a numerator degree of freedom of 1 (the number of groups minus one), one covariate, and an expected moderate effect size of .40, the analysis calculated and recommended that a sample size of 51 was needed for detection of a statistically significant result. A second, more conservative power analysis was run in the event that an expected effect size of .40 was too optimistic. Given an alpha level of .05, a power level of .80, a numerator degree of freedom of 1, one covariate, and an expected effect size of .25, the analysis calculated and recommended that a sample size of 128 was needed in order for detection of a statistically significant result. Figures B.1 and B.2 in *Appendix B* contain the power analysis outputs.

Procedure

This study included data from ten kindergarten classrooms. The first sample consisted of 94 students drawn from five kindergarten classrooms from 2008-2009,

forming the control group as Tools had not yet been implemented. This sample consisted of 56% boys and 57% free and reduced lunch status. The second sample consisted of 97 students from five kindergarten classrooms from 2010-2011, and formed the experimental group as the intervention, Tools, was in its second year of implementation. This second sample consisted of 51% boys and 63% free and reduced lunch status. The final sample contained a total of 191 kindergarten students. Table 3.1 depicts the breakdown of gender and free and reduced lunch status.

Table 3.1 Sample Demographics

	2008-2009	2010-2011
Males	56%	51%
Females	44%	49%
Free/Reduced Lunch Status	44%	52%

Previous Tools research has examined the efficacy of Tools in the first year of implementation (Barnett et al., 2008). One of the limitations identified with this study was that teachers may have been acclimating to a new curriculum and philosophy and results may have reflected this shift rather than the program's efficacy. In the present study, data were collected in the second year to control for first year adjustment to the new curriculum by teachers. In the next section, the research design is explained. Both groups include the same five teachers. The same teachers were present in group 1 and in group 2, but since each group is from a different academic year, the participants (scores) are from two independent samples.

Research Design (Quantitative)

This study used a mixed design (between group comparison), with two groups (teacher-directed & Tools), and potential interactions of the other variables including letter naming score, teacher fidelity and socioeconomic level (SES).

Reading achievement

Analysis of covariance (ANCOVA) was the statistical technique chosen as it controls for differences that exist between groups before a comparison of the within- and between-groups variance can be made (Gall, Gall, & Borg, 2003). Specifically, a Two-Way ANCOVA was conducted with measures of reading, as the letter naming score was the only beginning of the year measure obtained. No early math measures were obtained. Controlling additional variables, ANCOVA attempts to make the two groups equal. Because two groups are being compared, ANCOVA is an appropriate procedure because it controls for initial differences between the groups on one or more extraneous variables. ANCOVA allows for the determination of differences between the mean scores of both groups on the dependent variable (Gall, Gall, & Borg, 2003). Letter naming was the covariate, the reading score was the dependent variable, and the independent variables was SES, and YEAR (Tools or Not Tools). Before analyzing the data, the ANCOVA assumptions were considered. These assumptions included Independence, Homogeneity of Variance, Normality, Linearity, Fixed Independent Variable, Independence of the Covariate and the Independent Variable, and Homogeneity of the Regression Slopes (Rutherford, 2011). The data were examined to insure that the assumptions of ANCOVA were met.

Effect sizes were calculated using partial *eta squared* generated from SPSS. Effect size refers to the magnitude of the estimated difference or relationship (Gall, Gall, & Borg, 2003). In SPSS, only partial *eta squared* are generated. This statistic, using the formula $\eta_p^2 = SS_{\text{between}} / SS_{\text{total}} + SS_{\text{error}}$, represents a ratio of the variance that is accounted for by an effect in addition to that effect and in addition to the associated error variance within the ANCOVA study (Richardson, 2011).

Math achievement

To measure the effect of Tools implementation on math achievement, a Two-Way ANOVA was conducted with math score as the dependent variable, and SES and YEAR (Tools or Not-Tools) as the independent variables. ANCOVA was not appropriate for measures of math achievement as I did not have a covariate (beginning of the year measure of math). Before the ANOVA results were analyzed, the assumptions of the ANOVA were considered, including Independence, Normality, and Homogeneity of Variance (Rutherford, 2011).

Because five separate teachers are included in this study, the potential for variance based on teacher needs was also addressed. Each teacher may or may not implement the program with fidelity. To measure the influence of teacher fidelity on student scores, correlations were conducted using scores from the Tools year only. In the next section, a description of each measure is provided.

Measures

This study relied on previously collected reading and math data from both comparison years. Several measures were used in data collection.

Letter Naming Fluency

To assess how similar the two groups were prior to comparison of treatment effects, the results of *Letter Naming Fluency* tests were gathered as pretest data. In 2008-2009, the district used the Test of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002). This test enables a teacher to record how many letters named correctly in one minute with accuracy and speed. The validity of using this DIBELS *Letter Naming Fluency* has been widely tested (Good, Kaminski, Simmons, & Kame'enui, 2001). According to validity testing, Letter Naming Fluency at the kindergarten level has been found to predict later reading achievement. In 2010-2011, the district switched to AIMSweb (Edformation, 2005). AIMSweb also contains a test of *Letter Naming Fluency* and is measured in the same fashion as in the DIBELS assessment. Both DIBELS and AIMSweb are considered curriculum-based measures (CBM; Deno, 2003; Deno, Espin, & Fuchs, 2002). Curriculum-based measures provide benchmarks of expected progress and allow teachers to frequently monitor the progress of their students to inform their instruction. The choice to use the measure of Letter Naming Fluency in kindergarten was made due to its strong empirical support for its predictive value of future reading achievement (Clarke, Hulme, & Snowling, 2005; Duncan & Seymour, 2000; Foulon, 2005; Pennington & Lefty, 2001).

Measures of Academic Progress

To provide data regarding academic achievement, the Measures of Academic Progress (MAP; www.nwea.org) testing scores in reading and math were used. At the beginning, middle and end of the academic year, all students at this school were administered MAP testing. MAP scores evaluate knowledge of core math and reading

skills in Grades K-2. Within math, math reasoning and calculation are averaged. In reading, reading and comprehension scores are averaged. Each average is given a score, called a Rasch unit (RIT; nwea.org). A student's current achievement level and growth in reading or math is shown in the RIT score. MAP is a diagnostic and computerized adaptive assessment. The Northwest Evaluation Association (NWEA) has conducted several studies on the appropriateness of using MAP testing, including scale alignment studies which demonstrate the alignment of skills tested with states the test is given in (Cronin, Kingsbury, & Bowe, 2007). This alignment increases the validity of the use of MAP to measure achievement. Additionally, the NWEA uses rigorous standards of reliability. Most coefficients of test and re-test studies have fallen at the mid .80's to low .90's, consistently yielding correlations that are statistically significant between multiple tests by the same students (nwea.org).

Fidelity of Implementation Scale

Teacher fidelity to the implementation of Tools was measured using the Fidelity of Implementation Scale developed by Tools of the Mind (toolsofthemind.org). This observation tool was created specifically for observing fidelity in Tools kindergarten classrooms. The scale was a 5-point scale with "0" indicating absence of implementation of Tools activities and "5" indicating that fidelity is excellent. While technically ordinal, fidelity scores from this scale are treated as interval and this procedure has been supported in previous research (Lord, 1953). *Appendix C* includes a sample of the observation form used to obtain the fidelity scores. Three raters, including myself, were selected from within the school, based on their knowledge of the program. Because the data periods are in the past, the rating took place retrospectively. To measure inter-rater

reliability of fidelity ratings, an intraclass correlation coefficient was computed using SPSS. The intraclass correlation coefficient (Kappa) measures the reliability of ratings of two or more raters on ordinal or nominal scales (Cohen, 1960; Fleiss, Levin, & Paik, 2003). In this calculation, each subject (teacher) was rated by the same raters. The intraclass correlation coefficient was .96 for estimates of single ratings, .99 for estimates of averages of k ratings, indicating adequate agreement in ratings between raters. This level of agreement is considered in the “almost perfect range” (Landis & Koch, 1977). Figure D.1 in *Appendix D* contains the output of the intraclass correlation coefficient.

Socioeconomic Status

A measure of socioeconomic status was obtained from district data bases (doe.mass.edu). Free and reduced lunch status was used as a measure of SES. Free and reduced lunch status has been used in past research as a measure of SES (FRL; Stein, Berends, Fuchs, McMaster, Saenz, et al., 2008; van Ewijk & Slegers, 2010) and both types have been considered as a single unit. In this study, free and reduced lunch status were combined and coded as “free and reduced lunch” in comparison to “not free and reduced lunch.” Because this SES measure is only binary (free/reduced or not), it was not used as a covariate, but rather as an independent variable.

Collectively, letter naming scores obtained at the beginning of each year with end of year MAP reading and math scores should provide sufficient data for analysis. Letter naming scores could not be retrieved for end of year measures during the 2008-2009 year. MAP reading and math assessments were not conducted during the 2010-2011 year at the mid-year point. The availability of these scores would have enabled additional comparisons. Despite this lack of additional evidence, I was able to collect qualitative

data. In the next section, the qualitative research design is summarized.

Research Design (Qualitative)

Researcher profile

My interest in learning about the effectiveness of the Tools of the Mind curriculum began in the training phases. I was allowed to attend the three-day training during the summer prior to the first year of implementation. I realized during this training that a major pedagogical and philosophical shift would need to occur in the teachers and administrators given the sudden shift from highly directed curricula to Tools. In addition, I was ambivalent about letting go of the responses to problem behaviors that I had learned at so many trainings. With Tools came the information about self-regulation and the caution against using external mediators, such as behavior charts, which many educators had come to depend on. As I became more of a researcher through my doctoral program, I also questioned the lack of empirical evidence in support of Tools and how districts made decisions to implement new programs. These questions ultimately led me to pursue this inquiry as my doctoral research topic.

Data Collection

Data collection consisted of interviews of the five teachers that had been present in kindergarten before and after Tools implementation. Each teacher participated in a semi-structured, audio-taped interview, each lasting approximately 25 minutes. Interview questions focused on gaining an understanding of each of the teacher's experiences with training, implementation, reflection on state standards, and concerns. Questions and follow-up prompts were prepared ahead of time. Interview questions can be found in *Appendix E*. During the interview, questions were read to guide open-ended responses.

Notes were taken to document expressions or gestures that could not be captured in the audio-recording. Researcher notes were added immediately after each interview. Data were collected over a two week period.

Data Analysis

Data were analyzed using open and axial coding (Merriam, 2009). Construction of categories began immediately during reading the first observation notes and reading the first interview transcription. In the margins created for comments, notations were made, such as “enough training or not enough training.” During this process, coding was open, as any bit of data might be useful (Merriam, 2009, p. 178). These notations later transformed into categories and subcategories, based on how the data were being interpreted and what meaning was being attributed to each bit of data, resulting in axial coding (Merriam, 2009, p. 180). Color codes were assigned to the different categories of responses. Categories were then evaluated to ensure that they answered one of the research questions.

Trustworthiness

The goal of the trustworthiness of this study is to enable those who read it to use the findings to improve the experiences of teachers and of students who are participating in the Tools program. This study was conducted using sound methodological practices under the guidance of a trustworthy professor and committee members. Ethical guidelines were followed under the guidance of the University’s IRB. Special care was taken to use pseudonyms to protect the confidentiality not only of the individual participants, but also of the school. A final contributor to the trustworthiness of this study is the respect for sensitive position of teachers in this era of accountability. This study

respects the perspectives of teachers, who also need a voice to their experiences.

Summary

Together, quantitative and qualitative data enabled a more comprehensive assessment of the efficacy of Tools. Data captured through the interviews with teachers may further explain some of the obtained quantitative results. In the next section, the quantitative and qualitative results are discussed.

CHAPTER 4

RESULTS

In this section, an overview of the results from the study of the efficacy of the implementation of Tools on math and reading achievement in kindergarten is presented. Descriptive and inferential statistics allowed the researcher to describe and draw conclusions about the two groups (one that experienced Tools and one that did not). This section consists of two parts. Quantitative results are reported in the first section, while qualitative results are reported in the second section.

Quantitative Results

Summarized results were calculated using the Statistical Package for the Social Sciences Version 17 (SPSS, 2009) software program. Values including means, variance, standard deviations, and demographic information are reported in this section.

This quantitative portion of the research examined six questions. The final analysis for each research question is listed below:

Results for Research Question #1

What is the relationship between the Tools of the Mind curriculum and achievement on measures of reading?

H₀: There is no statistically significant difference in the reading achievement of kindergarteners in the Tools group as compared to kindergarteners in the non-Tools group.

H₁: There is a statistically significant difference in the reading achievement of kindergarteners in the Tools group as compared to kindergarteners in the non-Tools group.

To measure potential differences between Tools and non-Tools students' achievement scores in reading based on MAP testing, an ANCOVA model was conducted. In this model, the dependent variable was the reading score achieved on the MAP testing, and the year (Tools and non-Tools) and SES were the fixed factors/independent variables. The covariate in this study was the letter naming score that was obtained at the beginning of each year. In ANCOVA, *F* ratios are computed by dividing the explained variance between groups (e.g. year, as in Tools vs. non-Tools) by the unexplained variance within the groups. The ANCOVA enabled me to assess whether the covariate was linearly related to the MAP reading score. In addition, as a covariate, the letter naming score used in this model enabled me to assess how different the groups were at the onset of each year. This comparison was critical for the later analysis of treatment effect.

