Executive functions (including attentional shifting/flexibility, working memory, and inhibitory control) are strong predictors of children’s early school success (Blair & Razza, 2007; Espy et al., 2004). The current study explored questions related to measurement of executive functions in preschool-aged children. Convergent and predictive validity were assessed for two traditional executive function tasks (the Dimensional Change Card Sort and the Day-Night Stroop), a behavioral executive function task (the Head-Toes-Knees-Shoulders, HTKS), and teacher ratings of child classroom behavior (the Child Behavior Rating Scale, CBRS). All measures were low-to-moderately correlated for the full sample of preschoolers. The CBRS and the HTKS tasks were the most consistent predictors of emergent mathematics, vocabulary, and literacy, controlling for child age and Head Start status; however, all tasks were significantly related to each achievement outcome. Additionally, the convergent and predictive validity of the executive function tasks and teacher ratings were examined by Head Start status. Results show that the tasks were more closely related in non Head Start children. For predictive validity, the most
notable difference was for the Day-Night Stroop, which was a strong consistent predictor of academic outcomes for non Head Start children but not for Head Start children. Together, these findings provide insight to the convergent and predictive validity of executive function tasks during early childhood and differences in executive function associated with Head Start status.
Measures of Executive Function: Convergent Validity and Links to Academic Achievement in Preschool

by
Robert J. Duncan

A THESIS
submitted to
Oregon State University

in partial fulfillment of the requirements for the degree of
Master of Science

Presented May 31, 2012
Commencement June 2012

APPROVED:

Major Professor, representing Human Development and Family Studies

Co-Director of the School of Social and Behavioral Health Sciences

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

____________________________________
Robert J. Duncan, Author
ACKNOWLEDGEMENTS

First, and foremost, I would like to thank Drs. Megan McClelland and Alan Acock for their continued guidance and mentorship throughout the thesis process. They have each been instrumental in my development as a graduate student in numerous ways. For Megan, thank you for your all of your help in the writing process of my thesis, your contributions have taught me a great deal about writing scholarly research. Also, thank you for all the time you have invested in my development of thinking critically about child development, I could not be where I am at without your support. For Alan, thank you for your guidance on statistical methods, as well as other numerous conversations. Your mentorship has truly taught me a great deal and I look forward to continuing to learn from you.

I would also like to thank Dr. Lesli Richards for her contributions as committee member and Dr. Karen Higgins as graduate school representative. Your contributions make this process possible and they are truly appreciated. Additionally, thank you to Dr. Samuel Vuchinich for your guidance and help in furthering my thinking on statistical methods.

Finally, I would like to thank all my friends and family for their continued support throughout the process, and specifically, my mom, dad, and brother for their love and support.
CONTRIBUTION OF AUTHORS

As a co-author, Dr. Megan McClelland assisted in the design of this study and manuscript. She provided editorial comments and helped with the formulation of thoughts and ideas. Additionally, Dr. Alan Acock provided help with statistical analyses and interpretation. Data were collected as a part of Megan McClelland’s child development lab, including contributions from numerous graduate and undergraduate students. Specifically, the graduate students that assisted in data collection for this project included Megan Pratt, Lupe Diaz, Sara Schmitt, Derek Becker, Alicia Miao, Jenny Moss, and Catherine Gunzenhauser.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Literature Review</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Method</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Results</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Discussion</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>Conclusion</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>41</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>4.1. Descriptive Statistics for the Full Sample, Head Start, and</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>non Head Start Samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.2. Correlations Among Three Executive Function Tasks and</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Teacher Ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.4. Correlations Among Three Executive Function Tasks and</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Teacher Ratings</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Measures of Executive Function: Convergent Validity and Links to Academic Achievement in Preschool

Chapter 1. Introduction

Children’s early development has important implications for a number of outcomes throughout life. One area of importance is the transition from early childhood environments to formal schooling, where children’s executive functions are needed to be successful in the classroom. The behavioral aspects of executive functions include the ability to sit still in classrooms, pay attention to teachers, problem solve, and be goal-directive (Blair & Diamond, 2008; McClelland & Ponitz, 2011; Morrison, Ponitz, & McClelland, 2010). Executive functions (including attentional shifting/flexibility, working memory, and inhibitory control) have received considerable interest and are strong predictors of children’s early school success (Blair & Razza, 2007; Espy et al., 2004; Gathercole & Pickering, 2000). Although the importance of these skills is well established, measuring these skills has been challenging because convergent and predictive validity of different measures are often unclear. The current study focuses on understanding relations between cognitive and behavioral tasks of executive function in preschool children. It also examined relations between different executive function tasks and early academic achievement. Finally, differences in convergent and predictive validity of executive functions between Head Start children and non Head Start children were explored.

Defining Executive Function

Blair, Zelazo, and Greenberg (2005) define executive functions as the cognitive processes involving working memory, inhibition of responses, and shifting/sustaining
attention to achieve a goal. Executive functions are often used as an umbrella term in
cognitive psychology to describe any one of these skills or the combination of skills
working together to solve a problem (McClelland, Ponitz, Messersmith, & Tominey,
2010). Executive functions are sometimes divided into hot and cool categories, with hot
executive functions referring to processes that involve an emotionally arousing
component (i.e., a Delay of Gratification task), and cool executive functions being
cognitive processes without emotional arousal (i.e., a cognitive flexibility task such as the
Dimensional Change Card Sort; DCCS) (Hongwanishkul, Happaney, Lee, & Zelazo,
2005). Recent research has focused on disentangling the contributions of different
components of executive function to determine whether there is a general executive
ability or discernible components (Miyake, Friedman, Emerson, Witzki, & Howerter,
2000; Monette, Bigras, & Guay, 2011; Wiebe, Espy, & Charak, 2008). Findings,
however, are limited because of measurement issues. Specifically, it is often unclear how
well different measures of executive functions are related and how this might change
across development. Furthermore, the relations between cognitive and behavior aspects
of executive function are unclear.

