Executive Function Variance in a School-Based Behavioral Screener

Ryan Lenz

Follow this and additional works at: https://dsc.duq.edu/etd

Recommended Citation

This Immediate Access is brought to you for free and open access by Duquesne Scholarship Collection. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Duquesne Scholarship Collection.
EXECUTIVE FUNCTION VARIANCE IN A SCHOOL-BASED BEHAVIORAL SCREENER

A Dissertation
Submitted to the School of Education

Duquesne University

In partial fulfillment of the requirements for
the degree of Doctor of Philosophy

By
Ryan V. Lenz

May 2012
DUQUESNE UNIVERSITY
SCHOOL OF EDUCATION
Department of Counseling, Psychology, and Special Education

Dissertation

Submitted in partial fulfillment of the requirements
for the degree
Doctor of Philosophy (Ph.D.)

School Psychology Doctoral Program

Presented by:

Ryan V. Lenz
B.S. Psychology, Allegheny College, 1995
M.S. Ed. School Psychology, Duquesne University, 1996
Certificate of Advanced Graduate Study, Duquesne University, 1999

March 9, 2012

EXECUTIVE FUNCTION VARIANCE IN A SCHOOL-BASED BEHAVIORAL
SCREENER

Approved by:

Jeffrey A. Miller, Ph.D., ABPP
Professor and Associate Provost
Associate Academic Vice President for Administration
Duquesne University

Tammy L. Hughes, Ph.D.
Fr. Martin A. Hehir Endowed Chair for Scholarly Excellence
Professor & Chair, Department of Counseling, Psychology and Special Education
Duquesne University

Ara J. Schmitt, Ph.D.
Assistant Professor
Department of Counseling, Psychology, and Special Education
Duquesne University
ABSTRACT

EXECUTIVE FUNCTION VARIANCE IN A SCHOOL-BASED BEHAVIORAL SCREENER

By
Ryan V. Lenz

May 2012

Dissertation supervised by: Jeffrey A. Miller, Ph.D.

The implementation of school-wide behavior supports is considered typical educational practice. A main component of school-wide behavior support is the systematic screening for behavioral and emotional problems. Students’ ability to utilize executive functions greatly affects both academic and social success and it is associated with behavior and emotional problems (Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003). This study uses a sample of fourth grade students to analyze the relationship between executive function skills and a screening measure of behavior. This study will examine if deficits in executive function may account for variance in the first level of behavior support screening systems. This may help educators understand the relationship between behavior screening results and executive functions, and possibly explain how social and emotional difficulties may be affected by deficits in executive functions.
DEDICATION

I dedicate this dissertation to family, first of all, to my boys Connor, Bradan and Eli. I started this process when Bradan was just born, and Eli was born a year after I started. I love you boys, and am already proud of how you’ve grown into nice boys who like to learn. I dedicate this to the Torio family for their continued support. I dedicate this to Melanie and Will, two of my newest family members. I dedicate this to my Dad for always being there- although he was a steelworker, I think he could have been a good psychologist. I dedicate this to my brother Kurt for helping me and inspiring me throughout my life- from the sandbox, to Dan’s gym, to 22nd Street, to comparing notes on our careers. I dedicate this to my mother, who has always inspired my love of learning, and my career focus of helping children succeed in school. Lastly, I dedicate this to my wife Christine. We could never have done this without you. I thank you and I love you.
ACKNOWLEDGEMENTS

First and foremost I would like to thank the chair of my dissertation, Dr. Jeff Miller. He has mentored me as an advisor, leader of Directed Research, and national advocate for the field of School Psychology. Mostly, though, through the dissertation process, he helped me define and focus my research interest, and always knew when to encourage me and when to push me. It seemed like every time I felt overwhelmed, with pressures from work or family, and “spinning my wheels,” he would end an email with “you can do it.” Thank you also to my other committee members, Dr. Tammy Hughes and Dr. Ara Schmitt for their guidance and help with my final completion of the long dissertation process. Next I would like to acknowledge my research group, primarily Jess Blasik and Janice Decker, who helped design and implement the study, beginning three years ago. I would like to thank the administration at Hempfield (Randall Sarnelli and Ray Burk) and all the Hempfield teachers whose willing participation enabled a smooth process of research in the schools. Also, I would like acknowledge all of the School Psychology faculty at Duquesne. I am thankful that I was able to improve my skills and knowledge as a School Psychologist, and I feel blessed to have learned from a group who is so committed to research and practice in my chosen profession.
# TABLE OF CONTENTS

Abstract ........................................................................................................................................... iv
Dedication ........................................................................................................................................ v
Acknowledgement ........................................................................................................................... vi
List of Tables ..................................................................................................................................... x
List of Figures .................................................................................................................................... xi

## Chapter I: Introduction .............................................................................................................. 1

**Significance of the Problem** ........................................................................................................ 2

  - Purpose of Behavior Screening Methods .................................................................................. 3
  - Weaknesses of Behavior Screening Methods ............................................................................. 3
  - The Impact of Executive Function Deficits on Behavioral Screening ...................................... 4

**Theoretical Basis for the Study** .................................................................................................. 4
  - Baddeley’s Central Executive ...................................................................................................... 5
  - Executive Function Development .............................................................................................. 6
  - Measurement of Executive Function: Accuracy Versus Reaction Time ................................... 8
  - Response to Intervention ........................................................................................................... 8

**Problem Statement** .................................................................................................................... 9

**Research Questions and Hypotheses** ........................................................................................ 9

## Chapter II: Literature Review ..................................................................................................... 11

**Historical Background** ............................................................................................................. 11
  - Behavioral Screening in Schools ............................................................................................... 11
  - Executive Function Deficits ...................................................................................................... 12

**Theory of Executive Functions** ................................................................................................ 13
  - Working Memory ....................................................................................................................... 13
  - Baddeley’s Model ..................................................................................................................... 14
  - Central Executive/Executive Functions ..................................................................................... 15

**Development of Executive Functions** ....................................................................................... 16
  - General Cognitive Development .............................................................................................. 16
  - Central Executive- Structure and Development ......................................................................... 17

**Universal Screening for Behavior Problems in Schools** ............................................................ 19
  - Benefit of School-Based Screening Data .................................................................................. 19
Advantages of Preventative Strategies in Schools .......................................................... 20
Competing Models of Behavior Screening ........................................................................ 21
Link Between Executive Function Deficits and Mental Health Disorders ....................... 22
Review of Research of Executive Function Assessment .................................................. 29
Quantifying Executive Function Deficits: Accuracy vs Speed ........................................ 33
Studies Involving Analysis of Timed Measures ............................................................... 34
The Most Critical Executive Function - Inhibition .......................................................... 35
Response vs. Behavioral Inhibition ..................................................................................... 36
Conclusion ......................................................................................................................... 41

Chapter III: Methods ........................................................................................................ 43
Participants ......................................................................................................................... 43
Power Analysis ................................................................................................................... 44
Measures ............................................................................................................................. 45
  Central Executive/Executive Functions ........................................................................... 45
Behavioral Measure ........................................................................................................... 47
Research Design ............................................................................................................... 49
Procedures ......................................................................................................................... 49
Data Analysis ..................................................................................................................... 51
Research Questions and Hypotheses ................................................................................ 53

Chapter IV: Results ........................................................................................................... 55
Preliminary Analysis of Data ............................................................................................... 59
  Multicollinearity ............................................................................................................... 61
Research Question One Results ....................................................................................... 61
Research Question Two Results ....................................................................................... 73

Chapter V: Discussion ....................................................................................................... 82
Summary of Results ............................................................................................................ 82
Conclusions ......................................................................................................................... 84
  Implications ....................................................................................................................... 85
Limitations .......................................................................................................................... 86
Future Research .................................................................................................................. 87
References ........................................................................................................................... 89
LIST OF TABLES

Table 1. Teacher Ratings Completed in Each School……………………………………..48
Table 2. Participants in Each School and Each Classroom…………………………….55
Table 3. Means and Standard Deviations of Executive Function Measures- Accuracy…56
Table 4. Means and Standard Deviations of Executive Function Measures- Reaction
Time…………………………………………………………………………………………57
Table 5. Frequencies of Screener T-Scores……………………………………………..58
Table 6. Correlations Among Executive Function Measures…………………………60
Table 7. Correlations Between Executive Function Measures and BASC-2 Screener…60
Table 8. Effect of Stroop-RT and Stroop-A on BASC-2 Screener……………………65
Table 9. Effect of Category Switch-RT and Category Switch-A on BASC-2 Screener..69
Table 10. Effect of Letter-Memory-RT and Letter-Memory on BASC-2 Screener…..73
Table 11. Effect of all Reaction Times (Stroop-RT, Category Switch-RT and Letter-
Memory-RT) on BASC-2 Screener…………………………………………………..77
Table 12. Effect of All Accuracy (Stroop-A, Category Switch-A, and Letter-Memory-A)
on BASC-2 Screener……………………………………………………………………81
LIST OF FIGURES

Figure 1. Normal Distribution of Standardized Residuals for BASC-2 Screener with Stroop-Reaction Time and Stroop-Accuracy as IVs.................................................................62

Figure 2. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.................................................................................................63

Figure 3. Scatterplot of Residuals around Regression Line for BASC-2 Screener Satisfying Heteroscedasticity Assumption.......................................................................................64

Figure 4. Normal Distribution of Standardized Residuals for BASC-2 Screener with Category Switch-Reaction Time and Category Switch as IVs..............................................66

Figure 5. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.................................................................................................67

Figure 6. Scatterplot of Residuals around Regression Line for BASC-2 Screener Satisfying Heteroscedasticity..........................................................................................68

Figure 7. Normal Distribution of Standardized Residuals for BASC-2 Screener with Letter-Memory-Reaction Time and Letter-Memory-Accuracy as IVs..........................70

Figure 8. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.................................................................................................71

Figure 9. Scatterplot of Residuals around Regression Line for BASC-2 Screener Satisfying Heteroscedasticity Assumption...........................................................72

Figure 10. Normal Distribution of Standardized Residuals for BASC-2 Screener with Stroop-Reaction Time, Category Switch-Reaction Time and Letter-Memory-Reaction Time as IVs.................................................................74
Figure 11. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption………………………………………………………………..75

Figure 12. Scatterplot of Residuals around regression Line for BASC-2 Screener Satisfying Homoscedasticity Assumption…………………………………………..76

Figure 13. Normal Distribution of Standardized Residuals for BASC-2 Screener with Stroop-Accuracy, Category Switch-Accuracy, and Letter-Memory-Accuracy as IVs……………………………………………………………………………78

Figure 14. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption……………………………………………………79

Figure 15. Scatterplot of Residuals around Regression Line for BASC-2 Screener Satisfying Homoscedasticity Assumption…………………………………………80
Chapter I

Introduction

Executive function deficits in elementary age students have significant implications for behavioral, social and emotional development in the school setting. Executive functions are critical to a child’s ability to perform cognitive tasks such as problem solving, controlled attention, decision making, and mental manipulation both in school and in everyday activities (Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Mantyla, Carelli, & Forman, 2007). Additionally, school-age childrens’ ability to utilize executive functions has a significant impact on their social behavior (Best, Miller, & Jones, 2007). Presently, school systems are developing and implementing new ways to universally screen students in order to identify those who may need behavioral supports (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). However, it is not well understood how much another weakness, such as executive function may contribute a significant amount of variance at the first level of universal behavior screening.

Executive functions have been shown to relate to a variety of behavior problems (Nigg et al., 2006). In addition, executive functions correlate with learning and academic achievement as they are believed to be related to attention, attention deficits found in children with ADHD (Barkley, 1997), and learning across multiple environments (Gevins & Smith, 2000). In the school setting, it is important to understand both cognitive and behavioral weaknesses because either of these can produce students “at-risk” of academic failure.

Recent research is beginning to show that there is a link between problem behavior and deficits in cognitive processing (Young, Friedman, Miyake, Willcutt,
Corely, Haberstick, & Hewitt, 2009; Friedman & Miyake, 2004). Young and colleagues (2009) studied three different types of executive functions: inhibition, updating, and shifting among 12-year-olds. They found a significant correlation between all three executive functions, and measures of observable behavior problems. Each correlation (inhibition, -.47, updating, -.27, and shifting, -.20) showed a significant negative result, suggesting that executive functioning is negatively associated with behavioral problems.

Although literature suggests that there is a relationship between behavior problems and executive functions, the current study aims to provide a more detailed understanding of relationship between the fractionated components of executive functions and behavior ratings obtained from school-wide behavior screening. This may in turn provide new insight into which processes may increase a child’s risk for developing some type of mental health disorders later in adolescence or early adulthood. In particular, studies have pointed toward inhibition as being more closely related to some psychopathologies such as substance abuse (Nigg et al., 2006), attentional problems (Friedman et al., 2007), and overall externalizing behaviors (Young et al., 2009).

**Significance of the Problem**

Universal screening in schools is a growing practice (Severson et al., 2007). Because this practice is becoming more prevalent and valued, it is important to analyze other factors that may influence its variability. This study aims to build upon current practices of behavioral screening in public schools, by investigating whether deficits in executive functions are related to existing behavioral screening. With this improved understanding of universal behavior screening, school professionals, such as counselors,
teachers, principals, or school psychologists may be more effective at interpreting or understanding results from the first level of behavior screening.

**Purpose of Behavior Screening Methods**

School-wide behavior supports have become an important component of typical educational practice in order to improve school safety, prevent violence and disruptive behavior, and increase academic success (Walker, Golly, McLane, & Kimmich, 2005). As part of school-wide supports, behavior screenings typically involve various types of behavioral measures, including observations, teacher ratings, and record reviews of disciplinary referrals (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007).

**Weaknesses of Behavior Screening Methods**

Although these behavioral measures are certainly important they are limited to recording observable behavior. Although decision making and controlled attention are affected by deficits in executive functions, a student may not be recognized because the problem may not manifest as external or observable, as it is unclear if current screening methods directly assess covert cognitive processes. Severson et al. (2007) further describe the need to screen not only for behavioral factors that may be considered “precipitating” (p. 216) but also factors that could be “predisposing” (p. 216) such as executive functions deficits. A precipitating problem may be easily observed, such as a peer conflict immediately prior to a physical fight. However, a predisposing factor would
be a child’s history or individual differences not immediately visible to teachers or educators (Severson et al.).

**The Impact of Executive Function Deficits on Behavioral Screening**

A better understanding of the relationship between behavior and executive function assessment may provide valuable information to treatment teams for the development of academic and behavioral goals. The purpose is to help explain if a behavior screening may miss students who are not yet referred for services due to some level of executive function variance. Such students, who could be considered “behaviorally at-risk” for developing a mental health disorder, may not be identified unless their core cognitive functioning deficits are identified prior to symptom expression. This study examines if executive functioning may account for variance in existing behavior screening procedures. If this is better understood, this may enable school teams to identify students with manifest executive function and behavioral deficits and, by extension, provide preventive services, to reduce the chance that such students would either require specialized programming or develop significant mental health problems (Friedman et al., 2007).

**Theoretical Basis for the Study**

This study examines the relationship between teacher ratings of at-risk behaviors and executive functions in a general education fourth grade population. Most student screening methods (Walker et al., 2005; Severson et al., 2007) target observable patterns of student behavior and emotional functioning that may identify students as “at-risk” for developing significant problems such as depression, anxiety, violence and aggression, or substance abuse problems (Richardson, Caldarella, Young, Young, & Young, 2009;
Severson et al.). This study aims to build upon current practices of school-wide behavior screening by explaining if executive function deficits may account for some variance in a behavioral screener for Tier 1 Response to Intervention (RTI) assessment in a sample of 4th grade public school children.

The dependent variable is a behavioral screening version of the Behavior Assessment System for Children- Second Edition (BASC-2; Kamphaus & Reynolds, 2007) that is designed to measure behavioral competencies as well as problems. It is anticipated that scores on the behavioral screening instrument will be related to several different measures of executive function (Best et al., 2009), and therefore explain some level of variance in the first level screener.

The independent variables in the study are three different types of executive function— inhibition, updating and shifting. Although the executive functions are part of complex cognitive circuits, this study will focus on the three fractionated executive functions identified in Baddeley’s (2002) central executive component of his working memory model.

**Baddeley’s Central Executive**

Different types of executive skills have been described in the literature, including planning, organizing, and different aspects of attention (Miyake et al., 2000; Young et al., 2009), nevertheless, this study will focus on Baddeley’s (2000) working memory model. It includes the central executive, in which several types of executive functions perform cognitive tasks. Using the framework of Baddeley’s “central executive,” Miyake, and colleagues were able to identify 3 distinct but interrelated executive functions— inhibition, shifting, and updating. Inhibition is the ability to manage, or inhibit, prepotent
responses. The classic example is the Stroop test (Stroop, 1935), in which the individual is asked to read a series of color words (red, blue, green, etc.). Then the task alternates, and the subject is asked to identify the color of the print of a word. In the task, the words are different colors, but the text is still the color words (red, blue, green, etc.). The “prepotent” aspect is that the subject is likely to continue reading the text, but must inhibit this to identify the color of the text. The next executive function skill identified by Miyake et al. is shifting. This task involves the ability to shift the “rule” that is required during a task. For example, the subject may be asked to view a list of animals and objects, and identify if the item is small or large, then the subject is asked to “shift rules” and identify if the item is “living” or “not living.” The last executive function, updating, is the ability to update information into working memory. An example of this task may be some type of list learning task in which a subject is asked to recall increasing number of words heard aurally.