Before analyzing potential differences, the assumptions of ANCOVA were checked. The first three assumptions were met based on the design of the study. The dependent variable (MAP reading score) was interval. Both independent variables within this model were categorical and independent groups. Additionally, different participants were in each group, assuring independence of observations. The levels of the independent variables were set by the researcher. There was homoscedasticity and homogeneity of variances, as assessed by visual inspection of a scatterplot and Levene's Test of Homogeneity of Variance ($p = .365$), respectively. Satisfying the normality assumption, unstandardized residuals for the interventions and for the overall model were normally distributed, as assessed by Shapiro-Wilks test ($p > .05$), with the exception of the combination of Free/Reduced Lunch and 2009. In addition, skewness and kurtosis

fell within the required |2|. The ANCOVA revealed a linear relationship between letter name and the intervention types (Tools & non-Tools), as assessed by visual inspection of a scatterplot. To determine if the covariate and the independent variable were independent, independent samples *t*-test were run. There were no significant differences in the scores for free and reduced lunch ($M = 15.81, SD = 15.73$) and not free and reduced lunch ($M = 19.68, SD = 16.18$) conditions; $t_{(173)} = .725, p = .118$. Similarly, there were no significant differences in the scores for non-Tools ($M = 16.93, SD = 15.58$) and Tools ($M = 17.71, SD = 16.41$) conditions; $t_{(173)} = -.321, p = .749$. There was homogeneity of regression slopes as the interaction term was not statistically significant for YEAR *LetterName, $F(1,169) = .483, p = .488$ or for SES*LetterName, $F(1,169) = .640, p = .425$, and because the interaction plots look roughly parallel. Below, Table 4.1 contains the output of the ANCOVA.

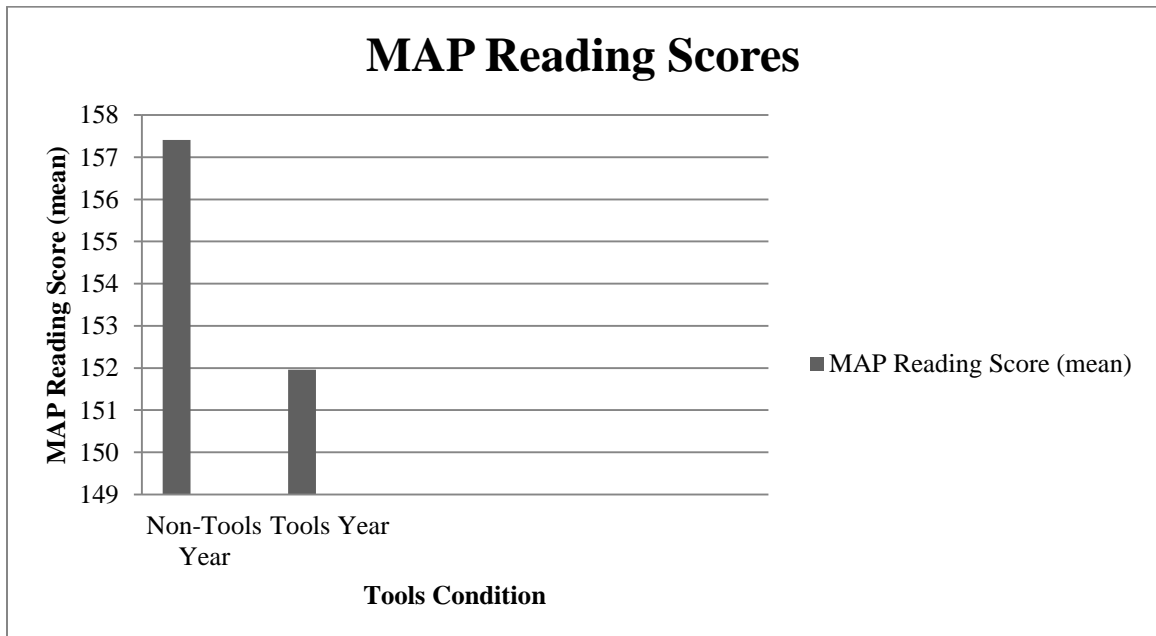
Table 4.1 Reading Achievement Scores by YEAR and SES

Source	SS	df	MS	F	p
Corrected Model	6816.917	5	1363.383	14.271	.000
Intercept	1605848.925	1	1605848.925	16808.905	.000
Year	498.654	1	498.654	5.220	.024
SES	42.363	1	42.363	.443	.506
LetterName	4393.25	1	4393.25	45.99	.000
Year*LetterName	46.158	1	46.158	.483	.488
SES*LetterName	61.02	1	61.102	.640	.425
Error	16145.518	169	95.536		

Appendix F contains output graphs and tables that illustrate satisfaction of these assumptions (see Figures F.1-F.6 and Tables F.1-F.6).

A statistically significant interaction between YEAR and SES was not observed, $F(1,170) = .025, p = .876, \eta_p^2 = .000$. This result is illustrated in Figure 4.1. After adjustment for letter name, there was a statistically significant difference in MAP reading scores between Tools and non-Tools years, $F(1, 170) = 17.217, p = .000$, partial $\eta^2 = .092$, therefore rejecting the null hypothesis. Students in the Tools year obtained lower scores on MAP reading than students in the non-Tools year. The partial η^2 indicated that 9.2% of the variability of the dependent variable, reading achievement, could be explained by the independent variable, Tools year. The partial η^2 for SES was also included in this model and will be discussed in the next section.

Figure 4.1 MAP Reading Scores (means) Across Tools Conditions



Results for Research Question #2

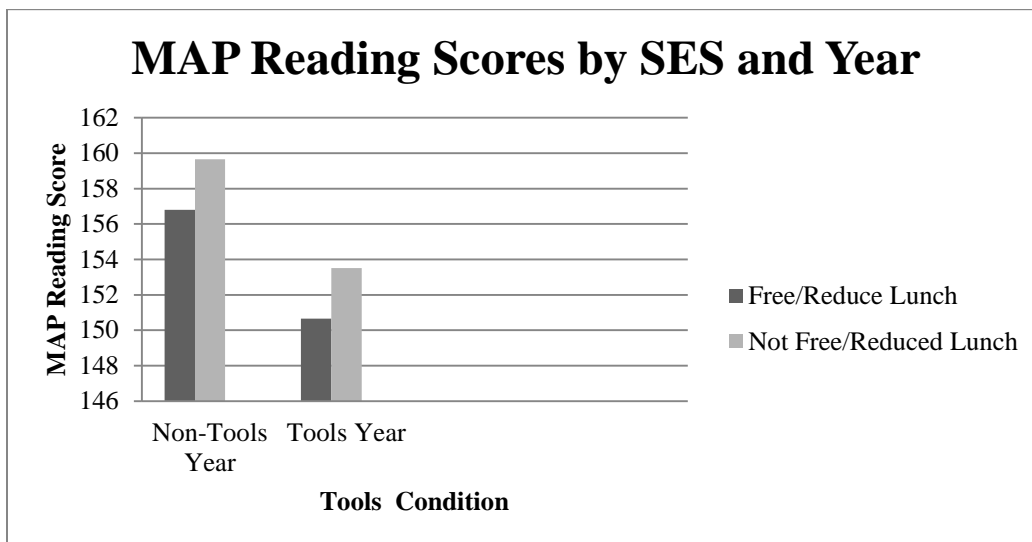
Is there a significant difference in MAP reading achievement across socioeconomic levels?

H₀: There is no significant difference in MAP reading scores across socioeconomic levels.

H₁: This is a significant difference in MAP reading scores across socioeconomic levels.

Analysis for this question was conducted using the same ANCOVA used in Question 1, where SES was the second independent variable. After adjusting for beginning of the year letter name scores, there was a statistically significant difference in MAP reading scores between SES levels as measured by Free/Reduced lunch status, $F(1,170) = 3.925, p = .049, \text{partial } \eta^2 = .023$. Students who were not Free/Reduced lunch status obtained higher scores on the MAP reading test than students who were not identified as Free/Reduced lunch status. The partial η^2 indicated that 2.3% of the variability of the dependent variable, reading achievement, could be explained by the independent variable, SES, in addition to the 9.2% explained by YEAR. Figure 4.2 illustrates the constant effect of SES on MAP reading scores.

Figure 4.2 MAP Reading Scores by SES and Year



Results for Research Question #3

What is the relationship between fidelity to Tools curriculum and achievement on measures of reading?

H₀: There is no relationship between achievement and teacher fidelity.

H₁: There is a relationship between achievement and teacher fidelity.

In order to examine the relationship between fidelity scores and MAP reading scores, a Pearson correlation was run, $r = .085$, $p = .246$. This result indicates no or negligible relationship. As fidelity increased, this result did not indicate that reading scores increase or decrease.

Results for Research Question #4

What is the relationship between the Tools of the Mind curriculum and achievement on measures of math?

H₀: There is no statistically significant difference in math achievement of kindergarteners in the Tools group as compared to kindergarteners in the non-Tools group.

H₁: There is a statistically significant difference in math achievement of kindergarteners in the Tools group as compared to kindergarteners in the non-Tools group.

To compare the mean differences in MAP math scores between the two group split on two variables YEAR and SES, a two-way ANOVA was conducted. The ANOVA was chosen instead of the ANCOVA due to the absence of a covariate. Unlike the reading model, there was no beginning of year math measure to provide the covariate. The two-way ANOVA allowed for observation of an interaction between the independent variables on the dependent variable. Before the ANOVA was analyzed for main effects and interactions, the assumptions of the ANOVA were checked. The first three assumptions were met based on the design of the study. The dependent variable (MAP math score) was interval. Both independent variables within this model were categorical and formed independent groups. Additionally, different participants were in each group, assuring independence of observations. Remaining assumptions were checked by running statistical tests using SPSS. Regarding the normality assumption, MAP math scores were normally distributed for all group combinations of SES and year (Tools and non-Tools), as assessed by the Shapiro-Wilks test ($p > .05$). There was homogeneity of variances, as assessed by Levene's Test of Homogeneity of Variance ($p = .273$).

Analysis of the ANOVA did not reveal a statistically significant interaction between SES and YEAR on MAP math scores, $F(1,186) = .049$, $p = .825$, partial $\eta^2 = .000$. The next step in the analysis involved testing the main effect of YEAR (Tool and non-Tools). In this analysis, I was testing for differences in MAP math scores between YEAR of students collapsed across SES, meaning that SES was in all intents and purposes, ignored and the model only considered YEAR. The results of the ANOVA

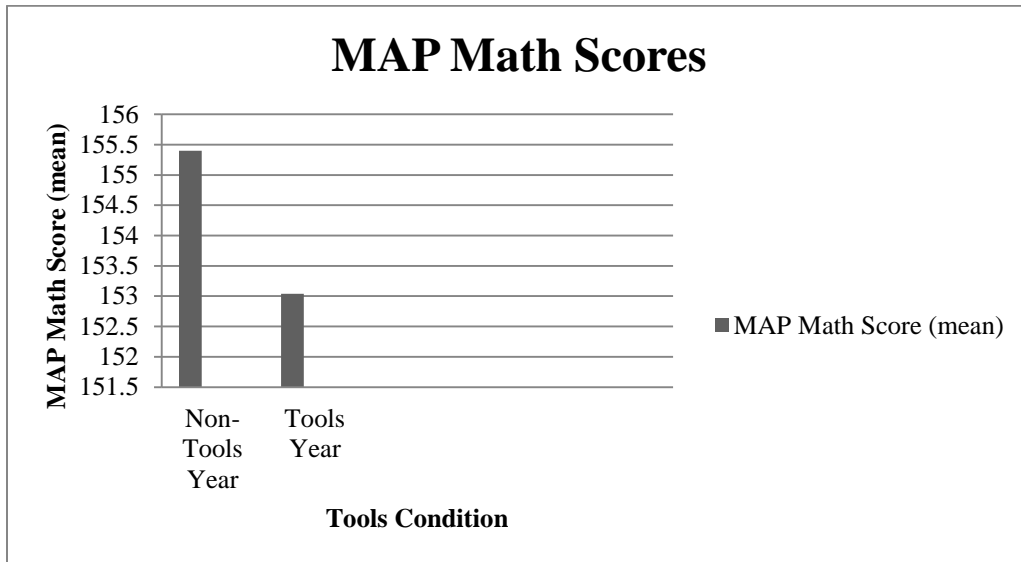
indicated that there was not a statistically significant difference in the math achievement of kindergarteners who had experienced Tools as compared with kindergarteners who had not experienced Tools. Therefore, the null hypothesis could not be rejected, $F(1, 186) = 1.256, p = .264$. Table 4.2 lists the ANOVA data for Questions 3 and 4. *Appendix G* contains tables and figures that illustrate the satisfaction of these assumptions (see Tables G.1-G.4 and Figure G.1).

Table 4.2 Math Achievement Scores by YEAR and SES

Source	SS	df	MS	F	Sig.
Corrected Model	2312.377	3	770.792	4.756	.003
Intercept	4346923.744	1	4346923.744	26820.919	.000
Year	203.617	1	203.617	1.256	.264
SES	2040.751	1	2040.751	12.592	.063
Year*SES	7.912	1	7.912	.049	.825
Error	30145.418	186	162.072		

Comparison of the means of MAP math scores indicates that math achievement scores were higher in the non-Tools condition (155.92) than in the Tools condition (152.99). While not statistically significant, results nonetheless indicate a drop in achievement scores in math. Figure 4.3 illustrates the difference in MAP math scores across Tools conditions.

Figure 4.3 MAP Math Scores by Tools Condition



Results for Research Question #5

Is there a significant difference in math achievement across socioeconomic levels?