Behavioral aspects of executive function relate to the observable behaviors of the
underlying cognitive skills that make up executive function, also referred to as behavioral
regulation (Ponitz et al., 2009). For instance, as children develop, they integrate their
cognitive abilities into behavioral responses. These overt, observable aspects of
executive functions, such as raising a hand or sitting quietly and listening to a teacher, are
related to successful classroom behavior (McClelland & Ponitz, 2011; Ponitz et al.,
2009). Moreover, these skills are important as children transition from home or
preschool environments to more formal school environments, where appropriate classroom behaviors have implications for learning and academic success (Morrison et al., 2010).
Chapter 2. Literature Review

**Early Childhood and Executive Function**

Early childhood is an important stage for understanding the development of executive function because of drastic changes in brain maturation, especially the prefrontal cortex. Maturation in the prefrontal cortex helps children integrate and demonstrate executive function skills, with maturation continuing through the early adult years as the skills continue to develop, albeit at slower paces (Best & Miller, 2010; Center on the Developing Child at Harvard University, 2011). A growth spurt during the first seven years of life in the prefrontal cortex allows children to move from simple inhibition, working memory, and attentional shifting capabilities to more complex, integrated problem solving skills (Garon, Bryson, & Smith, 2008). Children develop these skills at different rates, and to different degrees, and those children who demonstrate strong behavioral aspects of executive function by the time they enter kindergarten tend to be more successful in the classroom and on learning tasks (McClelland, Acock, & Morrison, 2006; Ponitz et al., 2009). The changes that occur in the brain make the preschool years important to understanding executive function, but our understanding is dependent on the measures used for assessing these skills.

**Measurement of Executive Functions**

Executive functions have been measured using a number of different tasks in early childhood (Carlson, 2005; or for a review, see Garon et al., 2008). Executive function tasks assess both simple and complex aspects of working memory, response inhibition, and attentional shifting, with some tasks assessing multiple components. Two traditional tasks of cognitive executive functions are used in the current study the
Dimensional Change Card Sort (DCCS) and the Day-Night Stroop task, and one behavioral executive function task, the Head-Toes-Knees-Shoulders (HTKS) task. The DCCS is considered an attentional/set-shifting task because it requires the child to sort cards based on changing rules (Garon et al., 2008; Zelazo, 2006). The Day-Night Stroop task is considered a response inhibition task because the child has to inhibit a natural response and say the opposite, for instance the child must say “day” when shown a moon and “night” when shown a sun (Gerstadt, Hong, & Diamond, 1994). Both of these tasks have been used in a number of studies demonstrating how executive functions relate to mathematics and verbal skills in early childhood (Blair & Peters, 2003; Hongwanishkul et al., 2005).

Finally, a commonly used measure for behavioral executive function is the Head-Toes-Knees-Shoulders task (HTKS), which requires the incorporation of executive functions and gross motor responses. In the HTKS, children must pay attention, remember the instructions, and demonstrate self-control in a behavioral motor response by touching their head when asked to “touch your toes” and by touching their toes when asked to “touch your head”. Currently, it is unclear how closely the HTKS is related to the more traditional, cognitive assessments of executive function. Together, the DCCS, the Day-Night Stroop, and the HTKS have all been developed to assess children’s executive function skills during the preschool years (Carlson, 2005; Garon et al., 2008; Ponitz et al., 2009). The current study assessed the convergent and predictive validity of these three direct assessments, in addition to teacher ratings of child behavior, during the child’s final year of preschool.

**Convergent Validity of Executive Function Tasks**
Although a number of studies have compared the components of executive functions, few have focused on comparing across measures of cognitive and behavioral aspects of executive function. The research that has been done on the components of executive function has generally found modest relations and a single underlying dimension (Blair & Razza, 2007; Hughes & Ensor, 2011; Wiebe et al., 2008). For example, Hongwanishkul et al. (2005) found a correlation of $r = .34$ between tasks measuring working memory (measured by the Self-Ordered Pointing Task) and attention shifting (measured by the DCCS) in 3- to 5-year-old children. Blair and Razza (2007) found similar correlations between attentional shifting (item-selection task, similar to the DCCS) and inhibition (peg tapping task), $r = .34$ in prekindergarten and $r = .41$ in kindergarten. These relatively low correlations may persist beyond the early childhood years. For example, Archibald and Kerns (1999) reported correlations of $r = .23$ for working memory and inhibitory control in children 7- to 12-years old, using Self-Ordered Pointing and the Go/No-Go Stroop task. Low correlations between the components of executive function raise questions over whether the tasks are assessing one latent construct or whether it consists of discernible components (i.e., inhibition, working memory, attentional shifting). Recent research has used confirmatory and exploratory factor analyses to assess the different components of executive function (Allan & Lonigan, 2011; Miyake et al., 2000; Monette et al., 2011; Wiebe et al., 2008).

Wiebe et al. (2008) used confirmatory factor analysis to assess if ten executive function tasks loaded onto a unifying dimension or whether models with multiple latent variables were better fits statistically. Their study focused on children aged 2- to 6-years old, with three tasks primarily tapping working memory and seven tapping inhibitory
control. The study found that a model with one latent variable was the best fit, however, generalizability of these findings need to be interpreted cautiously. First, there were low correlations among tasks that were intended to assess a common construct and some tasks were more closely related with tasks of a different construct than the one intended. For instance, correlations between the two components (working memory tasks and inhibition tasks) were larger than some correlations between tasks for a single component (i.e., inhibition). These findings suggest that some measures are not converging onto a common construct as intended and therefore the contributions of a given task onto a latent variable may be compromised. Second, the study found that factor loadings varied when models were separated into a younger and older age group within their sample of preschoolers (2- to 4-year olds and 4- to 6- year olds). This finding suggests that tasks may be differentially related based on the age of the child. Therefore, the dimensionality of executive function is likely dependent on both the age of the child and the measures used to assess the components. Executive function may become more differentiated as children age (Miyake et al., 2000); however, in early childhood most research has consistently found support for one underlying dimension of executive function, which is shown to be related closely to early academic achievement (Allan & Lonigan, 2011; Hughes & Ensor, 2011; Wiebe et al., 2008). The present study examined the relations and dimensionality of three direct assessments of executive function and teacher ratings in a sample of preschool aged children.

**Predictive Validity of Executive Function Tasks to Academic Outcomes**

As noted previously, executive function skills are strong predictors of academic achievement in early childhood (Blair & Razza, 2007; Bull & Scerif, 2001; Espy et al.,
2004; McClelland et al., 2007). Some research has focused on the specific components of executive function (i.e., inhibition, working memory, or attentional shifting) predicting achievement, while others have used composite measures or latent variables of executive function in predicting achievement. In one study, Espy et al. (2004) found that executive function, using a composite score of multiple tasks, was related to emergent arithmetic ability. Additionally, the authors found that inhibitory control predicted unique variance in early mathematics when controlling for other executive function abilities. Espy et al.’s (2004) study suggests that inhibition may be a key component of early academic achievement, which is consistent with findings from Blair and Razza (2007), and Bull and Scerif (2001). Blair and Razza (2007) found that inhibitory control predicted unique variance in mathematics, phonemic awareness, and letter knowledge in low-income kindergarteners. Additionally, Bull and Scerif (2001) showed that performance of 7-year-olds on inhibitory control and attentional shifting tasks were related to performance on a mathematical assessment.