As stated previously, executive functions impact problem solving, controlled attention, and decision making in academic settings and in everyday activities (Lehto et al., 2003; Mantyla et al., 2007). Although behavior screening may identify observable behaviors, a student may also have an associated executive function deficit that could go unnoticed. As a result, it is important to understand the relationship between strengths and weaknesses in this cognitive process, and the process of monitoring students for behaviors that are observable.

**Executive Function Development**

Developmental changes in frontal lobe processing and executive functions have been well documented (Alloway, Gathercole, Willis & Adams, 2004; Thorell, et al.,
From the preschool years to adulthood, changes occur in a variety of cognitive skills within Baddeley’s model of the central executive (Cowan et al., 2006). Specific to the structure of the central executive, Wiebe et al., (2008) found that due to developmental factors, each component of the central executive may not be distinguishable, or fractionated during the preschool ages (2-6). Rather, the system of executive skills may be more of a unitary general model during this young age, then become more fractionated as the child grows older (Friedman et al., 2004). Within the beginning school-age range (ages 6-10), the impact of cognitive development has increased implications for academic achievement, as children are now expected to demonstrate effective executive functions for the purpose of goal-directed behavior, and in order to perform mental tasks during academic work in the school setting (Alloway et al., 2003; Friedman et al., 2004). Within Piaget’s theory of cognitive development, most students acquire the capacity for concrete operations, which now enables them to begin to think logically, and utilize reasoning skills.

Most research on the topic of prevention and early intervention indicates that elementary age students should be universally screened for at-risk behavior at several points during each school year, and for several years during elementary school (Severson et al., 2007; Walker et al., 2005). This need to screen at multiple developmental stages aligns with executive function literature showing continued growth from preschool through adolescence (Lehto et al., 2003; Wiebe et al., 2008). Best and colleagues (2009) reviewed the importance of studying executive functions in school-age students. Although some of these skills may continue to develop through adolescence, some executive function development, such as inhibition, fully matures around age 8, and
while in the school-age range, the different types of executive functions become more distinguishable, rather than producing the single construct of executive function more visible in early childhood (age 2-6; Wiebe et al.). In addition, Best et al. conclude that school-age children can be assessed with a wider variety of cognitive tasks that may involve reading skills or more complicated directions. Given these findings, the current study will focus on fourth grade students between the ages of 9-10 years.

**Measurement of Executive Function: Accuracy Versus Reaction Time**

Essential to any type of assessment is how the student’s performance on a test is quantifiably measured. Among the many pieces of information that may be gathered on a psychological test, two very common measures are whether an item is correctly completed (accuracy) and the amount of time between a stimulus being presented and the subject beginning their response (reaction time; Sattler, 2001). Neuropsychological testing places more emphasis on the role of reaction time, as deficits in this process are often indications of a wide variety of weaknesses, including inattention, poor motor control, or sensory issues (Lezak, Howieson, Loring, Hannay, and Fischer, 2004). In terms of executive function assessment, accuracy (A) and reaction time (RT) are frequently used as dependent measures of performance (Friedman et al., 2007; Miyake et al., 2000). Because of the increased understanding of reaction time, some researchers are beginning to view RT variability as the core deficit in many neuropsychological tests (Epstein, Langberg, Rosen, Graham, Narad, & Antonini, et al., 2011).

**Response to Intervention**

Response to Intervention (RTI) is framework to guide educational practice that involves both academic and behavioral goals and strategies (Nunn, Jantz, & Butikofer,
2009). It is sometimes compared to a medical model in that if through monitoring it is learned that a patient is not improving after the administration of a particular dosage of medication, the amount of medication administered could be altered (Severson et al., 2007). The most common model of RTI uses three levels, or Tiers of support, with the first level, Tier I, being validated educational strategies utilized with the whole school population. In order to move to the next higher level of intervention, Tier II, students must be screened for weakness. The process of assessing the whole school population, known as universal screening, utilizes brief assessment data that can be easily gathered (Marchant, Anderson, Caldarella, Young, & Young, 2005). This study will examine the relationship between Tier I universal behavior screening data and measures of executive functioning.

**Problem Statement**

Studies have previously shown links between executive functions and behavior problems, but those studies have involved samples ranging in age from 17 to 24 years of age (Young et al., 2009; Friedman & Miyake, 2004). This study will extend the literature by examining if universal behavior screening in school-age children is related to three types of executive functions. This would allow for a clearer understanding of students’ underlying cognitive process that may be associated with problem behavior. The purpose of this study is to determine if executive function is related to teacher-completed behavioral screener. This may explain some weaknesses, or variance, that may exist in a first level universal behavioral screening measure.

**Research Questions and Hypotheses**

Research Question 1:
What is the relationship between the executive function measures and the behavioral measure?

Hypothesis 1: Reaction time of the executive function inhibition will account for more variance on the BASC-2 universal screener for behavior than accuracy.

Hypothesis 2: Reaction time of the executive function shifting will account for more variance on the BASC-2 universal screener for behavior than accuracy.

Hypothesis 3: Reaction time of the executive function updating will account for more variance on the BASC-2 universal screener for behavior than accuracy.

Research Question 2: Using measures of executive function (inhibition, shifting, and updating), is there a difference between reaction time and accuracy in the amount of variance contributed on the BASC-2 universal screener for behavior?

Hypothesis 4: The reaction time for inhibition, shifting and updating account for a significant amount of variance on the BASC-2 universal screener for behavior.

Hypothesis 5: The accuracy for inhibition, shifting and updating do not account for a significant amount of variance on the BASC-2 universal screener for behavior.
Chapter II

Literature Review

Historical Background

Behavioral Screening in Schools
The use of school-wide behavior support has become an important component of typical educational practice as policy makers and school officials attempt to improve school safety, and prevent violence and disruptive behavior (Walker, Golly, Mcclane, & Kimmich, 2005). In addition, school climate and school-wide supports are more frequently used methods to prevent significant mental health problems for children and adolescents. In order to target specific students the first step of the process is to screen for problems. This screening typically involves behavioral observation measures, such as teacher rankings of elementary students, student self-report questionnaires for older elementary aged students, and review of office disciplinary students for adolescent aged students (Severson et al., 2007). Although most models of school-wide plans begin with multiple levels of behavioral observations, consideration should be given to the strong relationship between behavior and higher level cognitive processing (Young et al., 2009).

These types of cognitive skills allow a child to focus on an intended thought or behavior, and sort through other information, and complete a goal directed task. A simple example of this would be the traditional Simon Says game in which a child has to remember directions, and only follow a directive when preceded by saying ‘simon says’ (Garon, Bryson & Smith, 2008). This type of process can be identified in many ways, but is most often classified as working memory (WM), or executive functioning (EF; Baddeley, 2002; Diamond, Carlson, & Beck, 2005).

**Executive Function Deficits**

Many studies have shown a link between cognitive control processes and social and emotional deficits. Some have focused on depression (Joormann, & Gotlib, 2008;
Harvey, Le Bastard, Pochon, Levy, Allilaire, Dubois, et al., (2004), whereas others have focused on anxiety. Several studies involving anxiety were narrowed down to specific subtypes, such as Posttraumatic Stress Disorder (PTSD; Leskin, & White, (2007). Still other studies have focused on Obsessive Compulsive Disorder (OCD; Bucci et al., 2007). Each of these studies utilized children and adolescents already diagnosed with some type of mental health disorder. However, the role of primary intervention and preventative programs is to intervene before more serious disorders are manifested (Walker, Golly, Mcclane, & Kimmich, 2005). Recent research has shown a link between behavioral indicators, and measurable deficits in cognitive processing (Young, Friedman, Miyake, Willcutt, Corley, Haberstick, & Hewitt, 2009). This presents an opportunity to explore whether some type of executive function may impact the screening components of an effective school-wide behavior support plan. To first understand the connections between impulse control, effective decision making and social and emotional functioning, the components of working memory must first be examined, as well as the aspects of executive functioning that are most salient to social and emotional functioning.

**Theory of Executive Functions**

**Working Memory**

Higher order cognitive processes that affect and guide behavior involve the ability to plan, anticipate situations, modulate one’s own behavior, and engage in two simultaneous tasks, such as performing a task while also processing visual and auditory information (Lehto, Juujarvi, Kooistra, Pulkkinen, 2003). Although different terms exist for similar cognitive tasks, all of these skills can be associated with working memory (Baddeley, 2002). Traditional understanding of memory usually involves long term
storage and retrieval. Working memory (WM) is the ability to mentally hold onto small amounts of information for a brief period of time and “work” with it, and complete some other task. Different theories or models of WM have been presented, and are usually considered either unitary or multi-modal.

**Baddeley’s Model**

Often cited in the literature, Baddeley’s model has gone through several adjustments over the years (Baddley, 2002), but maintains its multi-model structure. Some of the first descriptions of the model involve the central executive, and two “slave systems” or sub-systems known as the visual-spatial sketchpad and the phonological loop. The central executive was initially viewed as a simple control mechanism, or homunculus (described by Baddeley as a little man deciding how to utilize the slave systems, Baddeley, 2002). This component had adopted some of the features of Norman and Salice’s supervisory attentional system (Baddeley, 2002). The central executive component was later studied and described as having 3 components (focused attention, divided attention, and attention switching; Baddeley, 2002).

Baddeley’s two “slave systems” are different mechanisms for processing visual and spatial information (visual-spatial sketchpad) or auditory information (phonological loop). The phonological loop involves two components, the phonological store and the articulatory control system. The phonological store involves the brief storage of auditory information, such as single numbers or letters. The articulatory rehearsal system refers to process of sub vocalization of words in order to recall (Baddley, 2002). The visual-spatial sketchpad involves the ability to temporarily maintain and manipulate visuospatial information. At a later time, the episodic buffer was separated out as a fourth component,
which integrates information, and acts as an interface between the sub-systems (Altgassen, Phillips, Kopp, & Kliegel, (2007).

Central Executive/Executive Functions

Using Baddeley’s (2002) model as a starting point, the central executive is a crucial component in that it guides the processing of all stimuli, whether it is visual or auditory. Over the last several decades, research has begun to show the crucial roles and the key components of this “central executive,” also referred to as “executive functions.” While these two terms are sometimes interchanged with other processes, such as working memory, or the attentional system, these are all closely related to the functioning of the frontal lobe of the brain, and the prefrontal cortex. The process involves one’s ability to focus on a task, shift mental focus from one task to another, and engage in new learning, while ignoring, or inhibiting other visual or auditory stimuli (Miyake et al., 2000).

Executive functions typically involve the activation of the prefrontal cortex, commonly viewed as an active control mechanism of the brain, key to the integration of many other areas of the brain, as a child learns. It can refer to the process of higher order and metacognitive control processes needed for goal oriented behavior (Luby, 2006). Specific to younger children (preschool), the study of executive functions in children was typically neglected, as it was assumed that these skills developed later, due to how a young child presented as “dysexecutive” (Luby, 2006). Other researchers have put forth different models of EF. One of the models was developed as part of Baddeley’s (2002) multi-component working memory model. Other models, such the Supervisory Attentional System (SAS), place more emphasis on the role of attention processing as the overarching mechanism (Shalice & Burgess, 1991). While different types of executive
functions have been studied and analyzed over the years, researchers have recently been able to identify the primary components of the central executive. Miyake et al. (2000) were able to identify separate but interrelated components of the central executive.

Development of Executive Functions

General Cognitive Development

One of the most well-known theories of cognitive development is that of Jean Piaget (Ormrod, 1995). His developmental theory names multiple stages from childhood to adulthood. The first stage, from birth to about age 2, is the sensorimotor stage. This stage is characterized by behavior-based and perception-based schemas, rather than by involvement of the thinking process. The next stage, the preoperational stage, begins at about age 2 and continues to about age 6 or 7. At this time in life, a child begins to use language and the true “thinking” process, as he begins to develop internal schemas in the absence of the person or object. However, at this stage, the child’s thinking is sometimes illogical, at times attributing psychological phenomena to physical reality (such as monsters or the boogeyman). The third stage, concrete operations, begins around 6 or 7 and continues to about 11 or 12 years or age. At this stage, children begin to think logically, and are able to engage in conservation. While they are able to engage in logical operations, they are not yet able to apply logical operations with abstract or hypothetical situations. The last stage, formal operations, begins around 11 or 12, and continues for several years. The child begins to reason with abstract and hypothetical information or situations (Ormrod, 1995). While Piaget’s theory of cognitive development is widely accepted, some flaws are obvious, as evidenced by continuing changes in childhood cognitive skills (Best et al., 2009). This study will use a sample of children ages 9 and 10.
Because these ages fall within a broad range (6-11), Piaget’s theory does not differentiate between the cognitive capacities within this stage. The proposed study will contribute to the understanding of cognitive development within the stage of concrete operations.

**Central Executive- Structure and Development**

Recent research (Miyake, Friedman, Emerson, Witzki, & Howarter, 2000; Friedman & Miyake, 2004) has begun to show a clearer picture of this process of the “central executive” is organized. More specifically, these studies have shown the core process of EF as a fractionated process, with three key components, as opposed to four, as described by Baddeley.

Confirmatory factor analysis (Miyake et al., 2000) has shown three distinct, but interrelated components of the central executive: updating, shifting, and inhibition. Updating is involved in the encoding of new information; shifting is involved with the ability to shift mental focus from one task to another; and inhibition is the ability to inhibit other motor or auditory processes, in order to complete the original task.

Multiple studies (Miyake et al., 2000; Friedman et al., 2004; Friedman, et al., 2007; Young et al., 2008) have identified and supported these very clear, diverse components of the central executive. However, one of the weaknesses of these studies is that they do not consider the developmental sequence of executive function components during childhood. For example, Miyake et al. utilized a sampling of college students-who though considered by Piaget to be still in the developing stage, nevertheless represent a vastly different population from elementary school children. Subsequent studies have begun to examine the impact of developmental trajectory of executive functions during childhood.
Wiebe, Espy, and Charak (2008) conducted a study to attempt to establish the structure of executive control with children under the age of 6. In reviewing the literature, the authors noted that a variety of executive control measures for younger children have been developed. However, existing literature does not consistently describe the same structure and components of executive control. Wiebe et al. found little research had been conducted with preschool populations, but noted that Miyake et al. (2000) had been able to identify the fractionated components of the central executive, using Confirmatory Factor Analysis (CFA). As a result, Wiebe et al. used an approach similar to the Miyake group. They considered measuring a construct similar to shifting, but did not do so because they were not able to locate an assessment measure suitable to preschool aged children. The sample included 243 preschool children who were recruited through birth announcements, from local preschools, through the local health department, and by word of mouth. Children ranged in age from 2 years 4 months to 6 years. The sample was composed of 171 Caucasian, 43 African American, 9 Asian American, 1 Native American, 4 Hispanic, and 14 multiracial children; 1 child’s race was not reported.

Several assessments were specifically used to measure inhibition: Delayed Response task, in which treats were hidden in a pseudorandom order in two locations in the child’s view; Whisper task in which children whispered the names of a series of pictures of familiar and unfamiliar characters. Two subtests of the NEPSY (Korkman, Kirk, & Kemp, 1998) were administered, the statue and the visual attention task. Additional measures of inhibition included shape school, Tower of Hanoi, and a continuous performance task.
Factor analysis was conducted to determine the structure of preschool children’s executive control, and up to six increasingly restrictive models were tested. Results indicated that the simplest model, a single Executive Control factor was supported over other multifactor models. In discussing the findings, the authors note this finding contrasts with fractionated component models. They first considered that because of the more limited skills of the young children, the assessments may not have been diverse enough to measure different executive skills. However, the tasks did seem to be variable. The authors suggest that the single factor model may be more specific to the preschool aged child.

Universal Screening for Behavior Problems in Schools

Benefit of School-Based Screening Data

The practice of universal screening for behavior problems indicates that all students are assessed to determine if they need some type of behavioral support such as social skills training or counseling in the schools (Walker, Cheney, Stage & Blum, 2005). Teacher ratings of behavior have been shown to be a consistently reliable measure of behavior in the school setting (Kamphaus, Thorpe, Winsor, Kronche, Doudy & VanDeventer, 2007). In addition, studies have also shown that child and parent ratings do not substantially improve results of a behavior screening process (Kamphaus et al.). This would indicate that the use of teacher ratings of behavior is an effective means to screen for behavior problems and ultimately intervene before more significant mental health problems develop (Kamphaus et al.; Flanagan, Bierman & Kam, 2003). Flanagan et al. found that a systematic screening of behavior using teacher ratings with a sample of first
grade students effectively predicted behavioral and academic outcomes for these students two years later.