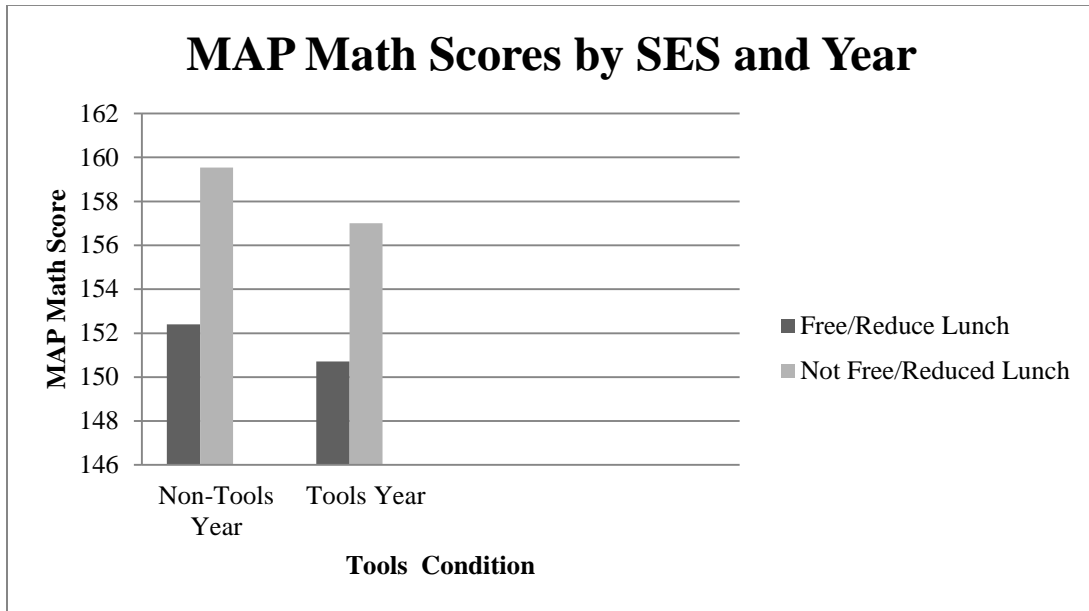
H₀: There is no significant difference in MAP math scores across socioeconomic levels.

H₁: This is a significant difference in MAP math scores across socioeconomic levels.

Testing for the main effect of SES involved testing for differences in MAP math scores between Free/Reduced lunch status and non-Free/Reduced lunch status students collapsed across YEAR. This collapse meant that YEAR was in essence ignored in order to simply observe SES differences. To answer this question, results of the ANOVA described above were analyzed. There was a statistically significant difference between SES levels for MAP math scores, $F(1, 186) = 12.592, p = .000, \text{partial } \eta^2 = .063$. As assessed by partial Eta, 6.3% of the variability of the dependent variable, math

achievement, can be explained by the independent variable, SES. Students not identified as Free/Reduced lunch status scored higher on MAP measures of math as compared to students identified as Free/Reduced lunch status. Figure 4.4 illustrates the MAP Math scores by Tools condition and SES.

Figure 4.4 MAP Math by Tools Condition and SES



Results for Research Question #6

What is the relationship between fidelity to Tools curriculum and achievement on measures of math?

H₀: There is no relationship between achievement and teacher fidelity.

H₁: There is a relationship between achievement and teacher fidelity.

In order to examine the relationship between fidelity scores and MAP math scores, A Pearson correlation was conducted, $r = .109$, $p = .135$. This result indicates no or negligible relationship. As fidelity increased, this result does not indicate that reading scores increased or decreased. See Tables H.1 and H.2 in *Appendix H* for Pearson

Correlation Results.

Qualitative Results

Data from this study were analyzed from a contextual ecological perspective. To describe the phenomenon of teaching Tools of the Mind in kindergarten, excerpts from interviews are provided to help describe the ecologies of each teacher. These excerpts will further delineate how teachers' perceptions of their own preparedness and factors related to the transition to Tools from a more teacher-directed curriculum and efficacy impact their instructional practices and student outcomes. Teachers are identified as Teacher 1, 2...(T1, T2...). Results are presented in sequence to answer two research questions, corresponding to Research Questions 7 and 8:

7. What are challenges that teachers experience with implementation of Tools?
8. What contributions do teachers perceive implementing Tools has made?

Results for Research Question #7

The first qualitative research question sought to discover teachers' perceptions of the challenges involved with implementation of Tools. Coding of the transcriptions resulted in the discovery of four main challenges related to training, personal cost, fidelity, and accountability.

Challenges Associated with Tools Implementation

“Challenges” is a category that identifies some of the difficulties articulated by teachers during their implementation of Tools, specifically in year two of implementation. The four main challenges included issues with professional development in preparation of and in support of Tools implementation, issues with fidelity to the program, and issues

related to accountability for student outcomes. Professional development challenges included trainings sponsored by Tools where teachers learned about the history, theory, and development of the curriculum. Implementation support and coaching included district and Tools of the Mind support and coaching regarding development of activities, creation of materials, and ongoing coaching related to implementation. This category also included personal costs of money and time expended by teachers. Additionally, this category includes evidence of the reflection, or lack of reflection, by teachers related to alignment of Tools to the state standards. Because the measure of reading and math progress was an assessment aligned with state standards, investigating the presence of such reflection may add to the discussion of the results obtained in quantitative analysis.

T1's Perceptions of Challenges. T1 characterized the initial three-day training as “very good” but that because the training took place in June, and implementation did not begin until August, some information had been forgotten. Subsequent trainings that took place during the year, and during the following two years, were characterized as helpful, and that Tools staff and district the district literacy coach were very helpful. This teacher expressed gratitude that the district had included the literacy coach in the training, which provided an in-house support throughout the year. The focus in teaching was placed on the fostering of self-regulation and teaching expectations, according to T1, although a lot of time was spent on making materials for the activities, which were neither provided by Tools or the district. While T1 believed that the focus on self-regulation activities was useful, she did not feel prepared to follow the expected guidelines of the program.

Additional issues identified by T1 were included in the challenges category. First, because Tools was only implemented in kindergarten, children had a hard time

generalizing those skills to first grade, which followed the more traditional, teacher-directed approach to teaching. Some students struggled to adjust to the new expectations when they made the transition to the next grade. Even within kindergarten, T1 recognized that some students struggled with the development of self-regulation and academics, and while most of these children were identified as needing special education services, not much was offered within the Tools framework. In addition, T1 shared that communicating with staff outside of those familiar with Tools was “a huge stumbling block.” A lack of consistency was observed in the programs as well. T1 stated, “Scott Foresman (used in grade 1) assumed some skills would be covered in kindergarten, and we found it wasn’t, so we found a gap.” Finally, T1 identified the expectation that teachers make an inordinate amount of materials challenging and personally costly.

With respect to alignment, T1 stated that because the manual said that the program was aligned with standards, the teachers did not pay attention to alignment, until they became aware of the gaps in the program. She expressed concerns that the district then assessed the students using a computerized assessment based on standards. T1 stated that her students may not have been sufficiently proficient with the computer to be assessed using a computerized test.

T2’s Perceptions of Challenges. T2 described the trainings as helpful, yet overwhelming, unlike T1. Trainings were described as helpful in that the program was broken down, they were shown how to implement the activities, and learned Vygotskian theory. The overwhelming components were depicted in this statement: “It was so different from anything I had ever done before, and I was blown away by the

expectations, the amount of materials that had to be made were overwhelming, the trainings did a nice job of overview, but did not dive into how activities were supposed to go.” T2 also described feeling unprepared. Support throughout the year was mainly obtained from peers, although each was only guessing, according to T2. Although questions could be emailed or asked at training sessions, so much information was loaded at trainings, which made it difficult to ask questions.

Other challenges were identified by T2, including the acknowledgement that some students who had been severely lacking in self-regulation did not seem to gain self-regulation with Tools. However, T2 stated that she did not think they would have succeeded in other programs either. Academically, “those children that needed the discreet intervention, small group with a teacher (which really wasn’t what Tools did) and needed those skills early just seemed get further behind.” According to T2, Tools did not provide interventions to be able to support these children to make the gains teachers had seen children in the past make with interventions. Another challenge identified was that the paraprofessional in the room was often pulled for other coverage, leaving the responsibility of scaffolding to one adult. Finally, T2 expressed concerns with the lack of consistency between Tools and other programs offered at this school. She stated, “It was as if we were speaking a different language, but I also felt that when I came to SST (pre-referral process), that I did not have a whole lot of data to present. I had writing samples, my own personal feelings...but not a lot of concrete data.”

With respect to alignment, T2 reflected that as a team, the teachers did not reflect on alignment with standards, stating, “we trusted Tools that it was based on our state curriculum.” In year two, the team began to notice gaps, especially after mid-year

benchmark scores were obtained, and concerns arose related to unexpectedly lower scores.

T3's Perceptions of Challenges. T3 also recalled feeling overwhelmed after the initial three-day training. She stated, “after the trainings we were blown away, because our traditional programs provided us with everything we needed to teach the program, and they were quite scripted. With Tools, I spent the rest of that summer making materials.” T3 also described her lack of confidence in her understanding of the program in the statement, “I did not feel like I knew what I was doing.” In addition, T3 stated that she did not make the connection between what she was doing and how it related to Vygotsky, although she felt she could execute the games and activities. When on-site trainings took place, T3 stated that the constructive criticism wasn't always gentle, deepening her lack of confidence.

Additional challenges were identified by T3. Tools was characterized by T3 as effective for some but not for other students. With respect to self-regulation, “there were outliers that it just didn't help” behaviorally or academically. Some students were not even ready for teacher regulation, let alone other- or self-regulation. T3 added that there was a noticeable decrease in the number of students with out-of-control behavior, but that she could not say with certainty if this reduction was due to the program or to the increase in behavioral support staff and programming. These students struggled with some of the harder academic pieces, such as the learning plans. T3 also identified the challenge of having to scaffold alone at times, when the paraprofessional was pulled to cover other rooms or other duties. This support was difficult to offer, especially when severely dysregulated students required so much scaffolding. The making of the

materials was also identified as a challenge. T3 further explained that the communication between kindergarten and others in the school was limited due to their limited understanding of the program. This limited understanding of the Tools by those in other grades gave T3 the impression that “others did not think we were teaching letters, that all we did was play.”

With respect to alignment, T3 recalled not sufficiently reflecting on standards until late in year two, stating, “I can’t recall a time to reflect on standards. We were in survival mode.” She recalled that a weekly scope and sequence was introduced by Tools in the third year.

T4’s Perceptions of Challenges. T4 described the Tools trainings as very helpful, providing a lot of information. However, she recalled that the manual was open to interpretation and that she found that all the teachers were implementing the program basically as it had been presented, but with their own flair. When asked if personalizing the program was allowed, T4 stated that the manual was supposed to be followed precisely. When the trainers conducted classroom visits, they commented if the program was not being implemented with fidelity. While this aspect of implementation was confusing, T4 recalled that she appreciated the visits and feedback. Despite this feedback, T4 expressed frustration with inconsistencies and what she characterized as insufficient training.

Additional challenges were articulated by T4. She described students who were not self-regulated or were identified as limited in their academic skills as not reaching expected levels of self-regulation. T4 reflected that this lack of self-regulation was not representative of all students. Additionally, “opportunities have to be placed in the

classroom for those children to succeed.” With this statement, T4 identified a challenge of meeting the needs of children at-risk students without other interventions. In trainings, teachers learned that Tools did not support the use of behavior charts for students as these external regulators might interfere with the development of self-regulation. T4 found this aspect of Tools limiting when working with students with behavioral challenges. In addition, T4 expressed concern about how staff outside of Tools had little understanding of the program, making communication across grade levels and support services difficult. T4 also stated that new teachers had not been assigned challenging behaviors, loading seasoned teachers with potentially disruptive students.

With respect to alignment, T4 did not consider the program adequately aligned with state standards, stating, “I did not feel comfortable that I was meeting the standard with that activity.”

T5’s Perceptions of Challenges. T5 characterized the training as overwhelming, as other participants had. She stated, “it was just a whole different learning approach, a lot thrown at us at once, and also knowing that we had to create all of the materials.” T5 described the initial training as not sufficient, but more helpful in subsequent years due to the learning that took place while actually implementing the program.

Together these findings suggest that teachers involved in the implementation of Tools at this school were dissatisfied with the amount of training provided, especially at the initial level of training. Unanimously, teachers related the overwhelming amount of new information to absorb as well as the philosophical shift that was required with the program to their perceptions as not ready or with sufficient competence to implement the program.

While T5 identified scaffolding as a strength of the program, she also identified it as a challenge. She stated, “Tools is heavily based on scaffolding at their (child’s) level, I think just trying to figure out how to be able to do it in the time span that we were given was a challenge. T5 also expressed concern about having to buy and make so many materials. A final challenge identified by T5 involved communication with staff outside of Tools. She cited poor transition from kindergarten to grade one resulting from a lack of understanding of Tools by grade one teachers. In addition, T5 found difficulty in explaining student progress or concerns in Special Education or parent meetings, as the language of Tools was unfamiliar to those outside of kindergarten.

With respect to alignment, T5 was able to recall the lack of reflection by teachers on alignment with state standards. According to T5, it was not until around the third year that the team reflected more on alignment, and found ways that the program is not aligned with standards.

Together, these data identified several challenges experienced by teachers while implementing Tools. These challenges included poor communication with staff outside of Tools, inconsistency of the program with other curricula, especially with reference to the transition between kindergarten and grade one, lack of reflection about alignment with state standards, making materials, and reliability of using MAP to assess student progress. With respect to alignment, these data clearly suggest that sufficient reflection did not occur between what was being taught and what was outlined in the state standards. While reflection increased at the end of year two and during subsequent years, during the period of time involved (year two) in MAP testing, there is no evidence of reflection of standards. Table 4.3 includes a summary of challenges with and

contributions of Tools experienced by the participants.

Table 4.3 Summary of Qualitative Results

Challenges	Contributions
Initial and ongoing professional development	Writing
Personal costs	Language development
Time	Cooperation and collaboration with peers
Lack of alignment with standards	Decrease in visibility of highly disruptive students
Too much information	Professional growth and understanding of development and play
Theoretical shift	
Lack of support interventions for struggling students	

Results for Research Question #8

Perceived Contributions of Tools

“Contributions of Tools” is a category that captures the successes and strengths of Tools experienced by teachers. This category captures these experiences as they relate to implementation and student outcomes.