Working memory has also been shown to predict mathematics and reading in early childhood (Bull, Espy, & Wiebe, 2008; Gathercole & Pickering, 2000). Bull et al. (2008) showed that working memory in 4 year olds uniquely predicted mathematics three years later, while controlling for reading ability and other executive functions. Additionally, short-term memory in 4 year olds uniquely predicted reading three years later, while controlling for mathematic ability and other executive functions. In general, there appears to be considerable support for cognitive aspects of executive function predicting both mathematics and verbal abilities in early childhood, but it is unclear which tasks do the best job of predicting each area of achievement. On the one hand,
some studies support specific aspects of executive function (e.g., inhibition tasks) as potentially being the strongest predictor of early achievement (Blair & Razza, 2007; Bull & Scerif, 2001; Espy et al., 2004). On the other hand, there is evidence that tasks requiring the integration of executive function skills and a behavioral response component may be the best predictors of children’s early achievement (Ponitz et al., 2009).

Tasks that tap behavioral aspects of executive function are strong predictors of achievement outcomes (McClelland et al., 2007; Ponitz et al., 2009). Although the tasks may tap some of the same underlying cognitive abilities, the behavioral component may become more important for school success as children enter classroom settings. McClelland et al. (2007) found that behavioral aspects of executive functions were predictive of gains in emergent mathematics, literacy, and vocabulary over the preschool year. These findings suggest that not only do these behavioral executive function skills relate to how well a child is able to do on achievement measures, but also how well the child is able to learn in the school environment over the course of the year. In a similar study by Ponitz et al. (2009), the authors found that higher scores on the HTKS predicted higher scores in mathematics, literacy, and vocabulary in kindergarteners. Additionally, those with higher initial HTKS scores entering kindergarten showed greater gains in mathematics throughout the year.

Behavioral aspects of executive functions may be more predictive of growth in academics than traditional executive function tasks because they incorporate the behavioral components necessary for successful participation in the classroom. It may also be the case that the cognitive aspects of executive function are the main influence on
behavior, with cognitive and behavioral executive function tasks predicting achievement equally well. The present study explored which tasks are the strongest predictor of achievement measures in preschool.

**Teacher Ratings of Behavior and Relations to Executive Function Tasks**

In addition to direct assessments of executive function, teacher ratings of behavior are important sources of information. If the direct assessments are validly measuring a child’s executive function, the teacher ratings of the child’s classroom behavior regulation (involving inhibitory control and attention) should be consistent with the findings from these tasks. One common measure for teacher ratings is the Child Behavioral Rating Scale (CBRS). Studies show that teacher ratings are significantly related to a child’s early achievement and executive functions (McClelland et al., 2006; Ponitz et al., 2009). Although teacher ratings can involve bias compared to more objective measures, they may do a better job of identifying a child’s broader ability to inhibit and attend based on a number of classroom experiences. The current study examined how teacher ratings of classroom behavior are related to each of the direct assessments and to children’s early achievement.

**Convergent and Predictive Validity of Executive Function Tasks as a Function of Head Start Status**

In addition to looking at convergent and predictive validity of executive function for all children, recent research suggests there may be differences in children’s executive function based on demographic risk factors (Rhoades et al., 2011; Sektnan, McClelland, Acock, & Morrison, 2010). For instance, low-income populations may be at increased risk for lower executive functions because of stress effects on the biological development.
of the brain (Blair, 2010; Kishiyama et al., 2009). Higher stress may cause increased neural reactivity in children and without the necessary supportive resources (i.e., high quality childcare or home environment) this may lead to lower executive functions (Blair, 2010). In a study by Sektnan et al. (2010), children with a number of family risk factors (being chronically low-income, of minority status, and having mothers with low maternal education) were more likely to have significantly lower rated behavioral executive function in preschool and kindergarten. Moreover, low maternal education predicted lower scores on behavioral regulation (at 54 months), kindergarten, and academic achievement at the end of first grade. Additionally, Wanless et al. (2011) found that lower family income and being an English-language learner were related to lower scores of behavioral regulation in preschoolers.

The current study builds on previous research by exploring if convergent and predictive validity of executive functions differ by Head Start status. Some recent studies have not found support for differences in factor loadings or item functioning of executive function tasks for children based on income status or maternal education (Wiebe et al., 2008; Willoughby, Wirth, Blair, 2010; Willoughby, Blair, Wirth, & Greenberg, 2011), suggesting the measures should be similarly related in Head Start and non Head Start children. Rhoades et al. (2011), however, found that demographic characteristics differentially influenced mean performance on an inhibitory control measure (but not attention shifting or working memory), suggesting possible inconsistencies in the construct of executive function across different populations. Specifically, children’s performance on an inhibition task was significantly different between low risk and poor African American children, but this difference was not significant for White children.
Although not directly related to the aims of the current study, these results suggest a need to further explore if executive function measures are converging consistently across different populations.

For predictive validity, no study to the author’s knowledge has explored differences by income or Head Start status. If there are differences in predictive validity of executive function tasks, they may be related to demand characteristics and the variability of the assessments. If children from Head Start perform worse on the executive function tasks (Howse et al., 2003; Rhoades et al., 2011; Sektnan et al., 2010; Wanless et al., 2011), more challenging assessments may be less sensitive in differentiating ability. Furthermore, less challenging assessments may be less sensitive in differentiating ability for the non Head Start children. A second possibility is that executive function relates differently to achievement for Head Start children, but no research has directly examined this. If the construct of executive function varies across different populations, it would influence both the convergent and predictive validity of different tasks and aspects of executive function. Thus, the current study explored if the three direct assessments and teacher ratings were similarly related and predicted achievement equally well for Head Start and non Head Start children.