**Advantages of Preventative Strategies in Schools**

Many national and statewide plans have begun to review the efficacy of early intervening to prevent behavioral and mental health problems. Walker, Golly, McLane, and Kimmich (2005) reviewed the effectiveness of a program implemented throughout the state of Oregon. The program, called *First Step*, is a school-wide screening and evidenced-based intervention program with multiple components, including parent education, social skills training and collaboration with the classroom teacher. The study reported by Walker et al. focused on students in kindergarten through second grade who were demonstrating some level of externalizing behavior. To demonstrate the effectiveness of the program, four dependent measures were used - two parts of the Systematic Screening for Behavior Disorders (SSBD) and the aggression subscale on the Child Behavior Checklist. Pre-test and post-test comparison after the intervention produced significant change with effect sizes ranging from .84 to 1.31. This would indicate that this school-wide intervention program was effective, as the students targeted for intervention demonstrated improved behavior.

Severson, Walker, Hope-Doolittle, Kratochwill and Gresham (2007) reviewed some of the relevant issues related to early screening and early behavior intervention practices in school districts and how this trend is increasing throughout the country. The authors stated three primary reasons for the trend. The first involves the strong reactions to the multiple schools shootings on the 1990’s, which led to pressures from parents, legislators, and the public to dramatically increase school safety, and to identify students
who might be at-risk for violent behavior. The second influence is the social pressure to find interventions that are proven to work (evidenced based) to prevent behavior and mental health problems in the schools. The third and final influence is the movement towards multi-tiered models aimed at preventing academic and mental health problems. Key to this process is the systematic and school-wide screening for students considered at-risk.

The authors conducted a survey and analysis of commonly used screening tools, in order to identify students who may have behavioral problems. After having reviewed these tools they concluded that best practice suggest three general models of screening, (1) multiple gating procedures, (2) teacher evaluation and rating(s) of all students in the classroom, and (3) teacher nomination of problem students followed by Likert rating(s) of their behavioral characteristics and social skills.

**Competing Models of Behavior Screening**

Walker, Cheney, Sage, and Blum (2005) also reviewed several models for school-wide behavior screening. They utilized several different methods (standardized rating scales, unstandardized scales, and review of discipline referrals) of screening students in several elementary schools. Each method identified students with behavior problems at different levels (primary, secondary and tertiary), with decreasing number of students at each level and the smallest number of students having the greatest need for intervention (tertiary). Some of the students had already been identified as needing intervention, and then were rescreened with a different observational method, and still found to be in need of intervention. They also found that kids with an “internalizing” profile were difficult to clearly assess and identify with observational measures, or review of discipline referrals.
The authors concluded that the results confirmed the need for ongoing screening, and that
the process is never “done.” A contrary view of the results, is that simply conducting
observational ratings and reviewing observational data is an insufficient method to screen
for social and emotional deficits. Walker et al. clearly stated that internalizing problems
are often overlooked with observational measures. Another flaw with this somewhat
complex method is an issue of cost-benefit and the resources that must be devoted to
screen and identify the right students. Some researchers (Kamphaus et al., 2007;
Flanagan et al., 2003) point out that multiple level and multi-gate screeners are effective
in identifying students in need of intervention. However, given the complexity, teacher
ratings (in the absence of direct observations or multiple “gates”) may be as effective and
more appropriate. Given the clear weaknesses of the current practice of behavior
screening (can be taxing in terms of resources, and the process can overlook an
internalizing student), the process may benefit from an additional component to the
process.

**Link Between Executive Function Deficits and Mental Health Disorders**

One group of mental health disorders that has been studied, internalizing
disorders, has been consistently linked to problems with executive functions. Some
articles and studies have focused specifically on depression (Joormann, & Gotlib, 2008;
Harvey, Le Bastard, Pochon, Levy, Allilaire, Dubois, et al., 2004), whereas others have
focused on anxiety. Several studies involving anxiety were narrowed down to specific
subtypes, such as Posttraumatic Stress Disorder (PTSD; Leskin, & White, (2007). Still
other studies have focused on Obsessive Compulsive Disorder (OCD; Bucci et al., 2007).
In assessing this link between executive functions and social and emotional problems, most studies seem to focus on broadly stated executive function/executive control, with a few focusing on working memory, both verbal and visual spatial. A variety of measures have been used to assess the deficits in working memory, ranging from a simple modified Sternberg task involving visually presented words in different colors (Joorman & Gotlib, 2008), or a test of facial memory, to more complex experiments involving the Wisconsin Cart Sorting Test (Bucci et al., 2007), or Trail Making Test (Emerson, Mollet, & Harrison, 2005).

Bucci et al. (2007), found that Obsessive-Compulsive Disorder does have an effect on executive functioning, particularly with visuospatial abilities. The authors concluded that this supported previous findings that this internalizing disorder (OCD) involves a dysfunction of frontostriatal circuits in the right hemisphere.

Another study of internalizing disorders (Joorman & Gotlib, 2008) focused on the effects of internalizing disorders on executive functioning, particularly in the area of depression and working memory. The authors’ research hypothesis was that the inability to remove negative thoughts from working memory was related to the tendency to ruminate on negative moods and events. They used a sample population of adults, one group recruited from outpatient clinics, and the other group recruited from the community to serve as a control group. Subjects ranged in age from 18 to 60. Subjects were diagnosed using the Structured Clinical Interview based upon the DSM-IV, and were included if they met criteria of Major Depressive Disorder. The dependent variable was the subjects’ performance on a computer based task. They were presented with several lists of words in red or blue, either negative or positive valence words. The words
then reappeared on screen, written in black, and surrounded by a box, that switched from each color. They were asked to indicate whether the word was from the associated list. Success was measured by the speed and accuracy of response. They attempted to show that depressed individuals experience an interference in the working memory process from irrelevant negative material. Results confirmed their hypothesis. In addition, findings also showed that not only did negative material interfere with the updating of working memory contents, but also they found that this interference pattern did not occur for positive material. This study by Joorman and Gotlib (2008) provides a link between EF deficits and emotional problems (depression). However, their subjects were all adults, and so the results may or may not generalize to the child and adolescent population.

Nigg et al. (2006) studied the link between executive functions and adolescent drug and alcohol use. Specifically, they were interested in the ability of poor response inhibition to predict drug and alcohol use. The authors’ research hypothesis was that weaknesses in response inhibition was predictive of drug and alcohol use in a high-risk sample of children, while taking into account IQ, parental lifetime alcoholism, and antisocial personality disorder. They were mainly interested whether poor response inhibition was a unique factor in the risk pathway. They used a sample population from an on-going “multiwave” community sample of high risk families, with children ranging in age from 12 to 17. The sample included 362 boys and 136 girls (498 total), from 275 different families. Subjects were assessed with a variety of measures. To assess ADHD, conduct disorder, and alcohol abuse, a variety of clinical scales and DSM-IV criteria were used. To assess IQ, participants were assessed at different times, or waves, by either the Stanford-Binet LM, or the WISC-R. In terms of executive functions, they were
assessed by one measure, the Stop-Signal, specifically targeting response inhibition. This computerized task requires the subjects to quickly press one of two buttons if either an X or O appeared on the screen, then withholding (or inhibiting) their response if a tone sounded. The dependent variable was calculated using response time, that is the difference between average stop-signal delay from the average response time. Executive functions were also measured with a series of more broad measures, to assess skills such as shifting or updating as well as inhibition. These tasks included a computerized version of the Wisconsin Card Sorting Test, a version of the Stroop test, a timed verbal naming task, a symbol-digit task, and the Tower of Hanoi. Results confirmed their hypothesis that response inhibition is an important indicator of drug and alcohol abuse. It was shown that ADHD and conduct disorder were related to drug and alcohol use, and parent substance abuse and antisocial features predicted substance abuse of adolescents.

More importantly, the researchers (Nigg et al., 2006) found that the while the broad EF measures did not predict drug and alcohol use (while considering the covariate), the response inhibition measure did predict drug and alcohol use, independent of the family and other well-established risk measured as part of the study. This study by Nigg et al. provides a link between this specific EF (response inhibition) and drug and alcohol use in adolescents. However, there were several weaknesses in the study. First of all, the subjects were families, including multiple siblings, with information gathered over time. Parent alcoholism information was gathered at baseline, which fails to account for the development of problems over time, and varying effects on different ages of siblings. In addition, although the study did compare other broad measures of EF, response inhibition was not effectively assessed by the other established component measures of EF.
(shifting, updating). This study was important however, in that a large sample was used (nearly 500 children) and a variety of established EF and drug and alcohol measures were used. Its significance is in its support of the importance of EF deficits, particularly clinical and theoretical response inhibition.

Another study focusing on the internalizing disorder of depression (Dalgleish et al., 2007) found positive associations between depressed mood and autobiographical memory, with frontal lobe and executive systems playing a key role. Multiple studies were completed by these researchers, in an attempt to associate several disorders, including eating disorders and depression, in order to examine the role of executive function in the process. Subjects were 32 adults diagnosed with an eating disorder. Although they did conclude there is a distinct role of executive function (EF) deficits with autobiographical memory, it was not clear which facets of EF were most important. The authors postulated that the most likely role of EF deficits is the effect of rumitive process and task-irrelevant thoughts associated with depressed mood. Again, though this study supports a link between executive functions and depression, however, there were several weaknesses. Multiple groups were used, but all mean scores for each group ranged from 26 to 40, indicating the results may not generalize to the adolescent population. In addition, of all the “executive control” measures used, little support was provided for their selection. Several were identified as measures of “divergent thinking,” not typically identified as part of the central executive (Alternative Uses Test and Design Fluency Test). The Block Design Subtest from the Wechsler Adult Intelligence Scale – Third Edition was administered, but what skills it measured was not described. Lastly, the Porteus Mazes Test was administered as a measure of planning. From this battery, it is
clear that it would be difficult to identify what type of EF was measured, if even it was assessed at all.

With respect to anxiety, Leskin and White (2007) found a link between PTSD symptoms and deficits in EF. Participants were grouped into varying levels of trauma and were assessed with several measures such as the Trail Making Test, Comprehensive Trail Making Test, and the Attentional Network Task (ANT). The ANT reflected deficits, whereas the other measures did not. This appeared to be consistent with previous research with similar samples (this study involved college students). This and previous research shows that the link between PTSD symptoms and EF deficits relates to functional and structural differences of the anterior cingulated and prefrontal cortex in PTSD patients. The authors describe their findings as a “subtle but specific” effect on EF by this internalizing disorder. Again, two weaknesses of this study were that the measures used have not been consistently cited in the literature as executive function measures, and the mean age of each participant group ranged from 20 to 21.

Emerson et al. (2005) looked at EF deficits with anxious and depressed boys. Thirty-eight boys, ages 9–11 enrolled in grades four and five of a suburban Chicago public school were selected to participate in the study, from a larger group of boys ($n = 65$), based on their scores on the Child Depression Inventory (CDI) and the Trait subscale of the State-Trait Anxiety Inventory for Children. Executive Functioning was measured using the Trail Making Test and the Concept Formation subtest of the Woodcock Johnson Test of Cognitive Abilities. They found that anxious depressed boys showed deficits in problem-solving tasks, sequencing and alternation tasks. Results support evidence for deficits in frontal lobe functioning. While this study did utilize school age
subjects in the study, the measures of executive functioning again have not been consistently cited in the literature.

Friedman et al. (2007) analyzed the link between executive functions and attentional problems. This study used subjects aged 7 to 14, who were part of a longitudinal study focused on whether attention problems show differential relations to multiple, separable EFs, and how developmental stability and change in attention problems relate to later EFs. They examined how attention problems, measured with the Teacher Report Form of the Child Behavior Checklist (TRF; Achenbach, 1991) at ages 7 to 14 years, related to inhibiting, updating, and shifting abilities, measured with laboratory tasks at age 17 years, and IQ at age 16. The EF tasks selected have been consistently used in the literature, and will be described in greater detail later in this paper. The researchers asked three primary questions: 1) do attention problems differentially relate to EFs; 2) how stable are the correlations between attention problems and EF from ages 7 to 14; and 3) how does development of attention relate to later EFs? The authors concluded that at each age (7 to 14), attention problems significantly predicted inhibiting and updating, and at some ages attention problems predicted shifting. Secondly, they concluded that attention problems were differentially related to the three EFs (shifting, updating, and inhibition). This study seems to provide strong evidence for the link between EF and behavior deficits.

Is there a component of the central executive that is more critical than the others? Many of the studies previously reviewed used a wide variety of assessments of frontal lobe functioning in order to determine how executive functions affect mental health disorders. However, some of the assessment measures are broad measures, such as
Wisconsin Card Sorting (Miyake et al., 2000), which makes it difficult to determine the importance of one component over the other two. In addition, it is also important to distinguish the differences between deficits in broad measures of cognitive processing, and deficits in some aspect of the central executive. From this point, it is important to begin to pinpoint which of the most frequently cited executive functions (shifting, updating or inhibition) is most critical when targeting children at-risk for social and emotional deficits.

**Review of Research of Executive Function Assessment**

In order to assess executive functions, different researchers have utilized a combination of neuropsychological tests, which seem to tap into the basic process. Each of the studies would assess the three commonly accepted components of EF, but to address the research question specific to this function, those that assessed inhibition will be reviewed in greater detail.

The study that first began to fractionate EF using confirmatory factor analysis, and analyze each component, Miyake et al. (2000), had multiple assessment measures, some of which were traditional “EF” measures, whereas, others specifically targeted the core components. The study utilized 137 undergraduate college students, who received college credit for participating in the study. Nine different computer based experimental tasks were administered, along with three other commonly used neuropsychological tests. The goal of the study was to determine if there was a connection between each of the EF components, while also demonstrating that each was a distinctly different measurable construct.
The first of the commonly used measures was the Wisconsin Card Sorting Test, a widely used, complex test, that measures more than one specific executive control task. Because part of its tasks may involve planning, it can be viewed as a broad measure of executive function, but in the context of this study, it assessed “shifting.” Random number generation was described as a measure of inhibition and updating. Finally, the Tower of Hanoi is also a broad measure of EF, but in the context of this study, it specifically assessed inhibition. Nine computer-based tests were used to measure each of the components. The three measuring shifting were plus-minus, number-letter task, and local-global task. The three that measured updating were tone monitoring task, and letter-memory task. Lastly, those that measured inhibition were stop-signal, Stroop, and antisaccade. Miyake et al. (2000) found that each of three executive functions (inhibition, updating and shifting) were related, but were distinguishable.

Lehto et al., 2003, also studied EF, and in their study described their model as assessing working memory and inhibition together. For the first part of the assessment, the children (ages 8 to 13) completed the standard pencil-and-paper version of the Trail Making Test. The next component was the Auditory Attention and Response Set (AARS) from the NEPSY-2. The AARS included two parts. Word Fluency task was a task chosen from the Finnish version of the NEPSY battery. Matching Familiar Figures (MFF). The participants also carried out the Mazes task included in the WISC–R.

Tsujimoto, Kuwajima, and Toshiyuki in 2007 conducted a study to explore the developmental fractionation of working memory and executive functions during childhood. They specifically targeted response inhibition, and employed one assessment measure, a go/no-go task. The utilized subjects ranging from age 5 to age 9. Based upon
the results of the assessment conducted that at age 5/6, the skills were more unitary in nature, and as they got older (8/9), the measures of working memory and EF became more fractionated.

Friedman et al., 2008, utilized a battery based upon Miyake et al. (2000). The three measures used to assess inhibition were the anti-saccade, the Stroop, and the stop-signal. The subjects utilized in the study were twins, aged 16-17. The purpose of the study was to identify the genetic influences of executive functions. The researchers concluded that individual differences in executive functions (shifting, updating and inhibition) are almost entirely genetic in origin. This study again utilized the same nine EF measures used by Miyake et al. (2000). Specific to inhibition, results also showed the inhibition factor had more shared variance to the common EF factor than shifting and updating.

Thorell, Lindqvist, Nutley, Bohlin, and Klingberg (2009) conducted a study with preschoolers. This study focused on intervention and developmental factors, but is important to review because it was very recent and it used a variety of assessment measures. Subjects were ages 4 to 5 years old, from 4 different preschools near Stockholm, Sweden. Four groups were formed. One was an experimental group that received training designed to improve inhibition, one experimental group trained to improve working memory, one active control group that played video games and lastly a passive control that did not engage in any activity. Each of the experimental groups and the active control group played computer games each day they attended preschool. The active control group played a commercially available computer game that was
chosen because it had minimal impact on working memory or inhibition. The experimental computer game was developed by the authors in conjunction with the company Cogmed systems.

The WM game, targeting visual-spatial working memory, required the children to view shapes on the screen and remember their location and order on the screen. The inhibition training program included two tasks (go/no go and flanker). The first task, had the children respond when a certain image was seen (like a fruit), then not respond when they saw another image (like a fish). The next task had them respond quickly when they saw one image, but not when followed by a different specified object. The last inhibition task asked had the children respond based upon the direction while ignoring another arrow on the screen.

Eight different pre- and post-test measures included the following: a Stroop-like task (to assess interference control), a go/no go task (to assess response inhibition, visual attention and response speed), the span-board task from the WASI-R-NI (to assess visual-spatial WM), a word span task (to assess verbal WM), an auditory continuous performance task to measure auditory attention, the Block Design subtest from the WPPSI-R (to assess problem solving).