T1’s Perceptions of Contributions. T1 described Tools as a successful program for the development of self-regulation and academic skills for most students. As for the Magic Tree House books, the students “were really engaged.” The program also focused on collaboration, and the students had to learn from each other with “buddy checks.” T1 stated that she observed a positive difference in the quality of the writing, language development, and play with Tools.

T2's Perceptions of Contributions. T2 identified a different perspective that Tools did focus on. In previous years, she could not recall assessing children's writing, play levels, or ability to join groups in the way that Tools made these so obvious. Without Tools, she stated that she might not have identified children in that regard or where other students were showing competence in this area. T2 stated that Tools had many strengths and was effective for most students.

T3's Perceptions of Contributions. Just as the previous two teachers had, T3 characterized Tools as effective academically and behaviorally for some students. As far as self-regulation development, the program helped many students to be in better control of their behavior and their learning. In particular, T3 identified the finger plays as "helping students attend." T3 did, however, recall that Tools implementation coincided with the addition of a behavioral support teacher to handle difficult behavioral situations. She was unclear if this person was taking care of the severely disruptive students or if in fact, students were more regulated in general.

T4's Perceptions of Contributions. T4 stated that the self-regulation piece worked for most. Unlike with prior programs, T4 was able to observe students working with a buddy, using mediators properly, taking turns, and other self-regulatory behaviors. She observed an increased level of play, and increased levels of writing with Tools.

T5's Perceptions of Contributions. The most striking success identified by T4 was the realization that she did not have to use a classroom behavior system as in years past. She identified the self-regulation activities as having made a tremendous difference for her students and how their self-regulation improved their academic performance. T4 also identified scaffolding as a strength of the program.

Summary

Overall, all of the teachers portrayed Tools as effectively helping some or most students develop self-regulation and academic skills. The common issues related to success included self-regulation activities and the observed increases in play, collaboration, writing, and language. Despite these gains, several challenges were also identified by teachers. Together, the perceived contributions of and challenges with Tools augment the results obtained through statistical analyses in the quantitative section. These results will support each other in the next section, which includes possible explanations for the results obtained.

CHAPTER 5

DISCUSSION

The purpose of this mixed methods study was to investigate the efficacy of the kindergarten curriculum, Tools of the Mind, on academic achievement. More specifically, this study investigated the existence of a statistically significant difference in reading and math achievement across a non-Tools year and a Tools year. A secondary purpose of this study was to uncover teachers' perceptions about challenges and successes related to the curriculum and its implementation. Quantitative data were collected and analyzed to inform the question of efficacy. Qualitative approaches were used to bring voice to the experiences of teachers as well as to offer potential explanations for the quantitative results. A review of the strengths and limitations of this study will follow the interpretations. Implications for educational practice will be discussed, followed by a review of implications for future research.

Findings and Interpretations

The previous chapter presented the data analysis to compare achievement on MAP testing across two separate years of kindergarten using two separate samples, resulting in the comparison of non-Tools and Tools groups. Specific quantitative findings will be explained, followed by explanations uncovered through qualitative interviews.

Impact of Tools on Achievement

The first and third research questions investigated the potential effect that Tools had on reading and math achievement, respectively. According to the main effects findings, Tools did not result in higher achievement in reading or math. Reading scores obtained on MAP testing at the end of the year were statistically significantly lower at the end of the year during which Tools was implemented as compared to the year without

Tools. Math scores, while not statistically significantly lower, were also lower during the Tools year. These results, however, may have been influenced by several factors that may have also contributed to the statistically significant drop in achievement after Tools was introduced. One factor was specifically related to reading. Prior to Tools implementation, the district participated in a nationally, federally funded reading program. This program funded additional Title 1 tutors to deliver scientifically research-based reading interventions. While critics have emphasized the problems with the over-emphasis of such techniques (Bredenkamp & Copple, 1997; Jarrett & Waite-Stupiansky, 2009), those methods were present during the non-Tools year and resulted in higher scores on MAP reading assessments. With Tools, those interventions were not available, and were instead replaced by self-regulation and literacy activities that were not explicit in nature. After the introduction of Tools, teachers articulated concerns about the lack of interventions, especially for those students who were struggling to gain early literacy. The potential exists for the higher original reading scores being due to these interventions and not necessarily the instruction using the typical reading curriculum.

Math scores, while not meeting significance levels, also dropped after Tools. In the non-Tools year, the average math score on MAP was 155.4 at the end of the year. Nationally, the average was 156.1 (nwea.org). After Tools, the score was even lower, at 153.0. These low math scores may reflect a general weakness in math for this district. Tools also identified math as an area of weakness (Leong, Bodrova, & Hensen, 2008). However, in the first few years, teachers reported focusing on learning the program and less on supplementing it. Teachers in this study recounted their struggles to acquire Tools theoretical and skill knowledge, a phenomenon also reported by Imholz and Petrosino

(2012). In addition, federal reading initiatives may have caused schools to pay less attention to math, resulting in lower overall math achievement. These programs may have hindered educators from realizing the relationship of early math skills to later achievement (Duncan et al., 2007).

Lower achievement after the introduction of Tools may also be explained by several other factors that pertain to both reading and math. These alternative explanations were formulated after analysis of the qualitative interviews with teacher. First, teachers articulated a general lack of reflection about the content of the Tools program and alignment with state standards. While the program manual stated that the curriculum was aligned with standards, all of the teachers interviewed reported not reflecting on standards until the second half of the second year of implementation or later. According to NWEA, MAP test questions are generated in alignment with state standards. If the instruction of the content was not aligned with state standards, then MAP scores could have reflected lack of exposure to items assessed on the test. Use of MAP relies on the assumption that the appropriate content has been taught. Second, students in the Tools year were not administered MAP at mid-year as the non-Tools group had. This difference may have given the non-Tools students an unfair advantage of a practice session. As one teacher commented, results of MAP testing for the latter group may have reflected the inexperience of the students with computerized assessments and lower computer skills in general. If lack of alignment with standards and inexperience with computers were contributing factors resulting in lower scores after Tools, the reliability of using MAP as an assessment tool should be called into question.

Impact of SES on Achievement

The second and fourth research questions investigated the existence of an effect of socioeconomics on achievement. A statistically significant effect of socioeconomics was found in both reading and math. In this study, socioeconomics levels were depicted as either free/reduced lunch or not free/reduced lunch. While Tools was described as a program that has the potential to buffer the early experiences of at-risk students (Bodrova & Leong, 2001), Tools did not have a buffering effect on academic achievement in the current study. Across both reading and math, students who received free/reduced lunch scored statistically significantly lower than those who did not receive free and reduced lunch.

Impact of Fidelity on Achievement

The third and sixth questions pertained to the potential relationship between fidelity to the program and achievement in reading and math. Fidelity results did not demonstrate a relationship with achievement. While the quantitative data revealed the absence of linear relationships between fidelity and reading and math achievement, additional and pertinent information was gleaned from the qualitative data. Qualitative interviews revealed a phenomenon that had occurred when Tools was first implemented. Tools implementation coincided with the addition of three new kindergarten teachers. Participants revealed that to the extent that could be controlled, new teachers were not assigned any students with known challenging behaviors. None of the new teachers were included as participants in this study. Deliberately placing students with challenging behaviors on these more experienced teachers may have interfered with the potential relationship between fidelity and achievement. If these classrooms had more students

with challenging behaviors, and the teachers were rated with high fidelity, the results may represent the impact of the negative behaviors on the achievement of those students and on the achievement of others. This latter suggestion again highlights the caveat in extant Tools literature related to the negative impact of peers on progress in school.

Additionally, the lack of variability in the reported fidelity scores across teachers would result in finding a statistically significant relationship. Finally, the small sample size ($n = 5$), further impacted the probability of finding significance.

Challenges Associated with Tools

Research question seven pertained to challenges experienced by teachers during Tools implementation. The challenges articulated by the teachers in this study align with those of the teachers in the other known qualitative Tools study (Imholz & Petrosino, 2012). In both groups, information was sought about the experiences of teachers in the second year of implementation. While both groups praised the quality of the actual trainings, both articulated a need for more training time. Both groups described a sense of overwhelm with the amount of new information. For both, this information was not congruent with their previous educational theory and methods experiences. This incongruence may have resulted in theoretical tension and confusion. For some of the participants in the current study, this sense of overwhelm may have had an immobilizing effect, resulting in literal adoption of the manual. This latter result may have impacted the ability of the teachers to adequately reflect on state standards and potentially adversely affecting achievement.

In a similar way, the anxiety of feeling ill-prepared coupled with insufficient theoretical understanding of the material could have contributed to the overall performance of students on achievement measures. Participants in the study conducted

by Imholz and Petrosino (2012) and in the current study expressed the lack of sufficient time to include all of the prescribed activities. In Imholz and Petrosino's (2012) study, teachers eliminated key aspects of the curriculum due to time constraints. While the current study did not include classroom observations, the possibility exists that the teachers also eliminated key components of the program due to time constraints. As Imholz and Petrosino (2012) speculated, elimination of key components may reflect lack of theoretical understanding of the role the activities play in the development of self-regulation. These decisions could have impacted achievement results as well. However, as one participant in the current study stated, the main theoretical training did not take place until after the second year of implementation. By the time the teachers acquired this knowledge, the quantitative data for this study had already been collected.

Another challenge articulated by both groups involved the transition to first grade. While one teacher showed interest in collaborating with first grade teachers (Imholz & Petrosino, 2012), all of the teachers in the current study articulated *concern* about that particular transition. Because Tools is a curriculum for preschool and kindergarten classrooms, the concern expressed by teachers may have been expressed in anticipation of the transition to the next grade. First grade teachers were not trained in Tools, but did receive a limited overview. This overview was intended to facilitate the transition between kindergarten and first grade. Despite the existence of this overview, kindergarten teachers' concerns were not allayed. According to one teacher, additional attention still needs to be paid to this transition, despite existing attempts to expose first grade teachers to Tools techniques.

In addition to the transition to first grade, teachers expressed concern about the

lack of awareness and understanding of Tools by personnel outside of kindergarten. This lack of understanding resulted in poor communication of student needs and progress at Special Education eligibility meetings and parent conferences. Tools teachers did not arrive at these meetings with the expected data from traditional programs, but rather with a new language that included dramatic play, self-regulation, dynamic assessment, and writing plans. This new language, being available only to a select few members of the school community, made collaboration with families and other school personnel challenging.

Finally, teachers reported a sense of helplessness with respect to struggling students, both academically and behaviorally. Traditional interventions were no longer available, and teachers reported a lack of adequate interventions to support students. It is possible that these interventions existed within the program, but were not yet accessed by these teachers in the phase of implementation analyzed. Students who struggled with attention did not appear to make gains in their social-emotional development with the program. This result was not expected, given the program's hallmark techniques to develop self-regulation. For these students, these self-regulation techniques did not appear to be within their ZPD's, resulting in frustration by students and teachers. One final observation regarding self-regulation was made when analyzing teacher interview transcripts. None of the teachers mentioned the use of private speech as a means to facilitate self-regulation. Given that in each classroom, some students appeared to not be at expected self-regulation levels, scaffolding private speech may have enabled students to regulate more easily, even with disabilities such as ADHD (Corkum et al., 2008).

These challenges offer insight into the experiences of teachers during the

implementation of Tools. In addition, these challenges offer highlight the supporting role of qualitative data to further explain and interpret the quantitative results. Despite the challenges, the participants identified several aspects of Tools that contributed to both student and teacher development.

Perceived Contributions of Tools

Teachers were able to articulate several important contributions enabled through Tools participation. Several of these contributions were also identified in the Imholz and Petrosino (2012) study. For example, in both studies, participants articulated that participation in trainings and in implementing the program resulted in growth in their own professional development and understanding of development. Praise was also observed for actual trainings in both groups. Both sets of participants also observed gains in student progress, but in areas not typically discussed with more traditional programs. Growth was observed in the ability of students to collaborate with peers and to use language.

Specific to the current study, teachers articulated satisfaction with observing less severely disruptive behavior with individual students. This finding may have been a result of competing factors. Because Tools implementation began at the same time as the addition of a behavioral support educator, students may have been less visible severely disrupting the hallways and classrooms; however, these students may have still been present and involved in traditional behavioral interventions. This possibility highlights the caution that is needed when describing the lower degrees of severe disruption, especially if this reduction is being attributed to Tools and self-regulation development. In theory, self-regulation would develop at higher rates in Tools classrooms reducing the

need for classroom teachers to constantly stop their instruction (Barnett et al., 2008). The potential removal of disruptive students from the classroom also created the same effect. Therefore, the question remains as to which program (Tools or behavior interventionist) caused teachers to observe less severely disruptive students. Additionally, the introduction of student specific behavioral interventions did not align with Tools and may have resulted in confusion for both the students involved and their teachers.

According to the results in this study, however, teachers reported frustration with not having adequate supports that were endorsed by Tools to assist them with disruptive behaviors. These concerns existed despite the self-regulation that was being developed in the classroom and despite the non-Tools behavioral interventions. Teachers reported worrying about the disruptive behaviors on the academic and social-emotional development of other students. The participants in this study did not report having additional time to scaffold, as implied in Barnett et al., (2008) but rather reported concern for not having more time to scaffold due to behavioral issues. These concerns were articulated by all of the participants. All participants shared, however, that these concerns involved a small number of children in each classroom, and that most of the students in their classrooms appeared to develop social-emotional, self-regulation, and academic skills according to expectations of the program. While gains were noted in many students in writing, collaboration, language, and play, these results highlight a significant caveat in the program and in the current Tools literature. Not only do teachers have to monitor the development of each student, but teachers also have to be aware of the potentially deleterious effects of disruptive behaviors of peers.