**Goals of the Present Study**

The current study has three main goals: *First, the study will assess the convergent validity of three direct assessments of executive function (the DCCS, the Day-Night Stroop, the HTKS) and teacher ratings (the CBRS).* Examining the relationships between cognitive and behavioral aspects of executive function measures will provide information on the degree to which the measures tap a common underlying construct. All three tasks
tap some of the same cognitive components, yet each is unique. For example, the Day-Night Stroop is primarily considered an inhibition task, the DCCS is primarily an attentional shifting task, the HTKS has a behavioral component, and the CBRS assesses children’s classroom behavioral regulation (Garon et al., 2008; Ponitz et al., 2009). Based on previous research, it is hypothesized that the correlations among the measures will be low-to-moderate, suggesting both an underlying unifying construct but also differing aspects of executive function in preschool aged children (Allan & Lonigan, 2011; Espy et al., 2004; Hughes & Ensor, 2011; Wiebe et al., 2008).

Second, the study will measure the predictive validity of three direct assessments of executive function (the DCCS, the Day-Night Stroop, the HTKS) and teacher ratings (the CBRS) for emergent mathematics, vocabulary, and literacy. By assessing which task is the most predictive of early academic achievement, researchers will be further informed on the skills that may be most important for a child’s early mathematics, vocabulary, and literacy development. Prior research suggests that all measures will be closely related to emergent mathematics (Bull & Scerif, 2001; Bull et al., 2008; Espy et al., 2004; Hongwanishkul et al., 2005; Ponitz et al., 2009). It is hypothesized that the HTKS and CBRS will be the best predictors across all three domains of achievement because they incorporate a behavioral component of executive function that may relate more closely with a child’s classroom behavior and learning-related skills (McClelland et al., 2006; Ponitz et al., 2009).

Finally, the current study will explore the differences between convergent and predictive validity of three direct assessments of executive function (the DCCS, the HTKS, the Day-Night Stroop) and teacher ratings (the CBRS) by Head Start status.
Research suggests that low-income children perform worse on executive function tasks compared to higher income peers (Howse et al., 2003; Rhoades et al., 2011; Sektnan et al., 2010; Wanless et al., 2011). It is unclear, however, if the convergent and predictive validity of these tasks are different depending on Head Start status. Although there is some support for possible differences in the construct of executive function across income statuses, findings are mixed and tend to suggest more similarities than differences (Rhoades et al., 2011; Wiebe et al., 2008; Willoughby et al., 2011). Therefore, it is hypothesized that the tasks should converge relatively similarly across Head Start status.

For predictive validity, it is unclear if there will be differences between the tasks because of the underlying ability (i.e., inhibition, working memory, attentional shifting) because no research to the author’s knowledge has explored this question. It is possible though that the tasks may vary in predictive validity because of measurement characteristics specific to the tasks. For example, tasks that are more cognitively demanding (i.e. the HTKS) will likely produce less variance in scores of the Head Start children, thus potentially reducing the predictive validity. Conversely, less cognitively demanding tasks (i.e., Day-Night Stroop) will likely produce less variance in scores for the non Head Start children, thus potentially reducing the predictive validity.
Chapter 3. Method

Method

Participants

The study consisted of 100 children (Head Start $N = 51$), with a mean age of 58.81 months ($SD = 4.23$). Participants were recruited from a diverse preschool in a university town in western Oregon. The preschool included classrooms with both Head Start and non Head Start children, a unique characteristic uncommon in Head Start programs. Approximately half of the preschool children ($n = 51$) were enrolled in Head Start, some of which were primarily Spanish speakers ($n = 15$) who received the tasks and assessments in Spanish. None of the non Head Start children were primarily Spanish speakers. Children and families were recruited by sending letters home with an explanation of the study and consent information. Consent forms were also available in Spanish and a Spanish speaker was available for Spanish speaking parents.

Procedure

After attaining consent, each child was tested in 2-3 testing sessions lasting about 10-15 minute each. All measures were counterbalanced for each child to avoid order effects. Trained research assistants administered each session. Spanish speaking children received all of the measures in a Spanish version from a fluent Spanish speaker.

Measures

Seven measures were used in the current study, three direct assessments of executive function, one teacher rating of the child’s classroom behavioral regulation, and three direct assessments of academic outcomes. For the three direct assessments of executive function, two are traditional measures of cognitive executive functions: the
Three-Dimensional Change Card Sort (DCCS) and the Day-Night Stroop task; and one measure of behavioral executive function: the Head-Toes-Knees Shoulders (HTKS). For the teacher rating of the child’s classroom behavioral regulation, the Child Behavior Rating Scale (CBRS) was used. The Woodcock-Johnson III Applied Problems subtest, Woodcock-Johnson III Letter-Word Identification subtest, and Woodcock-Johnson III Picture Vocabulary subtest were the three direct assessments of academic outcomes.

**Direct Assessments of Executive Function**

**Three-Dimensional Change Card Sort (DCCS).** The Three-Dimensional Change Card Sort (DCCS) measures children’s attentional and cognitive flexibility (Frye, Zelazo, and Palfai, 1995; Zelazo, 2006). The task requires children to sort cards based on three dimensions: size (small, medium, large), color (red, yellow, blue), and shape (dog, bird, fish). The game requires the child respond to rules by placing the card correctly in designated boxes. The child must utilize attentional shifting, working memory, and inhibition (to previous rules) to perform the task correctly, but it is mainly considered an attentional set-shifting task (Garon et al., 2008; Hongwanishkul et al., 2005). The task involves 18 trials, with six responses to sorting by shape, six by color, and six by size. The child is given a practice trial with shapes, and then tested six times. For color and size, the child is simply told the new rule and tested six times for each. A total of 18 responses are recorded, correct responses receive 1 point and incorrect responses receive 0 points, for a range of scores from 0-18. Scores have been shown to be reliable and valid in children ages 3-6 years (Hongwanishkul et al., 2005).

**Day-Night Stroop.** The Day-Night Stroop task (Gerstadt et al., 1994) is an inhibitory control task where the child must inhibit the natural response by responding to
a picture of a sun as “night” and a picture of a moon as “day.” The task involves inhibition of the natural response and some degree of working memory to remember the rules throughout the game. The task is measured with 16 trials, the child must say the opposite of what the picture is depicting, and for the first two incorrect responses the child is given a reminder about the rules of the game. No responses and incorrect responses are coded as 0, self-corrected responses are coded as 1, and correct responses as 2, with the range of scores from 0-32. The Day-Night Stroop has been shown to be reliable and valid in preschool aged children (Gerstadt et al., 1994).