Results of Thorell et al. (2009), showed no difference in post-test measures for either control group. There were significant differences after training on visual-spatial and verbal WM. Statistical differences, and moderate effect sizes were reported between control and visual-spatial WM (.89) and verbal WM (1.15). Measures of inhibition did not show significant improvement after training. The authors discussed the different effects the training had on inhibition, as opposed to WM. They proposed 4 main
explanations. First, they stated that the neuropsychological basis of WM and inhibition are different, in that the cortical area most involved with each process may differ in plasticity. Second, they suggested that the task of inhibiting a prepotent task takes only a few milliseconds, whereas holding information “in mind” takes a few seconds. Thirdly, they suggested that the measures differ in that the task difficulty can be more readily adjusted for WM tasks, than can the inhibitory tasks. Lastly, they explained that many of the children already performed well on the tasks, leaving little room for improvement, and that the segments of the training components may have been too short to have an effect.

**Quantifying Executive Function Deficits: Accuracy vs Speed**

As described by Sattler (2001), the four pillars of assessment involve norm referenced tests, interviews, observations, and informal assessment procedures. Most normed referenced tests involve “right or wrong” answers, that may be referred to as accuracy. More complex, however, is timed tasks. A timed task most likely assesses some type of information processing. Following Horn’s theory of intelligence, several types of abilities would be measured by time, Gs-Processing Speed, and Gt-Decision Speed (Sattler, 2001). Processing speed would be defined as the ability to rapidly scan and react to simple tasks, that may involve writing, printing, or may involve quick perceptual processing (Sattler, 2001). In contrast, Decision Speed would be the ability to provide answers quickly to tasks of slight or moderate difficulty, such as choice reaction time, or simple reaction time (Sattler, 2001).

Johnson and Deary (2011) further describe timed measures as two different ways of measuring processing speed. They describe one being psychometrically driven, in
which simple tasks are presented, and the variable is how many items are completed during a set time (Johnson & Deary, 2011). However, they describe others tasks that are more psychophysically based and drawn from experimental psychology, known as elementary cognitive tasks (ECT; Johnson & Deary, 2011). They describe 2 different ECTs, one being inspection time, in which a stimuli is presented for a set amount of time, then the subjects responds at their leisure. The other type of ECT would involve 2 different reaction time measures. The first would involve a simple reaction time, in which the subject is asked to quickly respond to either the presence or absence of a stimuli. The latter, choice reaction time, some minimal processing is required, in which 4 items may be presented on a computer screen, with 3 identical and 1 different, and the subject is asked to quickly denote which is different (Johnson & Deary, 2011).

Based upon previous design of executive function assessment utilized by Miyake et al. (2000), and based upon review of the literature review, choice reaction time, will be utilized for this study, and will be simply referred to as reaction time.

**Studies Involving Analysis of Timed Measures**

Miller and Vernon (1997) examined the developmental changes in processing speed with young children. As they discussed changes in speed of performance, they speculate that these changes may be due to increases in the stimulus identification and encoding, reductions in decision making time or faster response selection (Miller & Vernon, 1997). While these may in fact be different cognitive processes, they were categorized under the general construct of processing speed.

Miller and Vernon’s 1997 study involved two samples, one aged 4-6 (N=109), the other being college students (N=43), classified as adults. The measures were 8 computer
based assessments, involving simple tasks, such as indicating if 2 shapes were similar, or whether arrows were pointing one direction or another. Using a dependent variable of mean response times, the researchers found differences across all age groups, that mean time decreased with age, meaning tasks were completed more quickly. However, when reporting results, the researchers interchanged the terms reaction time, with processing speed. Incidentally, they also noted that noticeable improvements occurred with regards to reaction time, but not with accuracy; they noted that the only significant age improvements seemed to occur with more difficult tasks (Miller & Vernon, 1997).

As a related implication of reaction time, Epstein et al. (2011) found that not only is reaction time important to analyze, but also the variability of reaction time is linked to certain pathologies such as ADHD. Their study, using a sample of 151 children aged 7-11. There were 3 groups, one with inattentive ADHD, one with combined ADHD and the third being a control group. To measure reaction time, the researchers used 5 different computer-based tasks. Results indicated that when comparing each of the groups, there were significant levels of variability, with the ADHD groups having greater variability than controls.

**The Most Critical Executive Function- Inhibition**

As suggested by Friedman, Miyake, Young, DeFries, Corley, and Hewitt (2008), inhibition seems to be the most significant of the three components as it relates to real-world problems with behavior. As described by these authors, multiple studies have shown a stronger association between inhibition and real-world problems, as compared to updating or shifting. Some of these problems include ADHD, depression, externalizing behaviors, and substance abuse. As a result, weaknesses in inhibitory skills should be
examined more closely as a risk for the development of mental health disorders in childhood and adolescence.

As applied to the central executive, inhibition can be seen as the ability to inhibit dominant, automatic or prepotent responses (Friedman et al. 2008). Miyake et al. (2000) explains it as “one’s ability to deliberately inhibit dominant, autonomic, or prepotent responses when necessary” (p. 57). Sometimes referred to as response inhibition (RI), it can also be seen as the ability to withhold behavioral responses that are inappropriate in the current context, and is viewed as fundamental to various cognitive abilities and adaptive behaviors (Diamond, 2002). Kirkham, Cruess, and Diamond (2003) argued that working memory and inhibition together play a critical role in the ability to overcome “attentional inertia”—that is, focusing on the same previously relevant aspects of a stimulus even when contextual demands change.

**Response vs. Behavioral Inhibition**

Many theories have described inhibition in different ways. Beyond the scope of social or cognitive inhibition, Barkley (1997) describes behavioral inhibition as a key indicator of ADHD. Its function would involve several types of executive functions, but ultimately behavioral inhibition is involved with some level of motor control (Barkley, 1997). Given this aspect of motor control, most analysis of childhood pathology involves externalizing disorders such as ODD, and ADHD (Young, Friedman, Miyake, Willcutt, Corley, et al., 2009). In contrast, response inhibition refers to a cognitive control process that allows a child to inhibit prepotent, automatic or dominant responses (Young et al., 2009).
Young, and colleagues (2009) studied the genetic and environmental relationship between response inhibition and behavioral inhibition. They used a sample population of adolescents with externalizing behavior patterns. Their first research question focused on the structure and etiology of behavioral inhibition at two different points of adolescent development. The second question they asked was whether external behavior disorders can be explained by underlying cognitive processes (response inhibition). The third question was whether the relationship between behavior disinhibition and response inhibition can be explained by shared genetic factors, and if so, to what extent.

The study utilized a sample population taken from the Colorado Longitudinal Twin Study (LTS). The LTS sample was mostly Caucasian, residing close to Boulder, CO. This was a longitudinal study in which twins were assessed at age 12 then again at age 17. At age 12, the subjects were assessed for behavioral disinhibition, using several measures. In order to assess substance use, they were assessed with the Monitoring the Future questionnaire. In order to assess conduct disorder and ADHD, the subjects were assessed with the Diagnostic Interview Scale for Children, fourth edition (DISC-IV), combined with results from the Child Behavior Checklist (CBCL) and Teacher Report Form (TRF). Because this is part of a larger longitudinal study, ratings were taken from multiple years. Parent Scores from the CBCL attention and externalizing subscales assessed at 4, 7, 9, 10, 11, and 12 years of age were averaged, in order to establish lifetime pattern of symptoms. Teacher ratings of attention and externalizing subscales, which were not available for age 4, were taken from ages 7, 9, 10, and 12, and again averaged. The other type of externalizing was novelty seeking. At age 12, each twin completed the Junior Temperament and Character Inventory (J-TCI), an 18 item
questionnaire in which the twin answered questions about how often they engage in dangerous or risk taking behaviors. At age 17, substance use was assessed with the Composite International Diagnostic Instrument—Substance Abuse Module (CIDI–SAM; Cottler, Robins, & Helzer, 1989). The CIDI–SAM is a structured, face-to-face interview designed for administration by highly trained lay interviewers. The dependent variable for substance use at age 17 was the number of substances used repeatedly. Conduct and ADHD patterns were assessed, using the averaged data from the CBCL and TRF. Novelty seeking was assessed at age 17 using the Cloninger’s Tridimensional Personality Questionnaire—Short Form (TPQ), another 18 item questionnaire in which the teen was asked if and how often they engage in dangerous or risky behavior.

In order to assess executive functions, the researchers administered 9 subtests of computer based “laboratory tests.” Although the primary goal was to analyze response inhibition, other components (measuring updating and shifting) were also administered in order to help show that response inhibition does have discriminate validity in its relationship with behavioral inhibition. The keep-track, letter-memory, and spatial-2-back tasks were designed to assess updating (the ability to update the contents of working memory with new relevant information and delete no-longer-relevant information). With the keep-track task, participants saw lists of 15 words and at the end reported the most recently presented words belonging to two to four target categories. The letter-memory task had the participants continuously say aloud the three most recent letters in a serially presented list of unpredictable length, reporting the last three letters at the end of the list. Spatial-2-back task was a series of 25 locations flash and subjects had to indicate for each one whether it was the same as the one that had flashed two trials before. The dependent
measure for all three tasks was proportion correct. The number-letter, color-shape, and category-switch tasks assessed shifting (how efficiently participants could shift between two subtasks). Before each trial, a cue indicated which subtask to perform, with the cue sometimes being the same in two consecutive trials and other times being different. The dependent measures were the differences in response time between each trial. Lastly, the inhibition measures included anti-saccade, stop-signal, and Stroop task. These tasks were administered in the same fashion as the Miyake et al. study in 2000, described earlier in this chapter.

Statistical analysis was conducted using structural equation modeling. Regarding the first research questions, the results did show differences between monozygotic (identical) and dizygotic (fraternal) twins. Analysis showed that genetic factors do contribute in part to behavioral disinhibition, and that the etiology of behavioral disinhibition does undergo some changes during adolescent development. Several examples were cited. First of all, substance use was heritable during early adolescence (est. 58%) compared to later (est. 20%). Though the shift was not as drastic, the early adolescent sample regarding conduct problems showed higher heritability (est. 70%) compared to late adolescence (est. 49%). Lastly, the behavior of novelty seeking in early adolescence was estimated at 50% heritable, and by late adolescence estimated at 28%. Using the substance use as an example, the authors explain that while genes are still influencing substance use in late adolescence, there is a significant increase in the shared environmental component of variance of substance use from the early teenage years to the late teenage years. Regarding the overall latent variable of behavioral disinhibition,
results suggest that the behavioral disinhibition factor was highly heritable, with 59% of the variance explained by genetic influences (Young et al., 2009).

Regarding the second research question, statistical analysis was conducted using confirmatory factor analysis. At age 12, behavioral disinhibition significantly correlated (-.47) with response inhibition, (-.27) with updating, and (-.20) with shifting. Correlation with response inhibition was significantly larger than the correlation with the other two executive functions. Behavioral disinhibition also significantly correlated with response inhibition (-.39), updating (-.18), and shifting (-.17) at age 17. The correlation with response inhibition was again significantly larger than the correlation with updating and shifting. The results indicated that response inhibition shows discriminate validity as a component of executive control and is more closely related to behavioral disinhibition than the other two executive functions examined.

Regarding the third research question, results showed that the genetic connection between behavioral disinhibition and response inhibition was very similar at the two points in development (age 12 and age 17). In discussing the results, the researchers (Young et al., 2009) concluded that the strength of association between behavioral inhibition and response inhibition was quite remarkable, given the contrast of types of assessment used for each (computer based measures for response inhibition and rating scales clinical measures for behavioral inhibition). Young et al. (2009) further concluded that this was the first study to empirically show that the link between cognitive and behavioral inhibition “is driven by a shared biological vulnerability” (p. 126).

Given the strong links between EF, social/emotional problems and behavioral inhibition (Young et al., 2009), the next step may be to develop some type of link or
relationship between the process of school-based behavioral screening, and deficits in executive functions. Because some research has indicated a lack of fractionation of executive function in preschool aged children (Thorell et al., 2009; Lehto et al., 2003), it seems the earliest age feasible to recognize deficits may be ages 7-9, typically grades 2 through 4, while the children are still young enough to benefit from intervention. Another conclusion of these studies is that the nine EF areas most frequently used (Miyake et al., 2000; Friedman et al., 2004; Friedman et al., 2007; Young et al., 2009) have consistently produced a clear picture of the three most referenced executive functions (shifting, updating and inhibition). However, their measures do not have much ecological validity for a normal school setting. These measures are all computer-based, are all administered individually, and may take 2 hours to administer. It may be more feasible, however, to use some of the features of executive functions or an adaptive form of the 3 computer based measures, in order to explore the impact of executive function deficits. Typically the first step of the screening process is school-wide or classroom screening of observable behavior. While some research has shown this to be an evidenced based practice, there are limitations to using behavioral ratings and observational methods alone to identify children in need of intervention (Walker et al, 2005).

**Conclusion**

A possible improvement to universal screening is increased understanding of the role of executive functions, particularly inhibition, at all levels of screening, including primary, secondary or tertiary level of screening. Studies have clearly shown the link between externalizing disorders and executive functions (Friedman et al., 2007; Nigg et al. ,2006). A student who may develop outward behavior problems will be readily
obvious, and should be “flagged” by an effective screening process. However, there
seems to be evidence (Emerson et al., 2005; Joorman & Gotlib, 2008; Dalgleish et al.,
2007; Leskin & White, 2007) that internalizing disorders are also linked to EF deficits. At
the same time, reviews of screenings methods (Walker et al., 2005) indicate a common
weakness is that children who may have internalizing disorders may be overlooked. This
would suggest that executive functions may have a significant impact on the behavior
screening process.
Chapter III

Method

This study involved a sample of male and female students from two, K-5 public elementary schools in Western Pennsylvania, located within the same school district. The general purpose of the study was to determine if executive functioning explains variance in a behavioral screener for Tier I RTI assessment in a sample of 4th grade public school children. Most measures of behavior screening involve observations and ratings of observable behaviors. This study was designed to contribute to the understanding of behavior screening, by reviewing the impact of certain cognitive functions for 9 and 10 year-old students, and how these functions may explain some level of variance in behavioral screening measure.

Participants

Subjects for the study were recruited from two different elementary schools within the same Pittsburgh Area school district. All students were in grade four and were mostly 9 or 10 years old at the time of the study (one student was age 11). The sample consisted of general education classrooms, within a public school system. Each school building had 4 different fourth grade classrooms, so the students came from 8 different classroom teachers (4 from each school).

School records were reviewed to note that the participants’ hearing acuity, visual acuity, and developmental status had been assessed because these variables can have an effect on academic performance. Screening for hearing and visual impairments is part of school regulations and students who do not pass are addressed on an individual basis by the school. Exclusionary criteria for this study included a previous diagnosis of mental
retardation, severe autism, or uncorrected sight or hearing problems because these conditions would likely have a significant impact on the measurement of executive functions, independent of other factors.

**Power Analysis**

The purpose of conducting an a priori power analysis is to establish the minimum number of participants needed to achieve adequate power. When using multiple regressions analysis, Eisenhauer (2009) indicates that high statistical power is not needed for statistical significance. For the purpose of a regression equation, the required number of subjects needed is based upon the number of predictors (Stevens, 2007). This study included several analysis, with a maximum on 3 predictors. While preselection of predictors is not unusual, Stevens (2007) points out that this may result in the results being sample specific; this will be considered and reviewed in the discussion section. The power analysis was conducted using the power tables provided by Cohen (1988), with an alpha level set at .05. With \( u \) representing the number of predictors (\( u=3 \)), power set at .80, and \( \alpha = .05 \), a trial choice for subjects of \( N=60 \) indicates \( \lambda = 11.5 \). Using an equation to determine sample size, \( N=11.5/\alpha = 77 \) (Cohen, 1988). Degrees of Freedom, represented by \( v \), for this analysis indicates \( v=73 \). A more accurate value for \( N \) is reiterated by interpolating between \( N=60 \) and \( N=120 \). Using Cohen’s (1988) formula, \( \lambda = 12.03 \). The new value becomes \( N=12.03/\alpha = 80 \). It should be noted, however, the study will aim to achieve a strong regression, rather than just statistical significance; even weak regressions can produce statistical significance, and therefore not practically applied or generalized (Eisenhauer, 2009).
Measures

Central Executive/Executive Functions

Three executive function measures were administered using programmed E-Prime-2 computer software (Schneider, Eschman, & Zuccolotto, 2007). This software allows for the accurate measurement of both correctness and reaction time associated with responses. These were adapted from previous studies such as Friedman et al. (2008) and Miyake et al. (2000). However, to address the need for usability in the school setting, and for use with younger students in the present study (9 and 10 year olds) rather than adolescents and college students in previous studies such as Miyake et al., (2000), the battery was adapted. Rather than using 9 total EF measures, which would have taken 2 hours to administer, only 3 were utilized, with the battery requiring approximately 15 minutes to administer. Participants were seated approximately 18 inches from the computer screen and indicated answers verbally with the examiners pressing associated keys on the keyboard.