Strengths of the Current Study

Several aspects of this study demonstrate its strength as a research study. First, while other studies attempted to articulate the effects of Tools on student outcomes (Barnett et al., 2008; Diamond et al., 2007), none has been fully able to demonstrate the relationship between Tools and specific reading and math achievement. While the study conducted by Barnett et al. (2008) was able to show that in Tools classrooms, children scored higher in English and Spanish vocabulary, on teacher-reported social skills levels, on the quality of the classroom experiences, and on the frequency of scaffolding techniques, the study was not able to conclude that Tools resulted in higher achievement in emerging literacy or math. Similarly, while the study conducted by Diamond et al. (2007) was able to find higher levels of executive functioning and self-regulation in children in Tools classrooms than in children in non-Tools classrooms, academic measures were only collected from the students in the Tools classrooms. Without academic measures in both groups, the effect of Tools on achievement could not be determined. In the current study, both non-Tools and Tools classrooms were measured in reading and math achievement; therefore, this study is the first to actually measure the academic effect of Tools.

In addition to measuring the effect of Tools on achievement, this study contains other sources of strength. None of the previous research has studied students in kindergarten alone. The current study is also the first quantitative study that measures student outcomes in kindergarten. Unlike previous research that investigated Tools in its first year of implementation (Barnett et al., 2008), this study used data from the second year of implementation of Tools. This choice attempted to address concerns related to

first year acquisition of skills of the program by teachers, which may have limited the results in Barnett's study.

Other design choices also contributed to the strength of this study. First, ANCOVA was chosen for reading analysis in order to control for extraneous variables, such as baseline measures of reading. Second, in both reading and math analyses, two independent variables were selected, in order to control for as much of the variance as possible. Furthermore, no other study has combined quantitative and qualitative methods in the manner in which this study was conducted to attempt to provide a more comprehensive assessment of the effect of Tools. Despite these strengths, some limitations can be found with the study and its design.

Limitations

Several limitations can be found with the current study. These limitations are related to the design of the study, measures, the participants and the setting, and to the insider status of the researcher.

Design

While the mixed methodology of this study made this study unique thus far in the study of Tools, several limitations should be noted with the design. First, because the student participants were nested in classrooms with teachers, the impact of each teacher was not adequately controlled. A more appropriate statistic would have been Hierarchical Linear Modeling (HLM; Raudenbush, et al., 2004). This statistic would have allowed for the control of groups within the groups that themselves have qualities that influence the results, thereby potentially reducing sources of error. The small sample size of the current study resulted in the choice to use ANCOVA/ANOVA instead.

A second design issue pertained to the lack of data at the research setting. For example, the district did not assess math in a standardized manner at the beginning of the year. Baseline measures in math could not be obtained. Furthermore, MAP testing occurred at midyear and end of the year for the control group, but only at the end of the year for the Tools group. Finally, the curriculum-based measures could not be found for middle or end of the year assessments. Having three data points would have provided another source of student outcomes for analysis.

A final limitation pertains to the retrospective nature of the qualitative interviews. Participants relied on their memories from two years previous, and may have had difficulty separating year one and year two of implementation in their responses. Inquiring about previous years did not allow for observations of activities, fidelity, teacher scaffolding, and other areas of Tools. Additionally, having had insider status within the research setting, the possibility remains that my own objectivity was limited. The next section will highlight limitations related to the measures used in the study.

Measures

Limitations with measures pertain to MAP, fidelity, SES, and self-regulation. As mentioned above, several limitations were uncovered with the use of MAP testing. First, the control group was able to have a practice session at midyear while the Tools group did not. Results may have reflected the inexperience and/or motoric ability of students with computerized assessments. Second, because Tools instruction was not aligned with state standards, while MAP tests are aligned, results may have reflected the lack of instruction of concepts that were assessed. The limitations here are two-fold: validity of MAP testing with this Tools study and the validity of using Tools as the curriculum if

what is actually taught is not what is intended to be taught.

The measure of SES was used as an independent variable. The available data on SES were limited to free/reduced lunch status for this district. This data only captures one aspect of SES. Other variables may have been useful to include as measures of SES, such as mother's education level. This information was not available to the researcher.

Limitations are also found within the measures of fidelity. Because this study relied on previously collected data, and because fidelity measures were had not been collected during the initial years of Tools implementation, fidelity ratings occurred retrospectively. Three raters, including the researcher, were used. Excellent intercoder reliability was achieved. Even without counting my own ratings, the reliability of scores by raters was very high. Despite this high degree of reliability, none of the raters was considered an expert in Tools. We could have been rating degrees of fidelity based on other, less objective factors. Additionally, teachers could have received high ratings of fidelity by providing the correct types and amount of Tools activities. However, without effective classroom management and skill to scaffold, the development of self-regulation could be compromised, and as an extension, academic achievement.

A final limitation with measures pertains to self-regulation. The only data obtained in this study about self-regulation was limited to the qualitative data provided by the five teachers. Teacher reported that for the most part, they were able to witness the positive impact of self-regulation activities for most of the students, but not for all students, especially those who struggled with behavioral issues. Recollections of self-regulation may have also been influenced by how many severely disruptive students were in a classroom, by teacher management and personality style, and by teachers'

interpretations of the meaning of self-regulation. No standardized self-regulation or social skills measure was obtained, such as the ones used in previous research (Barnett et al., 2008; Diamond et al., 2007). Tools did have a measure of social-emotional development. Upon inspection of the available data, the results were deemed invalid. It was very clear that teachers were not rating in a way that demonstrated a unified understanding of self-regulation. This fact was clearly articulated in the qualitative interviews. Teachers reported personal growth in their understanding about self-regulation. In fact, teachers reported not fully understanding the theoretical information pertaining to self-regulation until the third year of implementation. For these reasons, the data on self-regulation was not deemed valid to use in this study. Because Tools is primarily about the development of self-regulation, this missing data in this study is a critical limitation. In addition to difficulty data collection and measures, other limitations were observed with the setting and participants.

Setting and Participants

Because data were collected from one single setting, the results of this study should be used with caution, as they may not generalize outside of this setting. While within each group, a sufficient sample size was obtained, only five teachers participated. The results obtained may be related specifically to characteristics of this school, of this team of teachers, and of the population of students in the district. Because the year assessed coincided with the addition of behavioral supports, mixed philosophies were being presented at the same time. Traditional behavioral techniques conflicted with Tools and may have caused confusion for a group of teachers attempting to learn a new educational philosophy and methodology. The challenges articulated by teachers may

have reflected attitudes that were opposed to the new program and could have inadvertently affected instruction and learning. Despite attempting to control for differences between the two groups in order to make comparisons after Tools, the possibility remains that the two groups were more different than they seemed. Other variables that were not controlled in the design, such as student attendance, English proficiency, and Special Education status, could have also influenced the results.

As with any study, this current study has strengths and limitations. From the understanding of these strengths and limitations, specific recommendations can be made regarding the implementation of Tools.

Recommendations

Stemming from the results of this study, several recommendations can be made that pertain to general educational practices and specific Tools-related issues. Following is a list of these recommendations:

1. Districts using Tools should ensure that the instruction is aligned with state and national standards.
2. Districts using Tools should ensure that assessment is associated with instruction and curriculum.
3. Districts that choose to use Tools should consider supplementing the curriculum with additional math and literacy programming, especially in the area of phonological awareness.
4. Before implementing Tools, districts may consider offering extended training opportunities. These opportunities may include visits to existing Tools classrooms to increase the exposure to Tools activities as well as to allow for

collaboration with teachers already experienced with Tools.

5. Tools teams and leaders should become increasingly familiar with the meaning of fidelity to the program.
6. Educational leaders should also participate in training to increase their understanding of Tools. This increased understanding may facilitate the collaboration across grade levels and in conferences with parents and other providers.
7. Additional training should be provided to facilitate the transition between kindergarten and first grade to maximize and extend the gains in self-regulation that may have been made during kindergarten.
8. Tools districts will need to decide on how purely to implement the program. While initially, providing additional academic and behavioral interventions to students in need was frowned upon, providing these additional supports might be essential for the success of students at risk for academic or behavioral difficulties.

These recommendations are based on the findings of this study. Specifically, these suggestions attempt to rectify some issues that were uncovered with Tools implementation in this study. In the next section, implications for further research are discussed.

Implications for Further Research

While several strengths were identified within this study, several limitations were also identified. Due to the small number of studies that have actually evaluated Tools, additional research is in need. Future Tools research should attempt to correct the limitations discovered in the design of the current study. Using larger samples from

multiple schools will allow the use of HLM to better account for teacher or other grouping variability. Researchers should include self-regulation measures in order to duplicate or extend previous research. Additional variables, including attendance rates, interval SES data, Special Education status, and English language proficiency, should be included. Future studies that will rely on secondary data should also attempt to use measures that have pre- and post-test data in order to increase the validity of comparisons. Because participants in this study continued to report that they were insecure with their theoretical understanding and with implementation of the program, future studies may choose to analyze data from the third year or later. Finally, future studies may investigate the potential long-term effects of the early self-regulation gains on later achievement through longitudinal studies. Because Tools is only designed as preschool and kindergarten curriculum, researchers may have difficulty extrapolating confounding variables in longitudinal research.

Conclusion

This study has attempted to assess the efficacy of the Tools of the Mind curriculum on reading and math achievement of kindergarteners. While results indicate that Tools did not have a positive effect on achievement, several confounding factors were discovered through the qualitative interviews that help further explain the decrease in achievement. These scores were limited to a small sample from only one school using a specific measure of achievement. The validity of using this measure in this study is limited. While it validly measures achievement, the MAP test assumes that appropriate instruction has taken place and that students have been adequately exposed to content being tested. Despite these results, this study was able to measure the efficacy of Tools

on achievement, demonstrate the enduring effects of SES across both curriculum types, and reveal challenges faced by teachers during implementation. Further analysis, however, is needed to determine whether Tools is providing the necessary and expected kindergarten skills to impact achievement.

APPENDIX A

CONSENT FORM FOR PARTICIPATION IN A RESEARCH STUDY

University of Massachusetts Amherst

Researcher(s): Patricia Mackay

Study Title: The Effects of Tools of the Mind on Kindergarten Math and Reading

1. What is this form?

This form is called a Consent Form. It will give you information about the study so you can make an informed decision about participation in this research.

2. Who is eligible to participate?

Kindergarten teachers who transitioned from a traditional curriculum to in Tools of the Mind are eligible to participate. Subjects must be at least 18 years old to participate.

3. What is the purpose of this study?

The purpose of this research study is to compare the achievement of students in reading and math across two instructional programs (traditional teacher-directed instruction and Tools of the Mind).

4. Where will this study take place and how long will it last?

Interviews will take place at the school (where participants are employed) either before or after school.

5. What will I be asked to do?

If you agree to take part in this study, you will be asked to participate one interview of up to 30 minutes, which will be recorded, transcribed, and analyzed. Interview questions will explore your experiences as you transitioned to Tools of the Mind and your experiences in teaching it since you began. You may skip any question you feel uncomfortable answering.

6. What are my benefits of being in this study?

You may not directly benefit from this research; however, I hope that your participation in the study may assist the advancement of knowledge.

7. What are my risks of being in this study?

I believe there are no known risks associated with this research study; however, a possible inconvenience may be the time it takes to complete the study.

8. How will my personal information be protected?

The following procedures will be used to protect the confidentiality of your study records. Interviews will be recorded and notes may be taken. Following the interview, the audio will be transcribed for analysis. The researchers will keep all study records, including any codes to your data, in a locked file cabinet. Research records will be labeled with a code. A master key that links names and codes will be maintained in a separate and secure location. The master key and audiotapes will be destroyed 3 years after the close of the study. All electronic files (databases and spreadsheets, transcription notes) containing identifiable information will be password protected. Any computer hosting such files will also have password protection to prevent access by unauthorized users. Only the members of the research staff will have access to the passwords. At the conclusion of this study, the researchers may publish their findings. Information will be presented in summary format and you will not be identified in any publications or presentations.

In the course of the interview, if information is disclosed that warrants breaching confidentiality, such as reports of abuse or neglect of a student, or intent to harm self or others, confidentiality cannot be guaranteed.

9. Will I receive any payment for being in this study??

N/A

10. What if I have questions?

Take as long as you like before you make a decision. I will be happy to answer any question you have about this study. If you have further questions about this project or if you have a research-related problem, you may contact the researcher, (*Patricia Mackay, at 508-949-7480*) or the Chairperson of this Dissertation Committee (*Dr. Kevin Nugent, at 857-218-4351*). If you have any questions concerning your rights as a research subject, you may contact the University of Massachusetts Amherst Human Research Protection Office (HRPO) at (413) 545-3428 or humansubjects@ora.umass.edu.

11. Can I stop being in this study?

You do not have to be in this study if you do not want to. If you agree to be in the study, but later change your mind, you may drop out at any time. There are no penalties or consequences of any kind if you decide that you do not want to participate.

12. What if I am injured?

The University of Massachusetts does not have a program for compensating subjects for injury or complications related to human subjects research, but the study personnel will

assist you in getting treatment.

13. Subject statement of voluntary consent:

When signing this form I am agreeing to voluntarily enter this study. I have had a chance to read this consent form, and it was explained to me in a language which I use and understand. I have had the opportunity to ask questions and have received satisfactory answers. I understand that I can withdraw at any time. A copy of this signed Informed Consent Form has been given to me.

Participant Signature:

Print Name:

Date:

By signing below I indicate that the participant has read and, to the best of my knowledge, understands the details contained in this document and has been given a copy.