**Head-Toes-Knees-Shoulders (HTKS).** The Head-Toes-Knees-Shoulders (HTKS) task assesses behavioral aspects of executive function and is a more complex version of the Head-to-Toes task (Ponitz at al., 2009). The task requires behavioral aspects of executive function by attending to rules in which the child must inhibit a natural response and do the opposite of a command. The task involves the child habituating to up to four commands (e.g. “touch your head”, “touch your toes”, “touch your knees” and “touch your shoulders”), then told he or she is going to play a game and do the opposite of each command. For example, when asked to “touch your toes,” the correct response is to touch the head instead, and when asked to “touch your shoulders,” the correct response is to touch the knees. Each incorrect response is coded as 0, self-corrected response as 1 point, and correct responses as 2 points; with 20 different commands there is a possible range of 0 – 40. The HTKS has been shown to have high inter-rater reliability (Ponitz et al., 2008), to be reliable and valid in different cultures, and to be significantly predictive of academic achievement (Wanless, McClelland, Acock, Chen, & Chen, 2011; Wanless, McClelland, Acock, Ponitz et al., 2011).
Teacher Ratings

Child Behavior Rating Scale (CBRS). The Child Behavior Rating Scale (CBRS) was used to assess children’s classroom behavioral regulation (Bronson et al., 1995). The questionnaire includes 10 items and assesses a child’s behavioral regulation in the classroom using a 5-point Likert scale. The same teachers assessed both Head Start and non Head Start children because of the unique classroom structure. An example item from the CBRS is, “Concentrates when working on a task; is not easily distracted by surrounding activities.” Previous research has shown the CBRS to be a reliable and valid measure for behavioral regulation and factor analyses of the task consistently identifies 10 items as tapping behavioral regulation (Lim, Rodger, & Brown, 2010; Ponitz et al., 2009; Wanless, McClelland, Acock, Ponitz et al., 2011).

Academic Outcomes

Emergent mathematics, literacy, and vocabulary. All children received either the Woodcock-Johnson Psycho-Educational Battery – III Tests of Achievement (WJ – III; Woodcock, McGrew, & Mather, 2001) or the Bateria Woodcock-Muñoz in Spanish (Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005). These achievement batteries measure children’s early mathematics, literacy, and vocabulary skills. Mathematic skills were assessed with the applied problems subtest of the battery that involves understanding quantities, simple calculations, and solving practical problems. Early literacy skills were assessed with the letter-word identification subtest, requiring the child to identify letters and pronounce words, increasing in difficulty (both receptive and expressive). Vocabulary skills were assessed by the vocabulary subtest (including both
receptive and expressive), where the child must point to or name a target picture.
Reliability and validity of these subtests are high (Mather & Woodcock, 2001).
Chapter 4. Results

Results

Analytic Strategy

Data analyses (including correlations, factor analyses, and regressions) were run using Stata 12 (StataCorp., 2011) and Mplus 6 (Muthén & Muthén, 2010). The first two research questions utilized the full sample of preschool children. The first research question examined the convergent validity for the three direct assessments of executive function and teacher ratings of child behavioral regulation, and was assessed using a correlation matrix and one-factor Confirmatory Factor Analysis (CFA). The second research question focused on the predictive validity of the three direct assessments of executive function and teacher ratings of child behavioral regulation, and was assessed using multivariate regressions in Mplus 6. The third research question explored the convergent and predictive validity of the three direct assessments of executive function and teacher ratings of child behavioral regulation by Head Start status (Head Start $N = 51$, non Head Start $N = 49$). The analyses included correlation matrices, one-factor CFAs, and multivariate regressions by Head Start status.

Missing Data. Overall, there was relatively little missing data. Of the original sample of 100 children, child age, Head Start status, and teacher ratings on the CBRS experienced no missing data. Five (3 Head Start, 2 non Head Start) of the 100 children did not participate in any testing sessions and therefore did not receive any direct assessments. For the remaining 95 children, all direct assessments experienced less than 6 percent missing data, which the exception of Applied Problems (8%). Missing data on the direct assessments for children occurred primarily for two reasons. First, children
were routinely absent on days of testing, which prevented them from completing the packet of assessments. The second reason is that there was a coding error. A coding error occurred most frequently in the Applied Problems subtest, resulting in the most missing data of any variable. Full Information Maximum Likelihood (FIML) was used to address the issue of missing data. FIML is shown to provide more accurate and reliable estimates than list-wise deletion and assumes data are missing at random (MAR) (Muthén & Muthén, 2010). The advantage of FIML is that all the data are used to estimate the model and additional auxiliary variables can be used to provide more accurate estimates than list-wise deletion (Acock, 2005).

Descriptive statistics. Descriptive statistics for child age, maternal education, direct assessments, and teacher ratings are shown in Table 1. The counts, means, and standard deviations are shown for the total sample and by Head Start status. Additionally, t-tests are shown for the differences in scores for each assessment by Head Start status. Across all measures, non Head Start children did significantly better than the Head Start children, except on the Day-Night Stroop task. The Day-Night Stroop was also the only task that could not go one full standard deviation above its mean before reaching the ceiling for the full sample.

Convergent validity among three direct assessments (the DCCS, the Day-Night Stroop, the HTKS) and teacher ratings (the CBRS). Correlations and a CFA were examined to answer the first research question on the convergent validity for the DCCS, Day-Night Stroop, HTKS, and CBRS. Table 2 shows the correlations among the three direct assessments of executive function and the teacher ratings. All tasks were found to be significantly related to each other ($\alpha = .01$) and ranged from $r = .30$ to $r = .61$. 

The HTKS and DCCS were the most related to each other and the Day-Night Stroop was the least related to the other assessments.