Inhibition was measured using a variation of the classic Stroop task (Stroop, 1935) and a variation of the task used by Friedman, et al. (2008). The Friedman et al. study’s subjects were approximately 17 years old. Because the current study used younger subjects, the number of trials was reduced. The Stroop task consists of color words presented in matching shades (i.e.: yellow printed in yellow ink) and non-matching shades (i.e.: yellow printed in blue ink). It involves the purposeful stopping of the prepotent response of reading the word rather than naming the color the word is printed in. Participants will be presented with three types of trials: color words printed in the matching color, color words printed in a non-matching color, and a string of asterisks
printed in one of the colors used. To account for the younger age of the subjects, the researcher will press the response button on the computer, rather than directing the subject to press the button (as was done in the Friedman et al. study from 2008).

To measure shifting this study used the Category Switch task, also employed by Friedman, et al. (2008). In this assessment, participants were required to switch between categorizing objects as either a living vs non-living object or bigger vs smaller than a soccer ball based on a symbol presented with the word. If the object was presented with a heart, participants were to classify the object as living or non-living. Meanwhile, when the object was accompanied by a plus sign, the subject was to indicate whether the item was bigger than- or smaller than- a soccer ball. Again, in order to modify this from the previous study, the number of trials was reduced. In addition, the level of vocabulary was reviewed to determine the appropriate grade level for this fourth grade sample. This test was administered using e-prime (Schneider et al., 2007). Again, to account for the younger age of the subjects, the researcher pressed the response button on the computer, rather than directing the subject to press the button (as was done on the Friedman et al. study from 2008).

Updating was assessed with a Letter Memory task (adapted from Friedman et al., 2008; adapted from Morris and Jones, 1990). In this task, several letters were presented for 2.5 seconds per letter as part of varying list lengths (five, seven, or nine letters). Participants were required to recall the last 3 letters in the series of letters, causing them to drop the letter that was fourth back while adding in the new letter. The letters were presented in the computer program e-prime (Schneider et al., 2007). Using the Friedman et al. (2008) study’s approach of ensuring that participants continue updating, participants
were required to rehearse the three most recent letters aloud. For example, if the letters presented were G, N, A, F, K, E, Q, the participants would have said “G…G-N…G-N-A…N-A-F…A-F-K…F-K-E…K-E-Q…” and then recalled K-E-Q at the end of the trial. The number of letters presented in each trial (5, 7, or 9) was varied randomly across trials to ensure that participants continue to update the information. Participants were required to provide a verbal answer and will earn a score based on the total number of letters recalled correctly. Total number of trials was reduced from the previous study’s number of trials (Friedman et al., 2008).

**Behavioral Measure**

The BASC-2 Behavior and Emotional Screening System (BASC-2-BESS; Kamphaus & Reynolds, 2007) was completed by the classroom teacher for each student. The BASC-2-Screener is a quick, standardized, and accurate predictor of behavioral, emotional and academic problems. It consists of 23 rating scale questions and a parent, student, or classroom teacher can complete it in approximately five minutes. Prior to the study, researchers met with teachers to inform them about the intervention and their role. Teachers were asked to spend approximately 5 minutes to complete the BASC-2-Screener for each student involved in the study. If any previously unnoticed problems are discovered with this screener, a follow up would have taken place at the school involving the teacher and school psychologist. Any additional concerns were handled following school procedures.

The BASC-2-Screener produces a single total score (T-score) that is classified on one of three levels (normal, elevated at-risk, or extremely elevated at-risk). It also produces validity measures that indicate results that are inconsistent or overly negative.
In terms of psychometric characteristics, the Teacher form has strong internal reliability (r=.96), and strong inter-rater reliability (r=.70). Validity measures were also strong as compared to more broad-based rating systems. For example, correlation with the Behavior Assessment System for Children (BASC-2) yielded a relationship with the BASC-2 TRS Composite/Externalizing (r=.80) and BASC-2 TRS Composite/Internalizing (r=.64). Correlations also exist with the Achenbach System of Empirically Based Assessments-Teacher Rating Form (total problems, r=.75) (Kamphaus & Reynolds, 2007). Similar to the full version of the BASC-2, several options for norms are available, including girls only, boys only and combined. Because gender was not a variable in this study, combined norms were used to calculate T-scores.

The homeroom teacher for each subject completed this measure one time for each student participant. This study utilized the T-score as the dependent variable, which will be a continuous variable, rather than using the ranges of classification described by the publishers (Kamphaus & Reynolds, 2007). As it is intended as a screening tool, each student will have 1 scale completed by their homeroom teacher (see Table 1).

Table 1.

*Elementary School 1*

Teacher 1- 1 rating completed for each student in the study  
Teacher 2- 1 rating completed for each student in the study  
Teacher 3- 1 rating completed for each student in the study  
Teacher 4- 1 rating completed for each student in the study

*Elementary School 2*

Teacher 5- 1 rating completed for each student in the study
Teacher 6- 1 rating completed for each student in the study
Teacher 7- 1 rating completed for each student in the study
Teacher 8- 1 rating completed for each student in the study

55 total subjects; 8 total teachers, from 2 different schools

**Research Design**

This study will use an explanatory research design, with 6 independent variables (IV) and one dependent variable (DV). The independent variables will be the accuracy totals and mean reaction times of each of the computer based assessments (inhibition, shifting, and updating). Because these are adapted from a previous study and are not nationally normed, mean and standard deviation will be calculated from this sample for each EF measure. These first two IVs measuring inhibition will involve mean reaction time (for each item) and total accuracy (I-RT, I-A) of the Stroop test. The next IVs, measuring shifting, will be based on the mean reaction time (for each item) and total accuracy (S-RT, S-A) using the Category Switch task (32 items/trials). The last IVs, updating, again based upon the mean reaction time and the total accuracy of the computer based Letter Naming task (20 items/trials). The dependent variable was the standardized T-score on the BASC-2 Screener for each subject.

**Procedures**

Subjects were recruited from a public school setting, from two elementary schools, both part of the same school district located in Western Pennsylvania. The sample was taken from the 4th grade only, of which there are approximately 100 students in each building (200 total). The general description of the study was submitted to and approved by the district administration. Letters describing the study were sent home for
parental review by each of the 4th grade homeroom teachers. Included with each was a
description of the study, contact numbers and emails for the Duquesne researchers, and
consent forms for the parents to complete and return to Duquesne.

Once volunteers were garnered and consents received, the Duquesne researchers
identified the subjects. Attempts were be made to balance the following: number of male
and female subjects, number of students from each school building, and number of
different students at different age groups.

Once identified, the researchers provided a list to each of the homeroom teachers,
along with blank forms of the BASC-2 screener, with each student’s name on it.
Homeroom teachers then completed a form for each student. Upon completion these were
gathered by the researchers. Once the BASC-2 screeners were received and scored, each
subject was de-identified by blanking out the name and substituting a number code.

The next stage of the test administration, involving the executive function (EF)
measures, began with a scheduled day of assessments. This was coordinated with each of
the building principals and classroom teachers to minimize disruptions, and to avoid
conflict with activities such as field trips or state assessments. Because this occurred in
the spring of the school year, the whole project was completed prior to the mandated state
“assessment window” for districts. The executive functions measures were administered
during the school day—one day for each school--by the Duquesne researchers.

Because this component was computer administrated, the codes matching the
BASC-2 screener forms were entered on the laptop. This ensured that responders were
de-identified, but could be matched to the scores on the BASC-2 screener. The EF
measures were designed to be brief measures, ones that could be used practically in a
normal public school setting. However, they were administered individually, lasting approximately 15-20 minutes for each subject. The students were escorted from class by one of the Duquesne researchers, to another room such as a conference room or empty classroom. Once in the room, the EF measures and the purpose of the study were briefly described by one of the researchers, and assent was requested from the student, who signed their name if they agreed to participate. They were told that they could withdraw or discontinue at any time. Once assent was received, the assessment was completed, and the student was then escorted back to the classroom.

**Data Analysis**

The first research question: What is the relationship between the executive function measures and the behavioral measure? Based upon the review of the literature, it was hypothesized that reaction time for the inhibition measure would account for more variance on the behavioral measure than the accuracy measure for inhibition. Secondly, it was hypothesized that the reaction time for shifting would account for more variance on the behavioral measure than the accuracy measure. Lastly, it was hypothesized that the reaction time measure for updating would account for more variance on the behavioral measure than the accuracy for the updating measure.

The second research question: Is there a difference between the combined variance of all of the reaction times for the EF measures (inhibition, updating and shifting), and the combined variance of the accuracy totals for the EF measures (inhibition, updating and shifting) on the BASC-2 universal screener? Based upon review of the literature, it was hypothesized that the reaction times would account for a
significant amount of variance, whereas the accuracy measures would not account for a significant amount of variance.

For use in each of the analysis, the total score (T-score) on the BASC-2 screener form were represented as a continuous variable to be compared with each of the three EF variables (inhibition, shifting and updating). The T-score from the BASC-2 screening form was compared with each of the three central executive components: inhibition, shifting and updating. Alpha levels were pre-set at $p<.05$. Assumptions that must be met include independence of cases, normality and homoscedasticity.

To address each of the first research question, hierarchical multiple regression was conducted to determine if the accuracy and reaction time of the EF measures (inhibition, shifting and updating) have significant relationships with at-risk behaviors (as measured by the BASC-2 screener); for this question, the reaction time and accuracy on the EF measures were the independent variables, while the score on the BASC-2 screener was the dependent variable. For each regression equation, a priori decision making entered in the order of inhibition, then shifting, then updating, based upon previous literature showing strength of correlation (Young et al., 2009). Hierarchical regression analysis was also conducted on the second research question, where the independent variables were the three reaction time measures for hypothesis one and the three accuracy measures for hypothesis two, and the dependent variable is the scores on the BASC-2 screener. Assumptions would include fixed independent variables, independent variables measured without errors, and a linear relationship between IV and DV. For this analysis, the concern would be multicolinearity, as each of the IVs may have some high intercorrelations.
Research Questions and Hypotheses

Research Question 1:

What is the relationship between the executive function measures and the behavioral measure?

Hypothesis 1: Reaction time of the executive function inhibition will account for more variance on the BASC-2 universal screener for behavior than accuracy.

Hypothesis 2: Reaction time of the executive function shifting will account for more variance on the BASC-2 universal screener for behavior than accuracy.

Hypothesis 3: Reaction time of the executive function updating will account for more variance on the BASC-2 universal screener for behavior than accuracy.

Research Question 2:

Using measures of executive function (inhibition, shifting, and updating), is there a difference between reaction time and accuracy in the amount of variance contributed on the BASC-2 universal screener for behavior?

Hypothesis 4: The reaction time for inhibition, shifting and updating account for a significant amount of variance on the BASC-2 universal screener for behavior.
Hypothesis 5: The accuracy for inhibition, shifting and updating do not account for a significant amount of variance on the BASC-2 universal screener for behavior
Chapter IV

Results

This chapter outlines the results of the analysis conducted for the hypotheses proposed in Chapter One. First, the participants included in this study are described. The measures that operationalize the constructs in the hypotheses are discussed. Finally, specific data analysis steps will be discussed.

Descriptive Statistics

There were a total of 55 participants involved in the study, all of whom were included in the database for analysis. The mean age for the sample was 9.71 years, with 17 nine-year-olds, 37 ten-year-olds, and one 11-year-old. There were 25 male participants and 30 female participants, with approximately half of each sex comprised of students from each of the two schools. There were four classrooms in each school, with 29 students from one school and 26 students from the other school (see Table 2).

Table 2.

Participants in Each School and Classroom

<table>
<thead>
<tr>
<th>School</th>
<th>Classroom</th>
<th>Males</th>
<th>Females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>16</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>School</td>
<td>Classroom</td>
<td>Males</td>
<td>Females</td>
<td>Total</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9</strong></td>
<td><strong>17</strong></td>
<td><strong>26</strong></td>
<td></td>
</tr>
</tbody>
</table>

The behavior rating forms were completed only once, by each subject’s homeroom teacher. Means and standard deviations for each variable are reported in Table 3 and Table 4.

### Table 3.

*Means and Standard Deviations of Executive Function Measures - Accuracy*

<table>
<thead>
<tr>
<th>Variable</th>
<th>M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibition (Stroop)</td>
<td>43.69 (1.49)</td>
</tr>
<tr>
<td>Shifting (Category Switch)</td>
<td>31.78 (3.14)</td>
</tr>
<tr>
<td>Updating (Letter Memory)</td>
<td>26.55 (6.56)</td>
</tr>
</tbody>
</table>

*Note. N = 55, M = Mean; SD = Standard Deviation*
Table 4.

**Means and Standard Deviations of Executive Function Measures**

**Reaction Time**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$ (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibition (Stroop)</td>
<td>193.58 (201.98)</td>
</tr>
<tr>
<td>Shifting (Category Switch)</td>
<td>140.59(325.24)</td>
</tr>
<tr>
<td>Updating (Letter Memory)</td>
<td>7796.63(3820.31)</td>
</tr>
</tbody>
</table>

*Note. N = 55, M = Mean; SD = Standard Deviation*

The Behavioral Screening Measure (BASC-2 Screener) was completed by each of the students’ classroom teachers. This measure allows for three different types of norms, male, female and combined. Because gender was not a variable in this study, and the design was to use the behavioral measure as a first level universal screener, the combined norms were used to calculate T-scores from each raw score. Each teacher completed one rating for their students involved with the study. The study used the screener as a continuous variable, using the T-score ranging from, with a Mean of 50, and a Standard Deviation of 10. Although this variable was not categorized, the publishers recommend a T score falling from 20 to 60 is “normal,” 61 to 70 has an “elevated” level of risk and 71 or higher presents an “extremely elevated” risk level. Using this classification, 49 T-scores (89.1 %) fell within the normal range, 5 T-scores (10.9%) fell within the elevated range, and none fell at the extremely elevated range. The total sample ranged from 35 to 70 and frequencies of T-scores are shown in Table 5.
Table 5.

*Frequencies of Screener T-Scores*

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
<th>% of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>2</td>
<td>3.6</td>
</tr>
<tr>
<td>36</td>
<td>2</td>
<td>3.3</td>
</tr>
<tr>
<td>37</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>38</td>
<td>2</td>
<td>3.3</td>
</tr>
<tr>
<td>39</td>
<td>2</td>
<td>3.3</td>
</tr>
<tr>
<td>40</td>
<td>4</td>
<td>6.7</td>
</tr>
<tr>
<td>41</td>
<td>7</td>
<td>11.7</td>
</tr>
<tr>
<td>42</td>
<td>2</td>
<td>3.3</td>
</tr>
<tr>
<td>44</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>48</td>
<td>3</td>
<td>5.0</td>
</tr>
<tr>
<td>49</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>51</td>
<td>2</td>
<td>3.3</td>
</tr>
<tr>
<td>52</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>56</td>
<td>2</td>
<td>3.3</td>
</tr>
<tr>
<td>57</td>
<td>5</td>
<td>8.3</td>
</tr>
<tr>
<td>59</td>
<td>1</td>
<td>1.7</td>
</tr>
</tbody>
</table>
### Preliminary Analysis of Data

Prior to conducting the analysis, the dataset was examined to determine if analytic assumptions were met. In order to examine the multicollinearity of independent variables, correlations were examined (see Table 6). In addition, plots of residuals and histograms of residuals were examined to determine linearity and homoscedasticity, and normality. Missing data was addressed first. It was found that the database did not include any missing data. One of the measures, Letter Memory Reaction Time had high scenes (5.26) and kurtosis (32.36) with values much higher than recommended. An examination of the data indicated one extreme outlier. This data point was deleted, and replaced with the group mean for the analysis, which brought the skewness and kurtosis to more reasonable levels.

Mahalanbis distances were used to examine data for multivariate outliers, with a critical value set at 13.82 of the Chi-Squared distribution for each of the analysis run. No values exceeded the established critical value.
Table 6.

*Correlations Among Executive Function Measures*

<table>
<thead>
<tr>
<th>Measure</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.St-A</td>
<td>-.134</td>
<td>.258</td>
<td>.235</td>
<td>.338*</td>
<td>-.257</td>
</tr>
<tr>
<td>2.St-RT</td>
<td>-.097</td>
<td>.034</td>
<td>-.209</td>
<td>.099</td>
<td></td>
</tr>
<tr>
<td>3.CS-A</td>
<td>.211</td>
<td>.245</td>
<td>.051</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.CS-RT</td>
<td>-.028</td>
<td>-.039</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.LM-A</td>
<td></td>
<td></td>
<td></td>
<td>-.344*</td>
<td></td>
</tr>
<tr>
<td>6.LM-RT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* *p* < .05.; St-A = Stroop, Accuracy (Inhibition); St-RT = Stroop, Response Time (Inhibition); CS-A = Category Switch, Accuracy (Shifting); CS-RT = Category Switch, Response Time (Shifting); LM-A = Letter Memory, Accuracy (Updating); LM-RT = Letter Memory, Response Time (Updating).

Table 7.