Signature of Person
Obtaining Consent

Print Name:

Date:

APPENDIX B
POWER ANALYSES

Figure B.1 Power Analysis for .40 effect size

XLSTAT 2013.2.03 - Power - ANOVA/ANCOVA

Goal: Find the sample size

Tests: ANOVA Factors and interactions

Determine effect size: Effect size

Inputs

Parameters	Inputs
Number of groups:	2
Num DF:	1

Results:

Parameters	Results
Power	0.8
alpha	0.05
Effect size	0.4
Sample size	51
Power (obtained)	0.800

Test interpretation:

H0: The means of the groups of the tested factor are equal.

Ha: At least one of the means is different from another.

The risk to not reject the null hypothesis H0 while it is false is 0.2.

For the given parameters, for an alpha of 0.05, the necessary sample size to reach a power of 0.8 is 51 observations.

Figure B.2 Power Analysis for .25 effect size

XLSTAT 2013.3.05 - Power - ANOVA/ANCOVA

Goal: Find the sample size

Tests: ANOVA Factors and interactions

Determine effect size: Effect size

Parameters	Inputs
Number of groups:	2
Num DF:	1

Results:

Parameters	Results
Power	0.8
alpha	0.05
Effect size	0.25
Sample size	128
Power (obtained)	0.801

Test interpretation:

H₀: The means of the groups of the tested factor are equal.

H_a: At least one of the means is different from another.

The risk to not reject the null hypothesis H₀ while it is false is 0.2.

For the given parameters, for an alpha of 0.05, the necessary sample size to reach a power of 0.8 is 128 observations.

APPENDIX C

TOOLS- K PROJECT FIDELITY OF IMPLEMENTATION SCALE

Scale:

- 0 – none of Tools activities are implemented. Tools materials are not present or are used in non-Tools activities

- 1 - very few activities implemented, those that are, may be missing key elements (example: reading with buddies but without reading strategies or decoders) Teachers current practice often runs counter to what Tools is after - teacher directed, whole group instruction

- 2- a number of activities implemented, some of them correctly, but many not - and implementation isn't consistent; other practices that don't support self-regulation continuing (worksheets, etc.)

- 3- most activities implemented, some going well, a few that still have errors. Children's progress in SW has been strong - scaffolding interactions are close to what Tools is after (in child's ZPD, just 'one more thing') Teacher is reflective - aware of what she's doing 'right' and what isn't quite 'right' and she's aware of children's development and thinks about it as she interacts with them and plans.

- 4- fidelity is good - most or all activities have been implemented and are implemented faithfully. The room feels like a 'Tools' classroom - a lot of peer scaffolding, independence, no to little teacher regulation needed, transitions are tight, minimal time in whole group, teacher scaffolding is on the mark.

- 5 - fidelity is excellent - on top of everything listed in 4, the nuances of how to implement activities and use tactics are obvious in practice - teacher may be taking her understanding into new activities she is innovating; teacher is attending to challenge level and children's ZPDs, self-regulation in the group feels high

APPENDIX D

INTRACLASS CORRELATION COEFFICIENT

Figure D.1 Intraclass Correlation Coefficient

Number of subjects (n)	5
Number of raters (k)	3
Model	The same raters for all subjects. Two-way model.
Type	Consistency
Measurements	Rater1 Rater2 Rater3

	Intraclass correlation ^a	95% Confidence Interval
Single measures ^b	0.9592	0.8143 to 0.9953
Average measures ^c	0.9860	0.9293 to 0.9984

^a The degree of consistency among measurements.

^b Estimates the reliability of single ratings.

^c Estimates the reliability of averages of *k* ratings.

APPENDIX E

INTERVIEW QUESTIONS FOR TEACHERS QUESTIONS 6 & 7

1. What is your recollection of the reason(s) Tools was chosen for this school?
2. Think back to the preparation you received prior to and during implementation.
What are your thoughts about training?
3. I would like you to think about the state standards, and if you could, please comment on how aligned you believe your instruction and the content of Tools was to these state standards.
4. What do you consider to be benefits of this program?
5. What do you consider to be challenges with this program?
6. Is there anything else that you would like to tell me about your experiences transitioning to Tools from the previous curriculum?

APPENDIX F
ANCOVA OUTPUTS

Figure F.1 ANCOVA Linearity Assumption Scatterplot-Overall

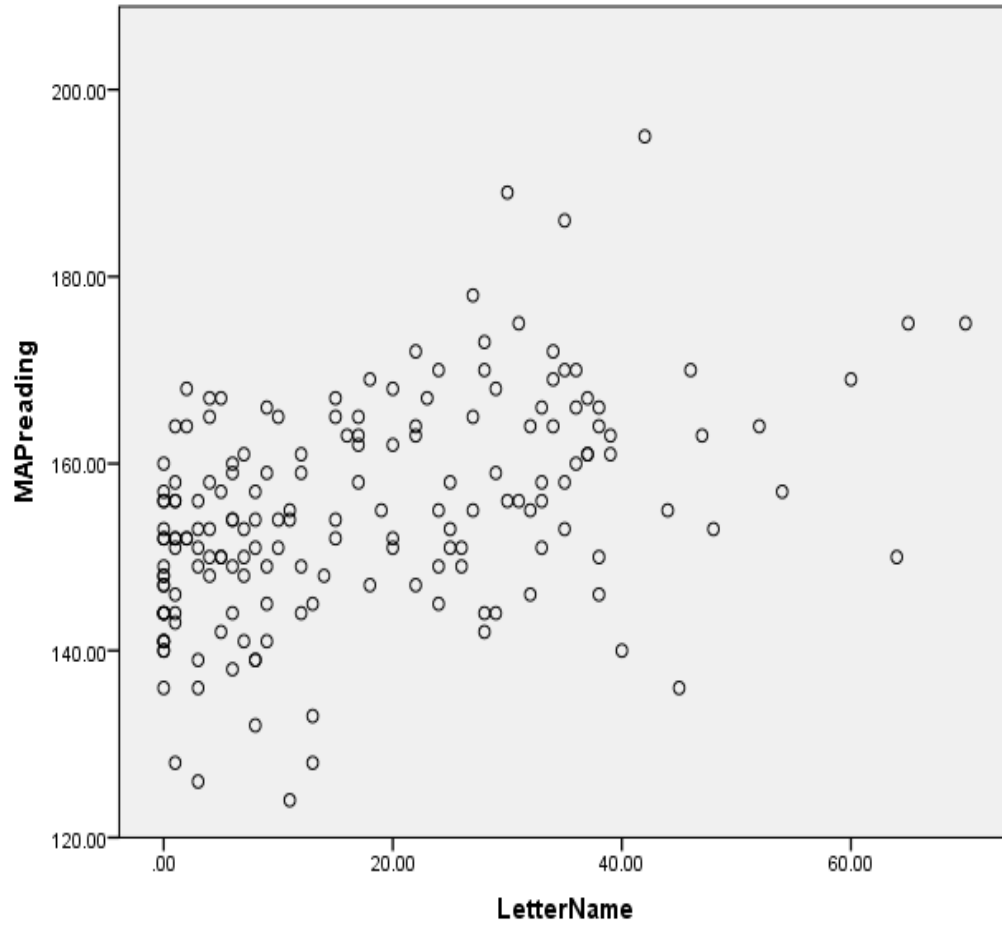


Figure F.2 ANCOVA Linearity Assumption Scatterplot-by Tools Year 2009

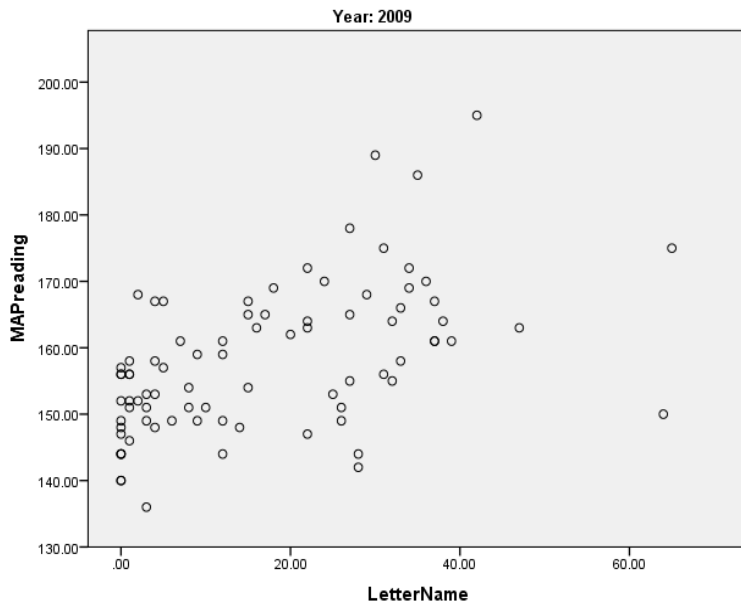


Figure F.3 ANCOVA Linearity Assumption Scatterplot-by Tools Year 2011

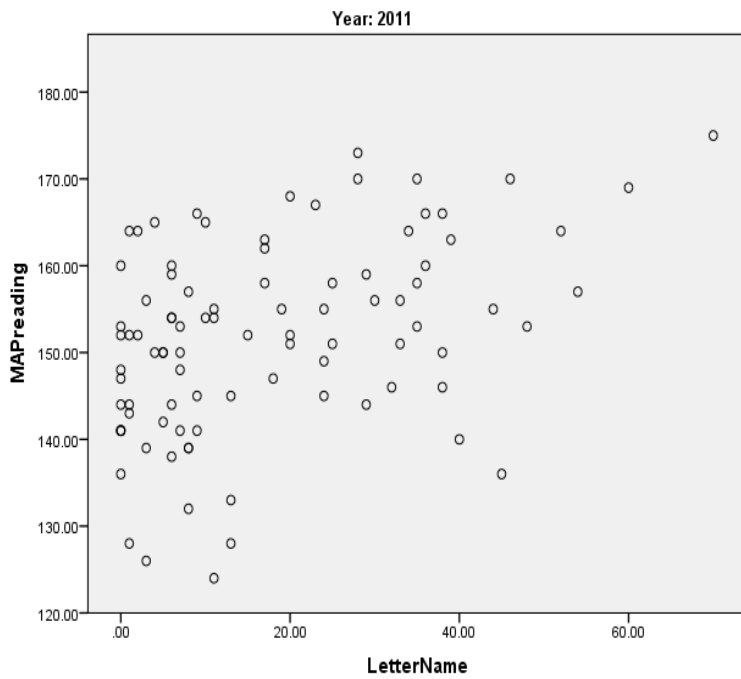


Figure F.4 ANCOVA Linearity Assumption Scatterplot-by SES Free/Reduced Lunch

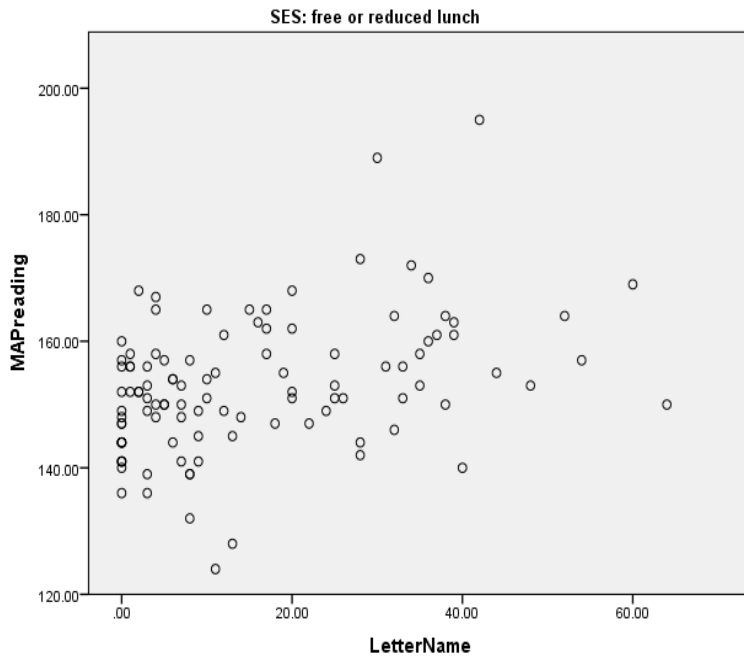


Figure F.5 ANCOVA Linearity Assumption Scatterplot-by SES Not Free or Reduced Lunch

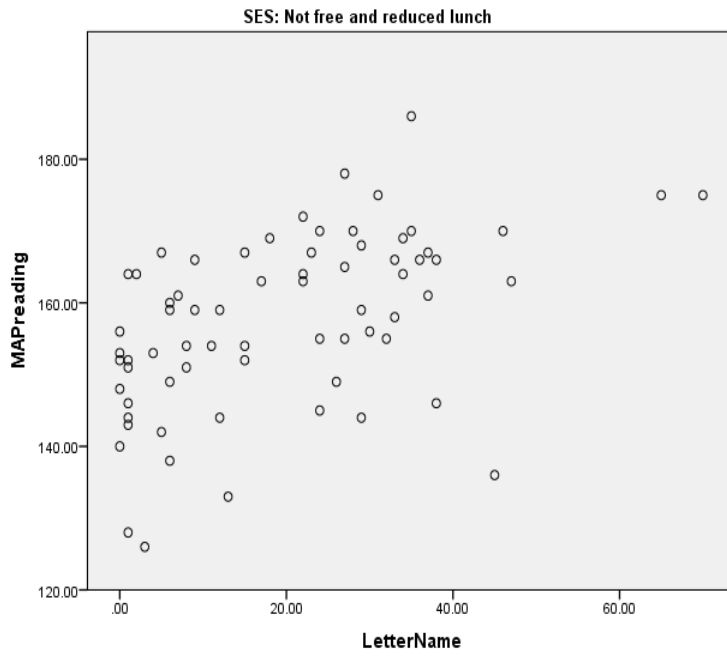


Table F.1 Full ANCOVA Output

Tests of Between-Subjects Effects

Dependent Variable: MAPreading

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6816.917 ^a	5	1363.383	14.271	.000
Intercept	1605848.925	1	1605848.925	16808.905	.000
Year	498.654	1	498.654	5.220	.024
SES	42.363	1	42.363	.443	.506
LetterName	4393.251	1	4393.251	45.985	.000
Year * LetterName	46.158	1	46.158	.483	.488
SES * LetterName	61.102	1	61.102	.640	.425
Error	16145.518	169	95.536		
Total	4214018.000	175			
Corrected Total	22962.434	174			

a. R Squared = .297 (Adjusted R Squared = .276)