In addition to a correlation matrix, a one-factor CFA was run to assess convergent validity of the four measures on a single underlying dimension. The CFA constrained the unstandardized factor loadings to be equal for the full sample, however, the intercepts for each task varied across the Head Start and non Head Start children. The intercepts varied across Head Start and non Head Start children because the mean performances on tasks were significantly different (except Day-Night Stroop) between the two groups. The primary interest, however, was in how well the tasks would load onto a latent factor with unstandardized loadings constrained equal between the two groups. All measures significantly loaded onto a single underlying dimension. Consistent with the correlation matrix, the HTKS and DCCS were the most related to the underlying dimension and Day-Night Stroop was the least related. By design, the standardized loadings were allowed to be slightly different between the two groups due to differences in the intercepts for each task (even though the unstandardized loadings were constrained equal). The standardized factor loadings for Head Start children were HTKS $\lambda = .73$ ($p < .001$), DCCS $\lambda = .67$ ($p < .001$), CBRS $\lambda = .54$ ($p < .001$), and Day-Night Stroop $\lambda = .42$ ($p < .001$). The standardized factor loadings for the non Head Start children were HTKS $\lambda = .78$ ($p < .001$), DCCS $\lambda = .76$ ($p < .001$), CBRS $\lambda = .64$ ($p < .001$), and Day-Night Stroop $\lambda = .51$ ($p < .001$). Fit indices suggest that the one-factor CFA fit the data well: $\chi^2(8) = 10.09$, $p = .26$, Comparative Fit Index = .97, and Root Mean Square Error of Approximation = .07. The only modification index given was correlating the errors between the CBRS and Day-Night Stroop in the non Head Start children, however, there
was not a strong theoretical justification for why these tasks would be more correlated in the non Head Start children. Thus, no modifications were made.

**Predictive validity of early achievement for direct assessments (the DCCS, the Day-Night Stroop, the HTKS) and teacher ratings (the CBRS).** For the second research question, regression analyses were run to examine the predictive validity of the DCCS, Day-Night Stroop, HTKS, and CBRS for emergent mathematics, literacy, and vocabulary. Each regression analysis included child age and Head Start status as covariates. Although gender was originally included as a covariate, it was dropped because it was not significant in any model. Twelve regression analyses were run: three for the DCCS, three for Day-Night Stroop, three for the HTKS, and three for the CBRS (see Table 3). Standardized betas and explained variance ($R^2$) were reported in order to make comparisons of effect sizes across measures and their predictability. Results for the entire sample of children showed that although there were differences among tasks for different achievement outcomes, the CBRS and the HTKS were the most consistent predictors across all achievement outcomes ($ps < .01$ for all achievement measures). Additionally, the tasks as a whole were most predictive of applied problems ($\beta$s ranged from $.31 - .51$), followed by picture vocabulary ($\beta$s ranged from $.25 - .47$), and letter-word identification ($\beta$s ranged from $.24 - .40$).

**Convergent validity of direct assessments (the DCCS, the Day-Night Stroop, the HTKS) and teacher ratings (the CBRS) by Head Start status.** The third research question was focused on how relations among the direct assessments and teacher ratings varied as a function of Head Start status (see Table 4). The hypothesis that tasks would converge similarly was only partially supported. In contrast to predictions, two key
differences between non Head Start and Head Start children emerged; first, the tasks were more closely related to each other in the non Head Start children than the Head Start children. Second, the Day-Night Stroop was significantly related to all tasks in the non Head Start children, but not significantly related to any other task in the Head Start children. The strongest correlation for both groups was between the HTKS and the DCCS, and the only other correlation that was significant for both groups was between the DCCS and the CBRS. Not only did the non Head Start children have more significantly related tasks, but the correlations were larger across all tasks. Furthermore, in the Head Start group, all correlations involving the Day-Night Stroop were small and non-significant ($r < .20$); however, in the non Head Start group all of these correlations were significant and moderate in size ($r > .36$). Thus, the Day-Night Stroop seems to be more closely related to other executive function tasks in the non Head Start children than the Head Start children.

When considering the factor analyses, the hypothesis for no differences between the two groups was supported. The CFA for the full sample (first research question) yielded a non-significant chi-square, thus the model did not significantly misfit. In other words, a model where the unstandardized loadings varied by Head Start status was not a significant improvement over the first more parsimonious model. However, the model provides some substantive information worth interpreting. The model with varied unstandardized loadings (not shown) did show a substantial improvement in reproducing the covariance matrices for both groups of children. This finding suggests it is more accurate to model the loadings separately by Head Start status, but not significantly better because of an increased number of estimated parameters.
**Predictive validity of direct assessments (the DCCS, the Day-Night Stroop, the HTKS) and teacher ratings (the CBRS) by Head Start status.** Differences in predictive validity of direct assessments and teacher ratings were also explored by Head Start status (see Table 3). The hypothesis that predictive validity would be different based on task variability and demand characteristics was not supported. For task variability, observed ranges were similar between Head Start and non Head Start children and none of the variances for tasks were significantly different between the two groups. For demand characteristics, the Day-Night Stroop task had the opposite effect from hypothesized. Instead of having a relatively easier task be a weaker predictor in non Head Start children, it was a stronger predictor for non Head Start children, despite the fact it experienced slightly greater ceiling effects. Furthermore, the HTKS and the DCCS followed similar patterns for predicting achievement in both Head Start and non Head Start children, even though the DCCS was a relatively easier task and the HTKS was a relatively harder task for the full sample, based on mean performance and scoring range (see Table 1).

For Head Start children, the CBRS was the most consistent predictor of children’s academic outcomes followed by the DCCS and the HTKS. The Day-Night Stroop task was not significantly predictive of any achievement outcome for Head Start children. Conversely, for non Head Start children, the Day-Night Stroop task was the most consistent predictor followed by the CBRS. Overall, the HTKS and the DCCS followed similar patterns of predictability in both groups of children, while the predictability of the CBRS and the Day-Night Stroop appeared to depend on whether the child was in Head Start or not. Analyses were also run as interactions between Head Start status and
executive function assessment (results not shown). These analyses revealed that the interactions between Day-Night Stroop and Head Start status were significant for the applied problems and picture vocabulary regressions (not letter-word identification). None of the other interactions between executive function task and Head Start status were significant for any other achievement outcome.
Table 1

*Descriptive Statistics for the Full Sample, Head Start, and Non Head Start Samples*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full Sample</th>
<th>Head Start</th>
<th>Non Head Start</th>
<th>T-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>Age (in months)</td>
<td>100</td>
<td>58.8</td>
<td>4.23</td>
<td>51</td>
</tr>
<tr>
<td>Maternal Education</td>
<td>76</td>
<td>14.3</td>
<td>4.06</td>
<td>33</td>
</tr>
<tr>
<td>Executive Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HTKS</td>
<td>92</td>
<td>15.3</td>
<td>14.1</td>
<td>47</td>
</tr>
<tr>
<td>DCCS</td>
<td>90</td>
<td>13.9</td>
<td>3.57</td>
<td>43</td>
</tr>
<tr>
<td>Day-N</td>
<td>93</td>
<td>24.2</td>
<td>8.83</td>
<td>47</td>
</tr>
<tr>
<td>CBRS</td>
<td>100</td>
<td>3.89</td>
<td>0.83</td>
<td>51</td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Problems</td>
<td>87</td>
<td>413.5</td>
<td>23.3</td>
<td>42</td>
</tr>
<tr>
<td>Letter-Word</td>
<td>95</td>
<td>346.1</td>
<td>35.1</td>
<td>48</td>
</tr>
<tr>
<td>Picture Vocabulary</td>
<td>93</td>
<td>467.1</td>
<td>15.2</td>
<td>48</td>
</tr>
</tbody>
</table>

Note.  HTKS is the Head-Toes-Knees Shoulders.  DCCS is the Dimension Change Card Sort.  Day-N is the Day-Night Stroop.  CBRS is Child Behavioral Rating Scale.  All t-tests are comparing mean performance between Head Start and non Head Start children.