*Correlations Between Executive Function Measures BASC-2 Screener*

<table>
<thead>
<tr>
<th>BASC-2 Screener</th>
<th>Accuracy Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.St-A</td>
<td>-.302*</td>
</tr>
<tr>
<td>2.CS-A</td>
<td>-.140</td>
</tr>
</tbody>
</table>
3.LN-A  -.295*

Reaction Time Measures

1.St-RT  .123
2.CS-RT  .010
3.LM-RT  .366

Note. *p<.05.; St-A=Stroop, Accuracy (Inhibition); St-RT=Stroop, Reaction Time (Inhibition); CS-A=Category Switch, Accuracy (Shifting); CS-RT=Category Switch, Reaction Time (Shifting); LM-A=Letter Memory, Accuracy (Updating); LM-RT=Letter Memory, Reaction Time (Updating).

Multicollinearity

According to Stevens (2007), high correlations (> .80) among multiple independent variables may present a risk for multicollinearity. An examination of the correlation matrix indicates that although several were statistically significant, none would be considered high, therefore multicollinearity is not problematic. The examination of assumptions, including normality, linearity, and homoscedasticity, was conducted using histograms and scatter plots after each research question.

Research Question One Results

The first research question examines the relationship between the executive function measures and the BASC-2 Screener. First of all, it was hypothesized that the inhibition-reaction time will account for more variance than inhibition-accuracy, on the behavioral screener. Using a hierarchical regression procedure, the inhibition reaction time (Stroop-RT) and the inhibition accuracy (Stroop-A) were entered as the independent
variables. The T-score on the behavioral screener (B-Sc) was entered as the dependent variable. Normality, linearity, and homoscedasticity are reflected in Figure 1, Figure 2, and Figure 3.

Figure 1. *Normal Distribution of Standardized Residuals for BASC-2 Screener with Stroop-Reaction Time and Stroop-Accuracy as IVs.*
Figure 2. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.
Figure 3. Scatterplot of Residuals around Regression Line for BASC-2 Screener

Satisfying Homoscedasticity Assumption

Result of the regression analysis indicates that the model did not account for a significant amount of variance, in the model summary and coefficient table represented in Table 8. The IVs accounted for 9.8% of variance in the Behavioral Screener, which was non-significant ($R^2 = .098$, $F(2, 52) = 2.83, p = .068$). The hypothesis was not confirmed, because the addition of the accuracy measure contributed more variance (Model 2, 8%) than did the reaction time measure (Model 1, 1%).
Table 8.

Effect of Stroop-RT and Stroop-A on BASC-2 Screener

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R_{adj}^2$</th>
<th>$R_{ch}^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.123</td>
<td>.015</td>
<td>-.003</td>
<td>.015</td>
<td>.371</td>
</tr>
<tr>
<td>2</td>
<td>.313</td>
<td>.098</td>
<td>.063</td>
<td>.083</td>
<td>.068</td>
</tr>
</tbody>
</table>

*Note.* Model 1- entered-Stroop-RT; Model 2- entered-Stroop-A

The second hypothesis was that shifting-reaction time measure would account for more variance on the behavioral screener than the shifting-accuracy measure. Again, using a hierarchical regression procedure, the shifting reaction time (Category Switch-RT) and the shifting accuracy (Category Switch-A) were entered as the independent variables. The T-score on the behavioral screener (B-Sc) was entered as the dependent variable. Normality, linearity, and homoscedasticity are reflected in Figure 4, Figure 5, and Figure 6.
Figure 4. *Normal Distribution of Standardized Residuals for BASC-2 Screener with Category Switch-Reaction Time and Category Switch-Accuracy as IVs.*
Figure 5. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.
Result of the regression analysis indicates that the model did not account for a significant amount of variance, with the model summary represented in Table 9. The IVs accounted for 2% of variance in the Behavioral Screener, which was non-significant ($R^2=.021, F(2, 52)=.561, p=.574$). The hypothesis stating that reaction time (RT) would account for more variance was again not confirmed, as Model 1, with only RT,
contributed near 0% variance, while Model 2, with the addition of accuracy, increased it to 2%.

Table 9.

*Effect of Category Switch-RT and Category Switch-A on BASC-2 Screener*

<table>
<thead>
<tr>
<th>Model</th>
<th>( R )</th>
<th>( R^2 )</th>
<th>( R_{adj}^2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.010</td>
<td>.000</td>
<td>-.019</td>
<td>.942</td>
</tr>
<tr>
<td>2</td>
<td>.145</td>
<td>.021</td>
<td>-.017</td>
<td>.574</td>
</tr>
</tbody>
</table>

*Note:* Model 1-entered-Category Switch-RT; Model 2-entered-Category Switch-A

The third hypothesis was that updating-reaction time measure would account for more variance on the behavioral screener than the updating-accuracy measure. Again, using a hierarchical regression procedure, the updating reaction time (Letter Memory-RT) and the updating accuracy (Updating-A) were entered as the independent variables. The T-score on the behavioral screener (B-Sc) was entered as the dependent variable. Normality, linearity, and homoscedasticity are reflected in Figure 7, Figure 8, and Figure 9.
Figure 7. Normal Distribution of Standardized Residuals for BASC-2 Screener with Letter Memory-Reaction Time and Letter Memory-Accuracy as IVs.
Figure 8. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.
Figure 9. *Scatterplot of Residuals around Regression Line for BASC-2 Screener*

*Satisfying Homoscedasticity Assumption*

Results of the regression analysis indicate that the model did account for a significant amount of variance, with the model summary represented in Table 10. The IVs accounted for 16.5% of variance in the Behavioral Screener, which was significant.
The hypothesis stating that reaction time (RT) would account for more variance was again confirmed, as Model 1, with only RT, contributed 13.1% variance, while Model 2, with the addition of accuracy, increased it by 3%.

Table 10.

*Effect of Letter Memory-RT and Letter Memory-A on BASC-2 Screener*

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R²</th>
<th>R² adj</th>
<th>R² cha</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.362</td>
<td>.131</td>
<td>.115</td>
<td>.131</td>
<td>.007*</td>
</tr>
<tr>
<td>2</td>
<td>.406</td>
<td>.165</td>
<td>.133</td>
<td>.034</td>
<td>.009*</td>
</tr>
</tbody>
</table>

*Note: Model 1-entered-Letter Memory-RT; Model 2-entered-Letter Memory-A*

*p<.01

**Research Question Two Results**

The second research question examines whether the reaction times of all the Central Executive Measures account for more variance than the accuracy of the Central Executive Measures on the BASC-2 Screener. First of all, it was hypothesized that the reaction times for all of the Central Executive Components (inhibition, shifting and updating) will account for a significant amount of variance on the behavioral screener. Using a hierarchical regression procedure, the inhibition reaction time (Stroop-RT), the shifting reaction time (Category Switch-RT), and updating reaction time (Letter Memory-RT) time were entered as the independent variables. The T-score on the behavioral
screener (B-Sc) was entered as the dependent variable. Normality, linearity, and homoscedasticity are reflected in Figure 10, Figure 11, and Figure 12.

Figure 10. Normal Distribution of Standardized Residuals for BASC-2 Screener with Stroop-Reaction Time, Category Switch-Reaction Time and Letter Memory-Reaction Time as IVs
Figure 11. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.
Results of the regression analysis indicates that the model did not account for a significant amount of variance, with the model summary and coefficient table represented in Table 11. Although non-significant, the IVs accounted for 13.9\% of variance in the Behavioral Screener (R^2 = .139, F(3,51) = .274, p = .053).
Table 11.

*Effect of All Reaction Times (Stroop-RT, Category Switch-RT and Letter Memory-RT) on BASC-2 Screener*

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2_{adj}$</th>
<th>$R^2_{cha}$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.123</td>
<td>.015</td>
<td>-.003</td>
<td>.015</td>
<td>.371</td>
</tr>
<tr>
<td>2</td>
<td>.123</td>
<td>.015</td>
<td>-.023</td>
<td>.000</td>
<td>.673</td>
</tr>
<tr>
<td>3</td>
<td>.373</td>
<td>.139</td>
<td>-.088</td>
<td>.124</td>
<td>.053</td>
</tr>
</tbody>
</table>

*Note:* Model 1-entered-Stroop-RT; Model 2-entered-Category Switch-RT; Model 3-entered-Letter Memory-RT

The final hypothesis was that the accuracy for all of the Central Executive Components (inhibition, shifting and updating) would not account for a significant amount of variance on the behavioral screener. Using a hierarchical regression procedure, the inhibition accuracy (Stroop-A), the shifting response time (Category Switch-A), and updating accuracy (Letter Memory-A) time were entered as the independent variables. The T-score on the behavioral screener (B-Sc) was entered as the dependent variable.
Figure 13. Normal Distribution of Standardized Residuals for BASC-2 Screener with 
Stroop-Accuracy, Category Switch-Accuracy and Letter Memory-Accuracy as IVs
Figure 14. Scatterplot of Standardized Residuals for BASC-2 Screener Showing Tenability of Linearity Assumption.
Result of the regression analysis indicates that the model did not account for a significant amount of variance, with the model summary and coefficient table represented in Table 12. Although the variance in the Behavioral Screener, was non-significant ($R^2=.134$, $F(3,51)=, p<.060$), it did account for 13.4% of the variance. Review of $R^2$ Change indicated that Stroop-Accuracy did statistically significant account for more explained variance (Significant F change, $p<.05$).

Figure 15. *Scatterplot of Residuals around Regression Line for BASC-2 Screener*

*Satisfying Homoscedasticity Assumption*
Table 12.

Effect of All Accuracy (Stroop-A, Category Switch-A and Letter Memory-A) on BASC-2 Screener

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2_{adj}$</th>
<th>$R^2_{cha}$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.302</td>
<td>.091</td>
<td>.074</td>
<td>.091</td>
<td>.025</td>
</tr>
<tr>
<td>2</td>
<td>.308</td>
<td>.095</td>
<td>.060</td>
<td>.004</td>
<td>.074</td>
</tr>
<tr>
<td>3</td>
<td>.366</td>
<td>.134</td>
<td>.083</td>
<td>.039</td>
<td>.060</td>
</tr>
</tbody>
</table>

*Note:* Model 1-entered-Stroop-A; Model 2-entered-Category Switch-A; Model 3-Letter Memory-A
Chapter V

Discussion

This study examined the relationship between a measure of universal behavioral screener for school age children and executive functions, specifically the fractionated components of the central executive (Miyake et al., 2000). This was done using a series of executive function measures adapted from those used by Miyake and colleagues in 2000, along with teacher-completed behavior screeners (BASC-2 Screener; Kamphaus & Reynolds, 2007).

Summary of Results

Although methods for behavioral screening in schools have become more common as typical education practices, considerable variability persists, regarding which methods are most effective and accepted in public schools (Walker et al., 2007). As some of these methods are designed differently, or target different types of behaviors (internalizing or externalizing), some rates of error, or variance have been shown (Severson et al., 2007). For this study, it was hypothesized that several different ways of measuring executive functions (reaction time and accuracy) would explain some level of variance in the behavioral screening tool.

The first research question related to relationships between the behavioral screener and each specific executive function. Results of the first analysis to examine the relationship between inhibition and the behavioral screener, indicated that the amount of variance explained was non-significant. The second analysis, looked at the relationship between the behavior screener and shifting, using measures of accuracy and reaction time. For this analysis, the amount was non-significant, but it did account for 13.4% of
variance on the behavioral screener. The third analysis did yield significant results. In this case, the amount of variance explained by the executive function updating was significant (p < .01). In addition, the hypothesis was confirmed that the reaction time explained a greater amount of variance than accuracy.

The second research question hypothesized that the reaction times for all the executive function measures would account for a significant amount of variance, whereas the accuracy for all executive function measures would not account for a significant amount of variance. The first analysis indicated that the reaction time for all of the executive function measures did not account for a significant amount of variance. Comparing of each executive function, updating accounted for the most variance, followed by the inhibition and shifting. The last analysis confirmed the research hypothesis, as the accuracy totals for all of the executive function measures did not explain a significant amount of variance on the behavioral screener. However, in this case, the accuracy for inhibition did explain the greatest amount of variance, and was statistically significant (p < .05).

Although it was not a research question, the correlations between the executive function measures and the behavioral screener were notable. While these correlations were not statistically significant, all six had consistent trends. All three accuracy measures of executive function were negatively correlated with the behavioral screener. This would indicate that as teacher reported behavioral concerns increased for each, the accuracy decreased. Likewise, for reaction time, the executive function measures were all positively correlated with the behavioral screener. As the behavioral problems became
more elevated, increased amount of reaction times were shown (it took more time for the
subject to react and provide their answer).

**Conclusions**

Some of the results were unexpected given how strong some previous studies
have shown the relationship is between the fractionated components of the central
executive and objective measures of behavior (Friedman et al., 2004; Friedman et al.,
2007; Young et al., 2009). In particular, response inhibition has been shown to be the
most strongly related to behavior (Nigg et al., 2006, Young et al., 2009). However, many
of these studies, such as Young et al. in 2009, used multiple methods of behavioral
assessment, rather than brief behavioral screeners.

Severson et al. (2007) report many methods of behavioral screeners, some more
effective with different type of pathologies, but each has some level error, “missing” 15-
30% of students with behavior disorders. The BASC-2 screener itself reports a validity
coefficient of less than .20, indicating that with such an internalizing disorder, the
screener may not have adequate predictive power (Kamphaus, 2007).

One significant result that was unexpected was that the executive function with
the strongest relationship to the behavior screener was updating. This may suggest some
developmental differences, with updating having a stronger link to behavior at the ages of
9 and 10 years old, with other studies such as Young et al. (2009) showing the strongest
link between behavior and inhibition at age 12.
**Implications**

This study shows that a behavioral screener does have some relationship to laboratory based measures of executive functioning. Because executive functions involve the ability to engage in good decision making, shift between mental sets, and inhibit prepotent responses, problems with behavior in school are implicated (Friedman et al., 2007). This link between behavior screening and executive functions may provide not only an understanding of the weaknesses in predictive power of screeners, but this may also suggest a link between specific executive functions such as inhibition, shifting and updating. However, this study does not support a strong link between response inhibition and behavior, as has been shown consistently in the literature (Nigg et al., 2006; Young et al., 2009). This may indicate some weaknesses in the design of this study, and therefore limit the generalizability. Regarding the other components of the central executive the relationship was the strongest between updating and the behavioral screener. While this result was not expected, there is support for the link between depression and weaknesses with updating (Joorman & Gotlin, 2008). Although speculation, it could be that at this age, more variance could be attributed to an internalizing disorder such as depression.

The study was designed to target an age in which the subjects were young enough to be flagged for behavioral intervention, but old enough to have developed some “diverse” skills in the fractionated components of the central executive. In 2007, Tsujimoto, Kuwajima, and Toshiyuki demonstrated this fractionation in ages 8 and 9, which led to the decision to use 4th grade students (primarily 9 and 10 years old) for the present study. In comparing the results of this study to the Young et al. study in 2009, the relationships between executive function measures and behavioral measures seem
different. While this may indicate some obvious differences in the study design and the measures used, it may also indicate some developmental differences. The results of the present study may suggest that the development of updating follows a different developmental trajectory that inhibition and shifting, and therefore has a stronger relationship to behavior at ages 9 and 10.

Another implication of this study is that, some support was added to the significance of reaction time over accuracy, as it relates to behavior, consistent with previous research (Epstein et al., 2011). Although the results were mixed, there appeared to be a consistent trend that as accuracy increased, behavior problems decreased, and as reaction time increased, behavior problems also increased. Although this is partly speculation, this may add to practice implications for school psychologists. A wide variety of screenings and assessments used in schools measure involve time measurement as well as correct answers (accuracy). This may indicate a need to integrate some type of reaction time measure into school-based screenings, similar to how academic universal screenings are conducted.

**Limitations**

The study was designed for a general education public school population, at the third grade level. Several limitations of the study relate to the sample used for the study. First, the sample would be too small to generalize to any public school system. Secondly, the subjects were solicited from one Western Pennsylvania school district, therefore not allowing for a representative sample, including a variety of socio-economic status, or ethnicity.
Another limiting factor may be how each executive function construct was measured. First of all, response time was a key measure. In previous studies adult or college age subjects, the subject was asked to quickly push a button on a computer keyboard, which would then produce a response time. For this study, the procedures were adapted, to accommodate the younger (9-10 year) subjects. While this may have addressed variability on the part of the children, or confusion about the testing guidelines, a new confounding, unmeasurable may have been introduced with each researchers individual reaction time.

Another limitation may have been how the assessments were adjusted to be easily administered in a school setting, and for 9- and 10-year olds. The number of items on each of the executive function measures was decreased. This may have decreased the reliability and validity of the measure. Future research should compare more traditional or complex measures of executive functions to determine if these may be better able to explain some level of variance in behavioral screening systems.

Although the behavioral measure (BASC-2 Screener; Kamphaus & Reynolds, 2007) does have good psychometric properties, the study design was limited with the absence of another outcome variable, such as placement in special education, or later referral for mental health services. This additional variable would have allowed further analysis to determine if the executive function measure would have added predictive power beyond that of the behavioral screener.