Table F.2 ANCOVA Homogeneity of Variance Assumption: Levene's Test

Levene's Test of Equality of Error Variances^a

Dependent Variable: MAPreading

F	df1	df2	Sig.
1.412	3	171	.241

Figure F.6 ANCOVA Homogeneity of Variance Scatterplots

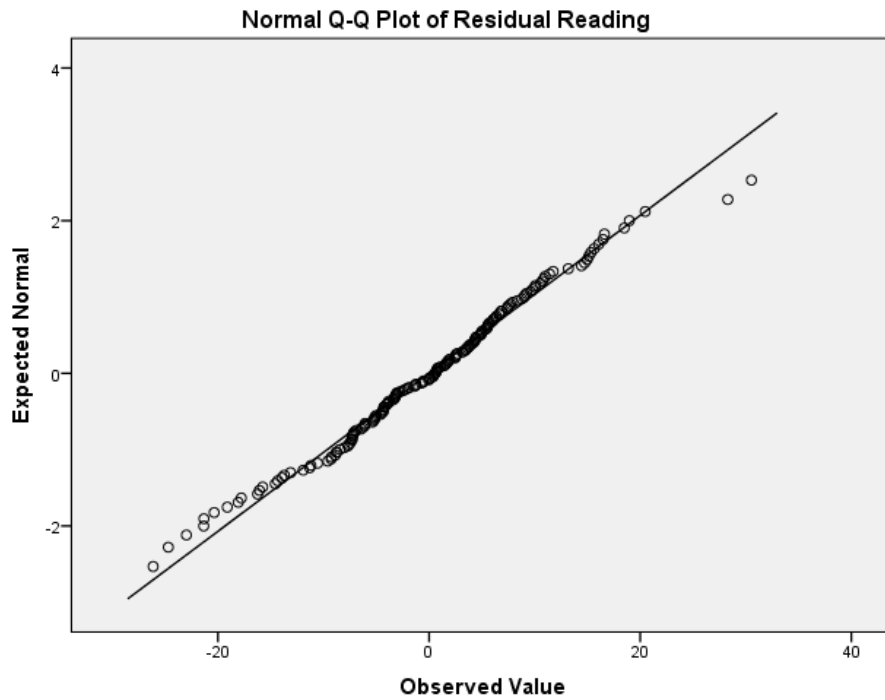


Table F.3 ANCOVA Normality Assumption-Shapiro-Wilk

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual Reading	.047	175	.200 [*]	.991	175	.317

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table F.4 ANCOVA Normality Assumption-Skewness & Kurtosis

Descriptives			Statistic	Std. Error
Residual Reading	Mean		.0000	.73050
	95% Confidence Interval for Mean	Lower Bound	-1.4418	
		Upper Bound	1.4418	
	5% Trimmed Mean		.0622	
	Median		.5454	
	Variance		93.385	
	Std. Deviation		9.66360	
	Minimum		-26.14	
	Maximum		30.54	
	Range		56.68	
	Interquartile Range		12.01	
	Skewness		-.032	.184
	Kurtosis		.556	.365

Table F.5 ANCOVA Independence of the Covariate and the Independent Variable Assumption for SES

IV=SES

Independent Samples Test					
		Levene's Test for Equality of Variances		t-test for Equality of Means	
		F	Sig.	t	df
LetterName	Equal variances assumed	.124	.725	-1.572	173
	Equal variances not assumed			-1.563	142.487

Table F.6 ANCOVA Independence of the Covariate and the Independent Variable
Assumption for Tools Year

IV= Tools Year

Independent Samples Test					
		Levene's Test for Equality of Variances		t-test for Equality of Means	
		F	Sig.	t	df
LetterName	Equal variances assumed	.164	.686	-.321	173
	Equal variances not assumed			-.322	172.537

APPENDIX G
ANOVA OUTPUTS

Table G.1 Full ANOVA Output Table

Tests of Between-Subjects Effects

Dependent Variable: MAPreading

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2322.312 ^a	3	774.104	6.214	.000	.091
Intercept	4358832.953	1	4358832.953	34990.827	.000	.995
SES	879.284	1	879.284	7.059	.009	.037
Year	1328.194	1	1328.194	10.662	.001	.054
SES * Year	32.823	1	32.823	.263	.608	.001
Error	23170.156	186	124.571			
Total	4568259.000	190				
Corrected Total	25492.468	189				

a. R Squared = .091 (Adjusted R Squared = .076)

Table G.2 ANOVA Normality Assumption: Shapiro-Wilk

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual MATH	.056	190	.200 [*]	.993	190	.474

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table G.3 ANOVA Normality Assumption-Skewness and Kurtosis

Descriptives				Statistic	Std. Error
Residual MATH	Mean			.0000	.91623
	95% Confidence Interval for Mean	Lower Bound		-1.8073	
		Upper Bound		1.8073	
	5% Trimmed Mean			.0841	
	Median			.1475	
	Variance			159.500	
	Std. Deviation			12.62931	
	Minimum			-31.00	
	Maximum			32.59	
	Range			63.59	
	Interquartile Range			17.07	
	Skewness			-.094	.176
	Kurtosis			-.164	.351

Table G.4 ANOVA Homogeneity of Variances: Levene's Test

Levene's Test of Equality of Error Variances^a

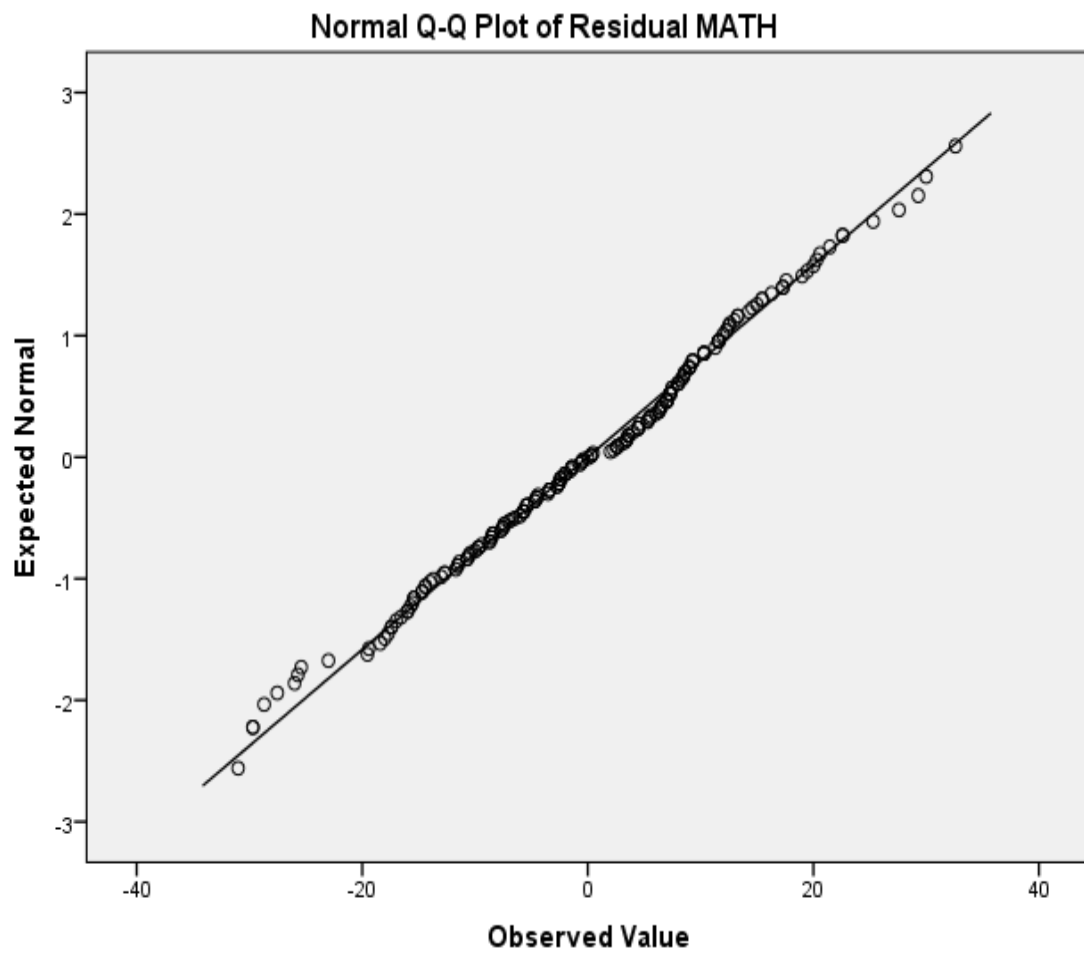
Dependent Variable: MAPreading

F	df1	df2	Sig.
1.065	3	186	.365

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.^a

a. Design: Intercept + SES + Year + SES * Year

Figure G.1 ANOVA Homogeneity of Variances Scatterplot



APPENDIX H

FIDELITY CORRELATIONS

Table H.1 Correlation between Fidelity and MAP Reading Scores

		Correlations	
		Fidelity	MAPreading
Fidelity	Pearson Correlation	1	.085
	Sig. (2-tailed)		.246
	N	190	190
MAPreading	Pearson Correlation	.085	1
	Sig. (2-tailed)	.246	
	N	190	190

Table H.2 Correlation between Fidelity and MAP Math Scores

		Correlations	
		Fidelity	MAPmath
Fidelity	Pearson Correlation	1	.109
	Sig. (2-tailed)		.135
	N	190	190
MAPmath	Pearson Correlation	.109	1
	Sig. (2-tailed)	.135	
	N	190	190

REFERENCES

- Agina, A. M., Kommers, P. A. M., Steehouder, M. F. (2011). Group Awareness in CSCL Environments. The effect of the external regulator's absence on children's speech use, manifested self-regulation, and task performance during learning tasks. *Computers in Human Behavior, 27*(3), 1118-1128.
- Armbruster, B. B., Lehr, F., Osborn, J., Adler, C. R., & National Institute for Literacy (U.S.). (2001). *Put reading first: The research building blocks for teaching children to read: kindergarten through grade 3*. Washington, D.C.: National Institute for Literacy, National Institute of Child Health and Human Development, U.S. Dept. of Education.
- Au, K. H. (2005). Negotiating the slippery slope: School change and literacy achievement. *Journal of Literacy Research, 37*(3), 267-288.
- Barkley, R. A. (1997b). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65-94.
- Barkley, R. A., Murphy, K. R., Fischer, M. (2008). *ADHD in adults: What the science says*, (171 - 175). New York, Guilford Press.
- Barnett, W., Jung, K., Yarosz, D., Thomas, J., Hornbeck, A, Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: a randomized trial. *Early Childhood Research Quarterly, 23*(3), 299-313.
- Berk, L. E., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education*. Washington, DC: NAEYC.

- Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21*(4), 327-336.
- Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., & ... Gill, S. (2008). Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program. *Child Development, 79*(6), 1802-1817. doi:10.1111/j.1467-8624.2008.01227.x
- Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*(2), 111-127. doi:10.1037/0003-066X.57.2.111
- Blair, C., & A. Diamond. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure.” *Development and Psychopathology, 20*, 899–911.
- 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*(2), 647-663.
- Blair, C., Zelazo, P., & Greenberg, M. T. (2005). The Measurement of Executive Function in Early Childhood. *Developmental Neuropsychology, 28*(2), 561-571.
doi:10.1207/s15326942dn2802_1
- Bodrova, E., & Leong, D. J. (1996). *Tools of the mind: The Vygotskian approach to early childhood education* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Bodrova, E., & Leong, D. J. (2005). Why Children Need Play. *Early Childhood Today, 20*(1), 6-7.

- Bodrova, E. & Leong, D. J. (2007). *Tools of the mind: The Vygotskian approach to early childhood education*. 2nd ed., Columbus, OH: Merrill/Prentice Hall.
- Bodrova, E., Leong, D. J., & Akhutina, T. V. (2011). When everything new is well-forgotten old: Vygotsky/Luria insights in the development of executive functions. *New Directions For Child & Adolescent Development*, 2011(133), 11-28.
doi:10.1002/cd.301
- Bodrova, E., Leong, D., & International Bureau of Education. (2001). *Tools of the mind: A case study of implementing the Vygotskian approach in American early childhood and primary classrooms*. Geneva: International Bureau of Education.
- Bredenkamp, S., & Copple, C. (Eds.). (1997). *Developmentally appropriate practice for early childhood programs*. (Revised edition). Washington, DC: NAEYC.
- Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, 24(3), 337-349.
- Bronfenbrenner, U. (1989). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Brown, Thomas E. (2005). *Attention deficit disorder: The unfocused mind in children and adults* (pp 20 - 58). New Haven, CT: Yale University Press Health and Wellness.
- Bruner, J. S. (1960). *The process of education*. Cambridge: Harvard University Press.
- Bruner, J. S. (1972). Nature and uses of immaturity. *American Psychologist*, 27, 687-708.

- Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability: Inhibition, Switching, and Working Memory. *Developmental Neuropsychology, 19*(3), 273-293.
- Bull, R., Espy, K., & Wiebe, S. A. (2008). Short-Term Memory, Working Memory, and Executive Functioning in Preschoolers: Longitudinal Predictors of Mathematical Achievement at Age 7 Years. *Developmental Neuropsychology, 33*(3), 205-228. doi:10.1080/87565640801982312
- Clarke, P., Hulme, C., & Snowling, M. (2005). Individual differences in RAN and reading: A response timing analysis. *Journal of Research in Reading, 28*, 73–86.
- Clark, C., Pritchard, V., & Woodward, L. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology, 46*(5), 1176-1191.
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Collette, F., Hodge, M., Salmon, E., & Van Der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience, 139*, 209-221.
- Corkum, P., Humphries, K., Mullane, J. C., & Theriault, F. (2008). Private speech in children with ADHD and their typically developing peers during problem-solving and inhibition tasks. *Contemporary Educational Psychology, 33*, 1, 97-115.

- Cronin, J., Kingsbury, G.G., Dahlin, M., & Bowe, B. (2007, April). *Alternate methodologies for estimating state standards on a widely used computer adaptive test*. Paper presented at the American Educational Research Association, Chicago, IL.
- Curwood, J. S. (2007). What happened to kindergarten? *Instructor*, 117(1), 28-30.
- Daugherty, M. & White, C. S. (2008). Relationships among private speech and creativity in head start and low-socioeconomic status preschool children." *Gifted Child Quarterly*, 52(1), 30-9.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A., (2006). Advances in Developmental Cognitive Neuroscience. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037-2078.
- Deno, S. D. (2003). Developments in curriculum-based measurement. *Journal of Special Education*, 37, 184-192.
- Deno, S. L., Espin, C. A., & Fuchs, L. S. (2002). Evaluation strategies for preventing and remediating basic skill deficits. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventative and remedial approaches* (pp. 213-241). Bethesda, MD: National Association of School Psychologists.
- Department of Elementary and Secondary Education. Retrieved from doe.mass.edu
- Dewey, J. (1938). *Experience and education*. New York: Macmillan.

- Diamond, A. (2006). The early development of executive functions. In E. Bialstock & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70-95). Oxford: Oxford University Press.
- Diamond, A., Barnett, S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*(30), 1387-1388.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning language with literacy: Young children learning at home and school*. Baltimore, MD: Brookes Publishing.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 6, 1428-46.
- Duncan, L. G., & Seymour, P.H.K. (2000). Socio-economic differences in foundation level literacy. *British Journal of Psychology*, *91*, 145-166.
- Edformation. (2005). AIMSweb progress monitoring and improvement system. Available from <http://www.aimsweb.com/>
- Elias, C. L., & Berk, L. E. (2002). Self-regulation in young children: Is there a role for sociodramatic play?. *Early Childhood Research Quarterly*, *17*(2), 216-38.
- Elkonin, D. B. (2005). Chapter 3: Theories of play. *Journal Of Russian & East European Psychology*, *43*(2), 3-89.
- Ertmer, P.A., & Newby, T.J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, *6*(4), 50-72.

- Fantuzzo, J., Sekino, Y., & Cohen, H. L. (2004). An examination of the contributions of interactive peer play to salient classroom competencies for urban head start children. *Psychology in the Schools, 41*(3), 323-336.
- Fernyhough, C., & Fradley, E. (2005). Private speech on an executive task: relations with task difficulty and task performance. *Cognitive Development, 20*(1), 103-120.
doi:10.1016/j.cogdev.2004.11.002
- Fleiss, J. L., Levin, B., & Paik, M. C. (2003) *Statistical methods for rates and proportions* (3rd ed.). Hoboken, N.J.: John Wiley & Sons.
- Gersten, R., Woodward, J., & Darch, C. (1986). Direct instruction: A research-based approach to curriculum design and teaching. *Exceptional Children, 53*(1), 17-31.
- Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Longmont, CO: Sopris West.
- Good, R. H., Kaminski, R. A., Shinn, M., Bratten, J., Shinn, M., Laimon, L., Smith, S., & Flindt, N. (2004). *Technical Adequacy and Decision Making Utility of DIBELS* (Technical Report No. 7). Eugene, OR: University of Oregon.
- Gunnar, M. & Quevado, K. (2007). The neurobiology of stress and development. *Annual Reviews of Psychology, 58*, 145-173.
- Hanline, M. F., Milton, S., & Phelps, P. C. (2008). A longitudinal study exploring the relationship of representational levels of three aspects of preschool sociodramatic play and early academic skills. *Journal of Research in Childhood Education, 23*(1), 19-28.

- Henderson, S. D., Many, J. E., Wellborn, H. P., & Ward, J. (2002). How scaffolding nurtures the development of young children's literacy repertoire: Insiders' and outsiders' collaborative understandings. *Reading Research and Instruction, 41*(4), 309–330.
- Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: An fMRI meta-analysis of 52 studies including 842 children. *Developmental Science, 6*, 876-885.
- Howse, R., Calkins, S. D., Anastopoulos, A., Keane, S. & Shelton, T. (2003) Regulatory contributors to children's kindergarten achievement. *Early Education and Development, 14*, 101-119.
- Hyson, M., Copple, C., & Jones, J. (2006). Early childhood development and education. In K. A. Renninger, I. E. Sigel, W. Damon, & R. M. Lerner (Eds.), *Handbook of Child Psychology (4)*, 3-47.
- Imholz, S. & Petrosino, A. (2012). Teacher observations on the implementation of the tools of the mind curriculum in the classroom: Analysis of interviews conducted over a one-year period. *Creative Education, 3*, 185-192.
- Jarrett, O. & Waite-Stupiansky. (2009). Recess: It's indispensable. *Young Children, 64*(5), 66-69.
- Karpov, Y. V. (2005). *The neo-Vygotskian approach to child development*. New York, NY: Cambridge University Press.

- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*(6), 781-791.
- Kozulin, A., & Presseisen, B. Z. (1995). Mediated learning experience and psychological tools: Vygotsky's and Feuerstein's perspectives in.. *Educational Psychologist, 30*(2), 67.
- Krafft, K. & Berk, L. (1998). Private speech in two preschools: Significance of open-ended activities and make-believe play for verbal self-regulation. *Early Childhood Research Quarterly, 13*(4), 637-658.
- Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.
- Lennon, J. M., Li-Grining, C., Raver, C. C., Pess, R. A., & Society for Research on Educational Effectiveness (SREE). (2011). *Long-Term Follow Up of CSRP: Understanding Students' Academic Achievement Post-Treatment*. Society for Research on Educational Effectiveness. 2040 Sheridan Road, Evanston, IL 60208.
- Leong, D. J., Bodrova, E., & Hensen, R. (2008). *Tools of the mind: Kindergarten manual*. 6th ed., Center for Improving Early Learning, Metropolitan State College of Denver, CO.

- Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology, 46*(5), 1062-1077. doi:10.1037/a0020066
- Lidstone, J. M., Meins, E., & Fernyhough, C. (2012). Verbal mediation of cognition in children with specific language impairment. *Development & Psychopathology, 24*(2), 651-660. doi:10.1017/S0954579412000223
- Lord, F. M. (1953). On the statistical treatment of football numbers. *American Psychologist, 8*(12), 750-751. doi:10.1037/h0063675
- Luria, A. (1961). The role of speech in the regulation of normal and abnormal behavior. New York, NY: Liveright Publishing Corp.
- Luria, A. R. (1973). *The working brain: An introduction to neuropsychology* (B. Haigh, trans.). New York: Basic Books.
- Luria, A. R. (1979). The making of the mind: A personal account of Soviet psychology. (M. Cole & S. Cole, Trans.). Cambridge, MA: Harvard University Press.
- Luria, A. R. (2002). L.S. Vygotsky and the problem of functional localization. *Journal of Russian and East European Psychology, 40*(1), 17–25.
- 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*(4), 471-490.
- 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*(4), 947-959.

- Merriam, S. B. (2009). *Qualitative Research: A guide to design and implementation*.
Jossey-Bass: San Francisco, CA.
- Miller, E. & Almon, J. (2009). *Crisis in the kindergarten: Why children need to play in school*. Retrieved from http://steinerireland.org/articles/kindergarten_report.pdf.
- NAEYC (2009). *Developmentally appropriate practice in early childhood programs serving children from birth to age 8*. Washington, DC: Self-published.
- Nicolopoulou, A., Barbosa de Sá, A., Ilgaz, H., & Brockmeyer, C. (2010). Using the Transformative Power of Play to Educate Hearts and Minds: From Vygotsky to Vivian Paley and Beyond. *Mind, Culture & Activity*, 17(1), 42-58.
doi:10.1080/10749030903312512
- Noel Foulin, J. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing*, 18(2), 129-155.
doi:<http://dx.doi.org/10.1007/s11145-004-5892-2>
- Northwest Evaluation Association. Retrieved from nwea.org
- O’Hearn, K., Asato, M., Ordaz, S., & Luna, B. (2008). Neurodevelopment and executive function in autism. *Development and Psychopathology*, 20(4), 1103-1132.
- Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, 72, 816–833.
- Phillips, B., Clancy-Menchetti, J., & Lonigan, C. (2008). Successful phonological awareness instruction with preschool children. *Topics in Early Childhood Special Education*, 28(1), 3-17.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

- Pintrich, P.R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P.R. Pintrich, and M. Zeidner (Eds.), *Handbook of self-regulation*, 451-502. San Diego, CA: Academic Press.
- Ponitz, C., McClelland, M. M., Matthews, S. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, *45*(3), 605-619. doi:10.1037/a0015365
- Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology*, *58*, 1-23.
- Raudenbush, S. W. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, Ill: Scientific Software Int. Inc.
- Raver, C. C. (2004). Placing emotional self-regulation in sociocultural and socioeconomic contexts. *Child Development*, *75*, 346-353. doi: 10.1111/j.1467-8624.2004.00676.x
- Raver, C. (2012). Low-income children's self-regulation in the classroom: Scientific inquiry for social change. *American Psychologist*, *67*(8), 681-689. doi:10.1037/a0030085
- Raver, C. C., Li-Grining, C., Bub, K., Jones, S. M., Zhai, F., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: Self-regulation as a mediating mechanism. *Child Development*, *8*(1), 362-378.
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135-147.
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contributions of children's self-regulation and classroom quality to

- children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, 45(4), 958-972. doi:10.1037/a0015861
- Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, 21(5),491-511.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Rogoff, B. (1995) Observing Sociocultural Activity on Three Planes: Participatory appropriation, guided participation, and apprenticeships. In J.V. Wertsch, P. Del Rio, and A. Alvarez (Eds.). *Sociocultural Studies of Mind*, pp.139-164. Cambridge: Cambridge University Press.
- Rose, L. T., & Fischer, K. W. (in press). Dynamic systems theory. In R. A. Shweder (Ed.), *Chicago companion to the child*. Chicago: University of Chicago Press.
- Röthlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., & Roebbers, C. M. (2012). Improving executive functions in 5- and 6-year-olds: Evaluation of a small group intervention in prekindergarten and kindergarten children. *Infant & Child Development*, 21(4), 411-429. doi:10.1002/icd.752
- Rutherford, A. (2011). *ANOVA and ANCOVA: A GLM approach*. Hoboken, N.J: Wiley.
- Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development*, 81(1), 6-22.
- Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly*, 23(1), 7-25.

- Sektnan, M., McClelland, M. M., Acock, A., & Morrison, F. J. (2012). Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Childhood Research Quarterly, 25*, 4, 464-479.
- Shonkoff, J. P., Phillips, D., & National Research Council (U.S.). (2000). *From neurons to neighborhoods: The science of early child development*. Washington, D.C: National Academy Press.
- Singer, B. D., & Bashir, A. S. (1999). What are executive functions and self-regulation and What do they have to do with language-learning disorders? *Language, Speech, and Hearing Services in Schools, 30*(3), 265-73.
- Snow, K. L. (2006). Measuring school readiness: Conceptual and practical considerations. *Early Education and Development, 17*(1), 7-41.
- IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.
- Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Saenz, L., Yen, L., et al. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*, 368-388.
- Thelen, E. & Smith, L. B. (1998). Dynamic systems theories. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.). *The handbook of child psychology: Vol. 1. Theoretical models of human development*, 5th ed., 563-634. New York, NY: John Wiley & Sons.

- Thorell, L. B., Lindqvist, S., Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 12(1), 106-113. doi:10.1111/j.1467-7687.2008.00745.x
- van Ewijk, R., & Slegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134-150.
- Vygotsky, L. S. (1933/1962). *Thought and language* (E. Hanfmann & G. Vakar, Eds. & Trans.). Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1987). Thinking and speech. In R. W. Rieber, A. S. Carton (Eds.), & N. Minick (Trans.), *The collected works of L. S. Vygotsky: Vol. 1. Problems of general psychology* (pp. 37-285). New York: Plenum. (Original work published 1934).
- Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The Development of Cognitive Skills and Gains in Academic School Readiness for Children From Low-Income Families. *Journal Of Educational Psychology*, 102(1), 43-53. doi:10.1037/a0016738
- Winsler, A., Manfra, L., & Diaz, R. M. (2007). "Should I let them talk?": Private speech and task performance among preschool children with and without behavior problems. *Early Childhood Research Quarterly*, 22(2), 215-231. doi:10.1016/j.ecresq.2007.01.001
- Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Child Psychology and Psychiatry*, 17(1), 89-100.

XLSTAT, Version 2013.2. Retrieved from xlstat.com.

Yang, O. S. (2000). Guiding children's verbal plan and evaluation during free play: An application of Vygotsky's genetic epistemology to the early childhood classroom. *Early Childhood Education Journal*, 28(1), 3-10.

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