*p < .05.  **p < .01.  ***p < .001.
Table 2

*Correlations Among Three Executive Function Tasks and Teacher Ratings (N = 88-93)*

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. HTKS</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. DCCS</td>
<td>.61***</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Day-N</td>
<td>.33**</td>
<td>.30**</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>4. CBRS</td>
<td>.40***</td>
<td>.44***</td>
<td>.36***</td>
<td>–</td>
</tr>
</tbody>
</table>

Note. HTKS is the Head-Toes-Knees Shoulders. DCCS is the Dimension Change Card Sort. Day-N is the Day-Night Stroop. CBRS is Child Behavioral Rating Scale. *p < .05. **p < .01. ***p < .001.
Table 3

Regressions for Applied Problems, Letter-Word, and Picture Vocabulary

<table>
<thead>
<tr>
<th></th>
<th>Full Sample of Pre-K (N = 100)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Applied Problems</td>
<td>Letter-Word</td>
<td>Picture Vocabulary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>β</td>
<td>R²</td>
</tr>
<tr>
<td>HTKS</td>
<td>.65</td>
<td>.16</td>
<td>.40***</td>
<td>.35</td>
</tr>
<tr>
<td>DCCS</td>
<td>3.06</td>
<td>.66</td>
<td>.47***</td>
<td>.38</td>
</tr>
<tr>
<td>Day-N</td>
<td>.92</td>
<td>.25</td>
<td>.35***</td>
<td>.32</td>
</tr>
<tr>
<td>CBRS</td>
<td>14.41</td>
<td>2.56</td>
<td>.51***</td>
<td>.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Head Start Pre-K (N = 51)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Applied Problems</td>
<td>Letter-Word</td>
<td>Picture Vocabulary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>β</td>
<td>R²</td>
</tr>
<tr>
<td>HTKS</td>
<td>.76</td>
<td>.25</td>
<td>.44**</td>
<td>.19</td>
</tr>
<tr>
<td>DCCS</td>
<td>3.19</td>
<td>.98</td>
<td>.48***</td>
<td>.23</td>
</tr>
<tr>
<td>Day-N</td>
<td>.61</td>
<td>.39</td>
<td>.25</td>
<td>.07</td>
</tr>
<tr>
<td>CBRS</td>
<td>17.20</td>
<td>3.26</td>
<td>.64***</td>
<td>.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Non Head Start Pre-K (N = 49)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Applied Problems</td>
<td>Letter-Word</td>
<td>Picture Vocabulary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>β</td>
<td>R²</td>
</tr>
<tr>
<td>HTKS</td>
<td>.56</td>
<td>.19</td>
<td>.39**</td>
<td>.28</td>
</tr>
<tr>
<td>DCCS</td>
<td>2.98</td>
<td>.87</td>
<td>.45***</td>
<td>.34</td>
</tr>
<tr>
<td>Day-N</td>
<td>1.35</td>
<td>.29</td>
<td>.56***</td>
<td>.42</td>
</tr>
<tr>
<td>CBRS</td>
<td>10.17</td>
<td>4.02</td>
<td>.39**</td>
<td>.27</td>
</tr>
</tbody>
</table>

Note. HTKS is the Head-Toes-Knees Shoulders. DCCS is the Dimension Change Card Sort. Day-N is the Day-Night Stroop. CBRS is Child Behavioral Rating Scale. Full sample regressions controlled for age and Head Start status. Head Start and Non Head Start regressions controlled for age only.

†p < .10. *p < .05. **p < .01. ***p < .001.
Table 4

*Correlations Among Three Executive Function Tasks and Teacher Ratings by Head Start Status*

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. HTKS</td>
<td>-</td>
<td>.47**</td>
<td>.14</td>
<td>.26†</td>
</tr>
<tr>
<td>2. DCCS</td>
<td>.66***</td>
<td>-</td>
<td>.17</td>
<td>.39**</td>
</tr>
<tr>
<td>3. Day-N</td>
<td>.48**</td>
<td>.36*</td>
<td>-</td>
<td>.18</td>
</tr>
<tr>
<td>4. CBRS</td>
<td>.49***</td>
<td>.47***</td>
<td>.56***</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. HTKS is the Head-Toes-Knees Shoulders. DCCS is the Dimension Change Card Sort. Day-N is the Day-Night Stroop. CBRS is Child Behavioral Rating Scale. Head Start Children (N=51) are Across the Top Diagonal and Non Head Start Children (N=49) are Across the Bottom Diagonal.

†p < .10. *p < .05. **p < .01. ***p < .001.
Chapter 5. Discussion

Discussion

The current study had three overall goals: The first goal was to assess the convergent validity between three direct assessments of executive function and teacher ratings of child classroom behavioral regulation in a sample of prekindergarten children. Results showed that all tasks were positively and significantly related, and loaded onto a single underlying dimension. The second goal was to measure the predictive validity of each of the direct assessments and teacher ratings for early academic outcomes. All tasks were found to be significant predictors of early academic achievement; however, teacher ratings were the most consistent predictor across all areas of achievement and all direct assessments were strong predictors of emergent mathematics. Finally, the third goal was to explore whether there were differences for convergent and predictive validity by Head Start status. In terms of convergent validity, the tasks were more highly correlated to each other in the non Head Start children than in the Head Start children. In terms of predictive validity, inhibition (measured by the Day-Night Stroop) appeared to be a stronger predictor of achievement outcomes in non Head Start children compared to Head Start children, while other aspects of executive function were relatively strong, consistent predictors for both groups.