**Future Research**

To some degree, school-wide behavioral approaches are relatively new, as a commonly adopted practice in public schools (Severson et. al., 2007). At the same time,
this is a practice that has great implications for intervening early to prevent behavioral and emotional difficulties in school. Additional research is needed to determine what causes variations in predictive ability in variety of screening approaches. The study of school-wide behavior screening could also include some type of computer-based assessment of executive functions, to therefore increase the predictive power of screeners, particularly with school-age children who may have internalizing disorders, not as easily recognized through teacher ratings of behavior.

More specially, if individual differences in cognitive processes such as executive functions contribute to screening variance, future studies using different ages and different demographics, could examine if individual assessments of executive functions could better predict behavior problems in school.
References


APPENDIX A

IRB PROTOCOL

1. Statement of the research question

This study addresses research questions related to the effects of a mnemonics intervention as it applies specifically to working memory in fourth grade students.

Research question: What are the effects of a mnemonics intervention on the components of working memory? It is hypothesized that the components of working memory will be positively affected by a mnemonics intervention.

Another component that will be addressed is the relationship between behavioral patterns and the central executive of working memory. Research question: Is there a relationship between teacher-reported behavioral concerns and the central executive as measured by this working memory battery? Children who exhibit problematic behaviors in the classroom are hypothesized to have poorer performance on the three tasks designed to measure the central executive.

2. Purpose and significance of the study

Working memory is a memory system commonly described as a person’s ability to simultaneously store, manipulate and process information over a brief period of time (Baddeley, 2000). The current conceptualization of working memory is a four-component model with further fractionation within one of the components (Baddeley, 2000; Miyake, Friedman, Emerson, Witski, & Howter, 2000). The four main components are the episodic buffer, central executive and two slave systems, phonological loop and visuo-spatial sketchpad. Each component has its own specialized roles within the working memory system. In addition, the central executive has been fractionated into three distinct processes: inhibition, shifting, and updating (Miyake et al.).

The central executive is viewed as having both a common executive function mechanism as well as components that are partially dissociable, where both unity and diversity of the executive functions are necessary to their performance (Miyake et al., 2000). Baddeley’s theory explains working memory such that the central executive is responsible for manipulating information and controlling attention while Miyake and colleagues break the central executive into three functional domains: inhibition, shifting, and updating.

The first domain of central executive is the executive function of inhibition. Inhibition is defined as the ability to override dominant or automatic responses in order to complete the task at hand; the “deliberate, controlled suppression of prepotent responses”
Inhibition is associated with an individual's ability to maintain attention in the face of distracting stimuli and has been correlated with attention related abilities (Barkley, 1997). Shifting, the second domain of the central executive, is the ability to flexibly switch back and forth between tasks or mental sets, without integrating them together (Miyake et al., 2000). The third domain of the central executive is updating, which is the ability to actively monitor incoming information while appropriately replacing old, no longer relevant information with new, relevant information (Miyake et al., 2000).

Working memory is a rich topic of study in psychology, as it has correlates with learning and academic achievement (Schuchardt, Maehler, & Hasselhorn, 2008), attention deficits such as those found in children with ADHD (Barkley, 1997), and can provide an indication of other abilities such as intelligence (Conway et al., 2005), attention span (e.g., Kane, Poole, Tuhulska, & Engle, 2006), reading comprehension (Carretti, Cornoldi, DeBeni, & Romano, 2005), and other higher order cognitive functions (Bunting & Cowan, 2005). Although working memory has received significant attention and empirical support, the implications for understanding and measuring the constructs have yet to be comprehensively examined in children.

Given the strong relationship between working memory and the skills needed for academic success (e.g., attention; Gevins & Smith, 2000), it is of critical importance for practitioners to have evidenced-based working memory interventions. Currently there is limited research investigating the effectiveness of interventions on working memory. One purpose of this study is to examine the effects of a working memory intervention (i.e., mnemonic strategy) on the working memory components of fourth graders. Research has supported mnemonic devices as an intervention that improves performance on academic tasks (Verhaeghen, Marcoen, & Goossens, 1992). Researchers (e.g., Dehn, 2008) have reported mnemonics as being a compensatory strategy for improving the efficiency of working memory. However, there are currently no studies that examine the effects of a mnemonics intervention on the working memory components, so the compensatory nature of the intervention is speculation. Another purpose of this study is to investigate the claims of a mnemonics intervention as a compensatory strategy.

In addition to intervention effects on working memory, this study will investigate behavioral concerns. School wide behavior supports have become an increasingly important component of typical educational practice. These assessments attempt to screen for weaknesses in the cognitive control process that are observable via behavioral difficulties. To determine how those skills correlate with observable behaviors reported by a teacher completed behavior screening, an assessment for behavioral concerns is often included, representing ecological validity. Recent research has begun to show that there is a link between behavioral indicators and deficits in cognitive processing (Young, Friedman, Miyake, Willcutt, Corley, Haberstick, and Hewitt, 2009). This study will
expand these findings by exploring the relationship between teacher-reported behavioral concerns and working memory.

3. Research design and procedures

The research design employed will be an experimental design employing a matched group design where participants will be matched for gender across groups. Participants will be randomly assigned to one of two groups: treatment or control. Parental permission and student assent will be obtained prior to the first testing session. Testing and intervention sessions will occur during a period that is typically used as a study period. This is to prevent the students from missing class time.

**Pre-Intervention Phase**

Once permission and assent are obtained, all students will be assessed with the working memory battery. During that time, a behavioral screening rating form will also be distributed so that the teachers can complete them for the participating students. The teacher will complete the behavior screener one time for each student, prior to the start of the intervention.

**Intervention Phase**

Once all of the students are assessed, the intervention phase will begin. The delivery of the intervention will take place in an empty classroom. The treatment group will receive the keyword mnemonic intervention while the control group will not receive any services during this study. The intervention phase will last approximately four weeks, during which the researchers will offer a 20- to 30-minute training session two days per week during a time at the end of the day that will be discussed with teachers before the intervention begins. The training sessions will consist of scripted, manualized treatment plans for the implementation of the mnemonic intervention, as well as predetermined examples to practice. This includes a script to read each day (*Appendix 1*) and examples of mnemonic phrases with visual pictures provided (*Appendix 2*). All students in the intervention group will be required to attend at least five of the eight offered training sessions in order for their information to be used in the data analysis. Number of training sessions will not be analyzed beyond the minimal attendance policy in this particular study.

**Post-Intervention Phase**

All students will be assessed with the working memory battery. If the intervention is successful, the control group will receive the same keyword mnemonic intervention and post-tested again. The intervention that the control group receives will offer the same number of sessions the treatment group received within a similar timeframe.

4. Instruments

The six working memory components will be assessed in a neuropsychological battery designed to individually tap each construct. The instruments were selected based on previous research (Friedman et al., 2008; Garon et al., 2008) and are individually administered. Three of the assessments are administered with testing materials and will be recorded on paper protocols. Three of the assessments will be administered on a laptop computer using the program E-Prime 2 (Schneider, Eschman, & Zuccolotto, 2007). The
reliability coefficients that have been reported in the literature for the three computer administered measures have ranged from .85 to .91. A behavioral screener will also be utilized for additional behavioral information.

**Phonological Loop**

The assessment used to measure the phonological loop is the Digits Forward subtest found on the TOMAL-2 (Reynolds & Voress, 2008). This task requires the participant to recall strings of numbers that increase in length. This particular version of the digit span task is used because of its sensitivity to change. The Digits Forward subtest on the TOMAL-2 is scored based on individual digits correct as opposed to an all or nothing scoring system that is commonly found on other measures similar to this one.

**Visuo-Spatial Sketchpad**

The visuo-spatial sketchpad assessment being utilized is Memory for Location found on the TOMAL-2 (Reynolds & Voress, 2008). This assessment requires the participant to remember the location of dots on a black and white matrix board. Once the time has elapsed, the image is removed and the participant is given round bingo chips and a blank matrix where he/she is to place the bingo chips in the correct location. This subtest is scored based on accuracy of location. The reliability for Memory for Location ranges from .86 to .96.

**Episodic Buffer**

The measure being utilized for the episodic buffer is also a subtest on the NEPSY-2, the Word Generation subtest (Korkman, Kirk, & Kemp., 2006). This particular subtest requires participants to quickly and accurately name objects belonging to pre-determined categories over the time course of 60 seconds. Administration of this assessment involves the participant being instructed to name as many items as they can within the given category, such as words beginning with a particular letter, in a 60 second interval. Immediately before the timing starts, the participant is given the category to which the objects ought to belong. This requires the attentional component of maintaining the category while accessing previous knowledge of objects that may fall within that category. The reliability for Word Generation ranges from .60 to .77.

**Central Executive**

All three of these executive function measures will be conducted using programmed E-Prime-2 computer software (Schneider, Eschman, & Zuccolotto, 2007). This software allows for the accurate measurement of both accuracy and reaction time associated with responses. Participants will be seated approximately 18 inches from a computer screen that is approximately 15 inches wide and will indicate answers by
pressing associated keys on the keyboard or stating their answer aloud. Scores will be based on accuracy of response and overall time to complete the task.

**Inhibition.** Inhibition will be measured using a variation of the classic Stroop task (Stroop, 1935) and is a variation of the task used by Friedman, et al (2008). The Stroop task consists of color words presented in matching (i.e.: yellow printed in yellow ink) and non-matching (i.e.: yellow printed in blue ink). It involves the purposeful stopping of the prepotent response of reading the word rather than naming the color the word is printed in. Participants will be presented three types of trials: color words printed in the matching color, color words printed in a non-matching color, and a string of asterisks printed in one of the colors used. The task is to always state the color of the ink (*Appendix 3*).

**Shifting.** The measure of shifting utilized in this study was a Category Switch task, also used by Friedman, et al. (2008). In this assessment, participants are required to switch between categorizing objects as either a living vs. non-living object or bigger vs. smaller than a soccer ball based on a symbol presented with the word (*Appendix 4*). If the object is presented with a heart, participants are to classify the object as living or non-living. Meanwhile, when the object is presented along with a plus sign, they will indicate whether the item is bigger than- or smaller than- a soccer ball. This will be administered using e-prime (Schneider et al., 2007). Participants are required to answer by pressing stating their answer aloud. Answers will be scored for accuracy as well as overall time.

**Updating.** The first measure used to assess updating is a Letter Memory task (adapted from Friedman et al., 2008; adapted from Morris and Jones, 1990). In this task, several letters are presented for 2.5 seconds per letter as part of varying list lengths (five, seven, or nine) letters. Participants were required to recall the last 3 letters of the series of letters, causing them to drop the letter that was fourth back while adding in the new letter (*Appendix 5*). The letters were presented in the computer program e-prime (Schneider et al., 2007). Using Friedman et al. (2008) study’s approach of ensuring participants continued updating participants were required to rehearse the three most recent letters aloud. For example, if the letters presented were G, N, A, F, K, E, Q, the participants would have said “G…G-N…G-N-A…N-A-F…A-F-K…F-K-E…K-E-Q…” and then recalled K-E-Q at the end of the trial. The number of letters presented in each trial (5, 7, or 9) was varied randomly across trials to ensure that participants continued to update the information. Participants were required to provide a verbal answer and earn a score based on the total number of letters recalled correctly. Overall time will also be recorded for evaluation.

*Behavioral Screening Rating Form*

The BASC-2 Behavior and Emotional Screening System (BASC-2-BESS; Kamphaus & Reynolds, 2007) is a quick, standardized, and accurate predictor of
behavioral, emotional and academic problems. It consists of 20 rating scale questions and a parent, student, or classroom teacher can complete it in approximately five minutes. Prior to the study, researchers will meet with teachers to inform them about the intervention and their role. Teachers will be asked for approximately 5 minutes to complete the BASC-2 Screener for each student involved in the study. Any students who choose not to be part of the study will not receive any different treatment at school. If any previously unnoticed problems are discovered with this screener, a follow up will take place at the school involving the teacher and school psychologist. Any additional concerns will be handled following school procedures.

The BASC-2 Screener produces a single total score that is classified on one of three levels (normal, elevated at risk, or extremely elevated at-risk). It also produces validity measures that indicate results that are inconsistent or overly negative. In terms of psychometric characteristics, the BESS-Teacher form has strong internal reliability (r=.96), and strong inter-rater reliability (r=.70). Validity measures were also strong as compared to more broad-based rating systems. For example, correlation with the Behavior Assessment System for Children (BASC-2) yielded a relationship with the BASC-2 TRS Composite/Externalizing (r=.80) and BASC-2 TRS Composite/Internalizing (r=.64). Correlations also exist with the Achenbach System of Empirically Based Assessments-Teacher Rating Form (total problems, r=.75) (Kamphaus & Reynolds). Teachers of the student participants will fill out the behavioral screener (BESS) prior to the intervention process. The teacher will complete this measure one time for each student participant prior to the start of the intervention.

5. Sample selection and size

Participants will be recruited from the fourth grade at a local elementary school. This grade was chosen because it minimizes the developmental differences often seen in younger students while balancing the need to intervene as early as possible with regards to effective teaching and learning strategies. Participants will report their age and gender on a demographic cover sheet.

Based on the power analysis we need at least 40 participants. If we do not meet that requirement after the first information packet is sent home, we will send out that information packet again to the parents who did not respond.

6. Recruitment of Participants

Participants will be recruited from a Pennsylvania elementary school. The sample will be taken from the 4th grade only. A general description of the study was submitted and approved by the district administration (Appendix 6). Information describing the study will be sent home by each of the 4th grade homeroom teachers, for the parents to review. The information to be sent home will be a letter (Appendix 7) that describes the
study and provides contact numbers and emails for the Duquesne researchers, as well as a permission form (Appendix 8) for the parents to complete and return to the classroom teacher. Researchers will get the permission forms from the teachers. In addition, the participants will sign a student assent form (Appendix 9). The information packet may be sent out two weeks after the first packet to those who did not respond if we do not meet the required 40 participants needed for this study. The letter will provide a pre-determined cut-off date that is one week from when the letter was distributed to which a parent can submit a child’s name for participation in the study.

7. Informed Consent Procedures

The participants and their parents will be given a description of the project that will outline the experimental nature of the treatment, the procedure for random assignment of treatment and control groups, and the services that will be provided to the control group. If the intervention is successful, the control group will be offered the same intervention after the study is completed. Participants will be allowed to discontinue or withdraw from the study at anytime without any penalties. Parents will be required to return the permission form to researchers prior to beginning the pre-intervention phase. During the first testing session, the students will be informed that their parent agreed for them to volunteer and will receive an assent form to be signed as well. The researchers will describe the purpose of the study and answer any questions at that time. Students will be reminded that participation in this study is optional and involves minimal to no risk. Withdrawal will in no way affect treatment in school or their grades.

8. Collection of data and method of data analysis

Data will be collected both pre- and post- intervention to allow for the comparison of the performance on the working memory components. Once the data is collected, the E-Prime 2 computer program will store the data. A function of the program is its ability to merge data from multiple sources. Once it is merged, the data will be put into a database on one of the secure researchers computers. The other data collected on protocol forms will be stored in a secure location. Several researchers will be collecting data; each is trained and qualified to administer the neuropsychological measures being used with this population of students.

To evaluate the hypotheses, a multivariate analysis of covariance (MANCOVA) will be conducted to investigate a categorical independent variable with two levels (intervention and control) and multiple dependent variables (phonological loop, visuospatial sketchpad, episodic buffer, central executive inhibition, shifting, and updating). Two separate MANCOVA’s will be run according to the two theoretical models: phonological loop, visual-spatial sketchpad, and episodic buffer will represent one analysis while the three executive functions of the central executive will comprise the second. Multivariate analysis of covariance (MANCOVA) is a parsimonious statistical method that combines regression analysis and analysis of variance to provide information
which lends itself to more accurate conclusions (Stevens, 2002). The regression portion accounts for the correlation between two variables, the covariate, which may affect the outcome, such as the relationship between pretest score and post test score. Meanwhile, the analysis of variance examines the difference between levels of the independent variable, the treatment versus control group. Alpha levels will be pre-set at $p<.05$. The data analysis will include a comparison between the experimental and control groups.

Another use of the data will be in regard to the behavioral screener, data from the pretest results will be used. The total score on the BESS screening rating form will be represented as a dichotomous variable (normal or elevated/at-risk) to be compared with each of the three central executive variables (inhibition, shifting and updating). Final analysis will involve a Multivariate Analysis of Variance (MANOVA) to examine the relationship between the independent variable teacher-reported behavioral concerns as measured by the BESS (normal range and at-risk/elevated range) and dependent variables of executive functioning (inhibition, shifting, updating). The total score on the BESS screening rating form will be compared with each of the three central executive components, inhibition, shifting and updating. The different domains will be assessed with tasks adapted from Friedman and colleagues (2008) with the computer program e-prime (Schneider, Eschman, & Zuccolotto, 2007). Alpha levels will be pre-set at $p<.05$.

9. **Emphasize issues relating to interactions with participants and participants' rights**

Participants will be informed that all information collected will be entered into a private and confidential database, so that their information will not be identifiable to anyone with access to the data that is collected. Participation in this study involves minimal risk. The keyword mnemonic intervention is similar to a school-based lesson from the curriculum. The intervention is designed to be fun and engaging to children. However, if students become distressed in the course of the intervention, researchers will follow-up with the school and inform the parents. If a student becomes fatigued, he/she will be given the option to take a break that day.

The participants’ information will be kept confidential throughout the data collection and analysis process. Although researchers will be working directly with the students throughout the intervention phase, researchers will not be privy to the data collected in any way that can be traced back to the individual students so as to preserve the confidentiality of the participants and fidelity of the intervention. Researchers helping with data collection will sign a confidentiality affidavit (Appendix 10) that will be stored by the principal investigator. The information will be de-identified by removing the name of the student from the assessment results, and then entered into a database using numerical identifiers before any analysis takes place. Researchers will keep a master list of the students’ information in a separate location at Duquesne to be accessed by the researchers and used only when entering in the follow-up assessment data for the analysis.

References


Appendix 1

PLAN FOR INTERVENTION:

Sessions will consist of instruction from researchers provided to 8-10 students. Each session will be approximately 30 minutes and will be divided into two segments. The first part of each session will be teaching the mnemonic strategies and reviewing each time. The second part of each session will allow the students to apply a mnemonic of their own to provided words and states/capitals. This will provide an opportunity to practice using the keyword method.

1) Segment 1 – Technique-

“We are going to learn about some techniques that will help us learn and remember information better. The strategies are called mnemonics. You can create your own, or you can use one that was already created. Some of you may have used them, or heard of them. An example may be a way to remember something like the colors of the rainbow—remember Roy G. Biv—red, orange, yellow, blue, green, indigo, violet. We are going to go over a few ways you can help yourself remember other types of information. We’ll practice using one technique, one mnemonic, and help you with examples. Then you’ll get a chance to make some of these up on your own.

The first type of mnemonic is called the Letter technique. Teaching letter strategies involves the use of acronyms. Does anybody know what an acronym is? ——Acronyms are words whose individual letters can represent elements in lists of information, such as the word HOMES to represent the Great Lakes (write on a board or have this written out for them on a handout), Huron, Ontario, Michigan, Erie, Superior. Acrostics are sentences whose first letters represent to-be-remembered information, such as “My very educated mother just served us nine pizzas,” to remember the nine planets in order (e.g., Mercury, Venus, Earth, Mars).

Another technique is called Imagery. This has been used for thousands of years, beginning way back with the Romans. The story goes that Roman politicians and speakers used this method to remember parts of a speech given in front of a large crowd.
In order to remember the parts of the speech, they would visualize themselves walking through a house, and each room represented a part of their speech or story.

The last technique is the Keyword method. With this type, we create a picture to remember, along with a “keyword” that is kind of like a “code word” to help us remember. A teacher might teach a new vocabulary word by first identifying a keyword that sounds similar to the new word and is easily represented by a picture or drawing. Then the teacher would come up with a picture that connects the word to be learned with its definition. Here is an example:

A teacher is trying to teach her students the definition of the old English word carline. She would first identify a good keyword. In this instance, “car” is appropriate because it is easy to represent visually and it sounds like the first part of the vocabulary word. Carline means “witch”, so the teacher would show the students a picture of a car with a witch sitting in it. When asked to recall the definition of carline, students would go through four-steps:

1. Think back to the keyword (car),
2. Think of the picture (a car),
3. Remember what else was happening in the picture (a witch was in the car), and
4. Come up with the definition (witch)
2) Segment 2- Mnemonic Strategies-Applied

VOCABULARY WORDS

“So now we are going to apply the Keyword technique. We are going to use it to learn new vocabulary words, and to help us remember states and capitals. First of all, the vocabulary word:

The word is “bedlam.” And it means a state of chaos. Our keyword is “bed” and the image is a bunch of people running around on top of a bed.

The next word is “confer.” It means to meet and talk. The keyword is “fur” and the image is furry animals sitting around talking.

STATES AND CAPITALS

“Now we are going to use the technique for remembering states and capitals. We will have a keyword for the state, one for the capital, and one to link the two.

The first one is Salem, Oregon. The keyword for Salem is “sailboat,” the keyword for Oregon is “ore” and the image is a sailboat with an ore on it.

The next is Harrisburg, Pennsylvania. The keyword for Harrisburg is “hairy,” the keyword for Pennsylvania is “pen” so the image is a hair pen.

What would be some keywords for Trenton, New Jersey?”

<table>
<thead>
<tr>
<th>New Jersey</th>
<th>Trenton</th>
<th>A jersey hanging on a tent.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&quot;jersey&quot;)</td>
<td>(&quot;tent&quot;)</td>
<td>&quot;tent.&quot;</td>
</tr>
</tbody>
</table>
Appendix 2

Pennsylvania ("pen")

Harrisburg ("hairy")
Oregon
("ore")

Salem
("sailboat")
In this task, you will see a color word or line of stars on the screen.
Always say the color of the ink that you see.

You will complete several practice trials then the task will begin.

Please press the space bar when you are ready.
In this task, you will have to say whether something is larger or smaller than a soccer ball or whether it is dead or alive.

When you see a plus sign at the top of the screen you will say whether this thing in the middle of the screen is larger or smaller than a soccer ball.

When you see a star at the top of the screen you will say whether the thing in the middle of the screen is alive or not alive.

You will complete several practice trials then the task will begin.

Please press the space bar when you are ready.
In this task, you will see letters flash on the screen one at a time. As new letters appear on the screen you will always say aloud the LAST THREE letters you saw.

When you see the "???" you will say the last three letters that you saw then start over when you see a new letter.

You will complete three practice trials then the task will begin.

Please press the space bar when you are ready.
Appendix 6

I am a currently a 4th year doctoral student enrolled in Duquesne University’s School Psychology Program. As part of the program, I participate in a research group that studies neuropsychological issues related to teaching and learning. This is sometimes referred to as “brain-based learning.” Our group is interested in designing intervention strategies to be utilized in the school setting while also tapping into these neuropsychological functions. As a result, we are contacting your school district about participating in a research study relevant to this topic.

Response to Intervention (RtI) is a national and state initiative that most school districts are struggling to understand and implement. Many assessments and interventions have been developed to target basic reading skills. However, children have additional skill deficits such as behaviors related inattention as well as difficulties with planning and organizing their thoughts and behaviors. These problematic behaviors are typical of a child with deficits in a cognitive process known as Working Memory. Our research group has developed a series of brief “benchmark” assessments, along with targeted interventions to be used for fourth grade students. This grade was chosen because it minimizes the developmental differences often seen in younger students while balancing the need to intervene as early as possible with regards to effective teaching and learning strategies. Participation in this study would be voluntary, with each individual’s parent/guardian providing permission for his/her child to participate.
Appendix 7

Dear Parent or Guardian:

My name is Jessica Blasik and I am a doctoral student enrolled in Duquesne University’s School Psychology Program. As part of that program, I participate in a research group that studies neuropsychological issues related to learning. As a group, we attempt to investigate the ways that children learn. We base our research upon well-established findings and the increasing knowledge of how the brain processes information. As a result, we are in the process of setting up a research study involving fourth graders. In the school setting, many assessments and strategies have been developed to measure and improve basic reading skills. However, additional cognitive processes such as attention as well as planning and organizational skills also affect children’s learning. These aspects of learning fall under a cognitive process known as Working Memory. Our research group has developed a series of brief working memory tasks, along with training strategies to be used for fourth grade students. Participation in this study will be voluntary, with each individual’s parent giving permission for his/her child to participate. In addition, your child’s teacher will be asked to fill out a brief rating scale to provide researchers with information regarding classroom behavior. It is suggested that behavior concerns may be related to the cognitive processes associated with Working Memory. Researchers are also interested in studying the relationship of behavior and Working Memory.

The Duquesne University research group will be working directly with the students. Your child would meet with the researchers 10 times over a span of approximately six weeks. The first session will involve a series of working memory tasks. Three parts are administered on a laptop computer, and involve simple tasks such as recognizing colors of words, quickly identifying pictures on a computer screen, and repeating letters in a sequence. The other components involve recognizing patterns, verbally listing words, and repeating lists of numbers. During the following 4 weeks, your child would meet with the researchers two times per week to practice strategies to improve memory. Groups would consist of about 10 students and two Duquesne University researchers. At the end of the study, the working memory tasks will be completed again to see if any change has occurred in their working memory abilities. The study will be completed this spring. Researchers will offer a 20- to 30-minute training session two days per week during a time at the end of the day that will be discussed with teachers before the training begins.

If you do volunteer your child to participate, half of the volunteers will be randomly assigned as a “no treatment control” group. This means that this group of students would only participate in the “before” and “after” working memory tasks, and not the eight tutoring sessions. If there is improvement found, the training will be offered
to the no-treatment control group following the “after” phase of the study where those students will receive the same number of sessions within a similar timeframe. There is a deadline to volunteer for the study. Please complete the enclosed forms, and return them to your child’s homeroom teacher by February 19th, 2010. Permission forms will be collected at the end of that school day. In order to begin scheduling all of the sessions, additional participants cannot be included after that date.

Should you have any questions, please feel free to contact advisor Dr. Jeffrey Miller at (412) 396-4035 or researcher Jessica Blasik at (607) 765-2062. We look forward to working with you and your child. Thank you for your time and consideration regarding our study.

Sincerely,

Duquesne University Researchers
PERMISSION TO PARTICIPATE IN A RESEARCH STUDY

TITLE: MIND Program
Memory Intervention for Neuropsychological Development

INVESTIGATOR(S): Dr. Jeffrey A. Miller, Ph.D., ABPP
Professor of School Psychology
(412) 396-4035

Jessica Blasik, M.S.Ed.
1926 Larkins Way
Pittsburgh, PA 15203
(607) 765-2062

ADVISOR: (if applicable:) Dr. Jeffrey A. Miller, Ph.D., ABPP

SOURCE OF SUPPORT: This study is being performed as partial fulfillment of the requirements for the doctoral degree in School Psychology at Duquesne University.

PURPOSE: You are being asked to volunteer your child to participate in a research project that seeks to investigate research questions related to the effects of a memory training (specifically mnemonics) as it applies to working memory in elementary age students.
Participants will be randomly assigned to one of two groups: treatment or control. The treatment group will receive the training while the control group will not receive any services during this study. If the training is found successful, the control group will be offered the same training following the study. The same number of sessions will be offered to those students in the control group. Both groups will be evaluated during the week before and during the week after the training. The training will last four weeks, with two 20-30 minute sessions each week during a time at the end of the day that will be discussed with teachers before the training begins. Teachers will fill out rating scales for each individual participant to provide information related to classroom behaviors.

Only the treatment group will receive the training but all individuals will be assessed at the beginning and end of the training to examine its effectiveness in improving working memory processes. Each of the individual assessment batteries will include all working memory component measures, which should take approximately 20 minutes. These are the only requests that will be made of your child.

**RISKS AND BENEFITS:** There are no risks greater than those encountered in everyday life. The benefits include the potential to help participants use memory strategies in the classroom. Participation in the study may lead to more efficient memory and learning strategies.

**COMPENSATION:** Receiving a training beyond the regular school day will require no monetary cost to you. An envelope is provided for the return of your response to the investigator. Teachers will collect this and give it to researchers.

**CONFIDENTIALITY:** Your child’s name will never appear on any survey or research instruments. No identity will be made in the data analysis. All written materials and consent forms will be stored in a locked file on
Duquesne campus and will only be accessed by the members of the research group directly involved in the study. The response(s) will only appear in statistical data summaries. All materials will be destroyed at the completion of the research.

RIGHT TO WITHDRAW: Your child is under no obligation to participate in this study. He/she is free to withdraw from the study at any time with no penalties.

SUMMARY OF RESULTS: A summary of the results of this research will be supplied to you and your child, at no cost, upon request.

VOLUNTARY CONSENT: I have read the above statements and understand what is being requested of my child. I also understand that my child’s participation is voluntary and that he/she is free to withdraw at any time, for any reason. On these terms, I certify that I am willing to allow my child to participate in this research project.

I understand that should I have any further questions about my child’s participation in this study, I may call Principal Investigator, Dr. Jeff Miller, (412) 396-4035, Jessica Blasik, (607) 765-2062, and/or Dr. Paul Richer, Chair of the Duquesne University Institutional Review Board (412) 396-6326

_________________________________________ ____________________
Participant's Signature Date

_________________________________________ ____________________
Researcher's Signature Date
ASSENT TO PARTICIPATE IN A RESEARCH STUDY

TITLE: MIND Program
(Memory Intervention for Neuropsychological Development)

INVESTIGATOR(S): Dr. Jeffrey A. Miller, Ph.D., ABPP
Professor of School Psychology
(412) 396-4035

Jessica Blasik, M.S.Ed.
1926 Larkins Way
Pittsburgh, PA 15203
(607) 765-2062

ADVISOR: (if applicable:) Dr. Jeffrey A. Miller, Ph.D., ABPP

SOURCE OF SUPPORT: This study is being performed to help doctoral students at Duquesne University who are studying ways to help children learn better.
PURPOSE: You are being asked to participate in a research project that is trying to look at ways to improve memory for 4th graders using small group teaching games. Teachers will also fill out rating scales about you to provide information related to how you act in the classroom.

You will be asked to participate in some short problem solving activities before the training. You will learn how memory works and practice ways to use your memory. The training will last four weeks, with two 20-30 minute sessions each week during a time at the end of the day that will be discussed with teachers before the training begins. Then, you will be asked to complete some problem solving activities after the training. You will meet with the researchers about 10 times over two months. These are the only requests that will be made of you.

RISKS AND BENEFITS: There are no risks greater than those encountered in everyday life. Each session will be similar to tasks you would do in class. This may help you learn new ways to remember information.

COMPENSATION: Participation is voluntary and will cost you no money.

CONFIDENTIALITY: Your name will never appear on any survey or research instruments. Your name will not be used in the data analysis. All written materials and consent forms will be stored in a locked file on Duquesne campus. Your responses will only appear in statistical data summaries. All materials will be destroyed at the completion of the research.

RIGHT TO WITHDRAW: You are not required to participate in this study. You are free to take a break, stop, or withdraw at any time during the study. Withdrawal will not affect your grades in school or how your teacher treats you.
SUMMARY OF RESULTS: A summary of the results of this research will be supplied to you and your parent/guardian, at no cost, upon request.

VOLUNTARY ASSENT: I have read the above statements and understand what is being requested of me. I also understand that my participation is voluntary and that I am free to withdraw at any time, for any reason. On these terms, I am willing to participate in this research project.

I understand that should I have any questions about this study, I may call Principal Investigator, Dr. Jeffrey Miller, (412) 386-4035, Jessica Blasik, (607) 765-2062, and Dr. Paul Richer, Chair of the Duquesne University Institutional Review Board (412)396-6326

_________________________________________ __________________________
Participant's Signature Date

_________________________________________ __________________________
Researcher's Signature Date
CONFIDENTIALITY AFFIDAVIT FOR RESEARCH STUDY

I, ____________________________, am aware that the information I am helping collect is confidential material and will not share, distribute, disclose, or discuss the data or results with researchers not directly involved with the study. The principal investigators will have access to all of the de-identified raw data, but I may not be able to access this, because it will be kept in a secure location. This is consistent with the Ethical Standards of the American Psychological Association and IRB procedures.

I understand that the principal investigator will save this form representing a contract to maintain confidentiality of the participants. By signing this confidentiality affidavit, I am agreeing to uphold the confidentiality and therefore the integrity of the Memory Intervention for Neuropsychological Development study.

_________________________________________  ____________________________
Researcher's Signature                      Date
March 15, 2010
Dr. Jeff Miller
School of Education
Duquesne University
Pittsburgh PA 15282

Re: Memory intervention for neuropsychological development (Protocol # 10-10)

Dear Dr. Miller:

Thank you for submitting the research proposal of your student, Ms. Jessica Blasik, to the IRB.

After review by IRB members, Dr. Joseph Kush and Dr. Sarah Peterson, along with the entire Board, the study is approved under the federal Common Rule, specifically 45-Federal Code of Regulations 46.101 and 46.111. In addition, the study meets requirements set forth in subpart D, 46.404 (research with minors not involving greater than minimal risk).

Consent, permission and assent forms are stamped with IRB approval and one-year expiration date. Ms. Blasik should use the stamped forms as originals for copies that she displays or distributes.

The approval must be renewed in one year as part of the IRB’s continuing review. You and Ms. Blasik will need to submit a progress report to the IRB in response to a questionnaire that we will send. In addition, if the consent form is still in use in one year, it will need to be renewed by our office. In correspondence please refer to the protocol number shown after the title above.

If, prior to the annual review, you propose any changes in procedure or consent process, you must inform the IRB of those changes and wait for approval before they are implemented. In addition, if any unanticipated problems or adverse effects on subjects are discovered before the annual review, they must be reported to the IRB Chair before proceeding with the study.

When the study is complete, please provide us with a summary, approximately one page. Often the completed study’s Abstract suffices. You or Ms. Blasik should retain a copy of
research records, other than those that have been destroyed for confidentiality, over a period of five years after the study’s completion.

Thank you for contributing to Duquesne’s research endeavors.

If you have any questions, feel free to contact me at any time.

Sincerely yours,

Paul Richer, Ph.D.
C: Dr. Joseph Kush
    Dr. Sarah Peterson
    Ms. Jessica Blasik
    IRB Records