Convergent Validity among Three Direct Assessments (DCCS, Day-Night Stroop, HTKS) and Teacher Ratings (CBRS).

For the convergent validity of the direct assessments and teacher ratings, all correlations were found to be positive, low-to-moderate, and significant, with the HTKS and the DCCS the most closely related. These findings are consistent with our
hypotheses and prior literature suggesting correlations in the low-to-moderate range (Blair & Razza, 2007; Hongwanishkul et al., 2005; Wiebe et al., 2008). One unique contribution of the present study is that few prior studies have looked at how behavioral and cognitive executive function tasks relate. The correlation between the HTKS and the DCCS was higher than many previously reported values between executive function tasks (Blair & Razza, 2007; Hongwanishkul et al., 2005; Wiebe et al., 2008), which suggest that these tasks require similar underlying executive function skills. Despite a relatively high correlation between the HTKS and the DCCS, there was still a large amount of unshared variance between the tasks. One possible reason for the uniqueness is that the HTKS has a behavioral component (in addition to attentional shifting, working memory, and inhibition) and the DCCS primarily requires attentional shifting (Garon et al., 2008; Ponitz et al., 2009).

The Day-Night Stroop, an auditory inhibition measure, was most closely related with teacher ratings, but its relation with teacher ratings was still smaller than the correlations between the HTKS or the DCCS and the teacher ratings for the full sample of children. This finding suggests that the Day-Night Stroop is more closely related to teacher ratings of child behavioral regulation than it is with the other two direct assessments. Teacher ratings were related moderately and consistently across all direct assessments, which may be because teachers have a better global understanding of the child’s broader ability to attend to and inhibit behaviors (McClelland et al., 2006). Consistent with recent literature and the correlation analysis, factor loadings support the notion of a single underlying dimension (Allan & Lonigan, 2011; Hughes & Ensor, 2011; Wiebe et al., 2008). Although each task and teacher ratings tap unique aspects of
executive function, together they load moderately well onto a single underlying executive function dimension.

**Predictive Validity of Early Achievement for Direct Assessments (DCCS, Day-Night Stroop, HTKS) and Teacher Ratings (CBRS).**

Consistent with our hypotheses, regression analyses with the full sample of prekindergarten children showed that the CBRS and the HTKS were the most consistent predictors across achievement, with all tasks significantly predicting emergent mathematics. The findings are consistent with prior literature showing executive function as a strong predictor of early academic achievement, especially mathematics (Blair & Razza, 2007; Bull & Scerif, 2001; Hongwanishkul et al., 2005; Ponitz et al., 2009). One unique contribution of the current study is that it compared effect sizes across cognitive and behavioral executive functions, and teacher ratings, while including child age and Head Start status as covariates. The teacher ratings of child behavior (the CBRS) were found generally to be the most consistent predictor of academic achievement, followed by the HTKS and the DCCS, and the weakest predictor was the Day-Night Stroop. It is possible that the CBRS and the HTKS were the most consistent predictors because both involved a behavioral component that may more closely mimic children’s behavior in learning environments. Moreover, behavioral regulation has been found to be a strong predictor of growth in achievement (McClelland et al., 2007; Ponitz et al., 2009). Additionally, the DCCS was found to be a strong predictor of emergent mathematics and vocabulary, which is consistent with prior studies showing card sorting tasks relating to early mathematics and vocabulary (Hongwanishkul et al., 2005).
The Day-Night Stroop was the weakest predictor overall, although it was still highly significant for predicting emergent mathematics and literacy. One explanation for the weakness in predictability may be that it seemed to be less cognitively demanding and had less variability compared to some of the other direct assessments. For example, it was the only measure in the current study that could not go one full standard deviation above its mean before reaching the maximum score. This explanation, however, runs counter to the results found for the differences between Head Start and non Head Start children.

**Convergent Validity of Direct Assessments (DCCS, Day-Night Stroop, HTKS) and Teacher Ratings (CBRS) by Head Start Status.**

The current study unexpectedly found that the three direct assessments and teacher ratings were more related in the non Head Start children than in the Head Start children, suggesting some inconsistencies in convergent validity across the two groups. Lower socioeconomic status has been shown to be a risk factor for lower scores on executive function tasks (Howse et al., 2003; Rhoades et al., 2011; Sektnan et al., 2010; Wanless et al., 2011); however, little research has focused on how validity of tasks may be different by socioeconomic status (Wiebe et al., 2008; Willoughby, Blair et al., 2011). Although substantial differences were observed in the correlation matrices, non-significant differences for factor analyses are consistent with the hypothesis and prior research (Wiebe et al., 2008). The substantial differences still merit noting though, as significance is highly influenced by sample size. In the Head Start children, only two correlations were significant, the HTKS and the DCCS, and the DCCS and the CBRS.
Conversely, in the non Head Start children all tasks were significantly related to each other.

The Day-Night Stroop showed a large difference in relations to other tasks between the groups, which may be related to the fact it was the only task that did not have a significant difference in mean performance between the groups. The Day-Night Stroop was not significantly related to any other task in the Head Start children, yet was significantly related to all tasks in the non Head Start children. This finding is partially consistent with Rhoades et al. (2011), suggesting a differential relation between inhibition and other executive functions depending on income status, which could mean aspects of executive function develop differentially based on demographic characteristics. Despite notable differences in the Day-Night Stroop, the full sample CFA fit the data well, suggesting consistency in the convergent validity across both groups. Although a CFA that allowed varied loadings improved the chi-square, it could likely be due to chance alone and a larger sample would be needed to more accurately test these models. Additionally, the CFA with varied loadings improved the ability to reproduce the covariance matrices (more accurate), suggesting future studies should continue to explore discrepancies in latent variable modeling of executive function aspects (specifically inhibition) by income status.

Predictive Validity of Direct Assessments (DCCS, Day-Night Stroop, HTKS) and Teacher Ratings (CBRS) by Head Start Status.

The predictive validity of the three direct assessments and teacher ratings varied depending on whether the child was in Head Start, although not in the hypothesized way because of task variability and demand characteristic. As previously mentioned, the most
notable change in predictive validity was for the Day-Night Stroop. The Day-Night Stroop, an inhibition measure, was not significant for predicting any of the academic achievement measures in Head Start children, but was one of the strongest predictors for the non Head Start children, in spite of a slightly greater ceiling effect in the non Head Start children. This finding suggests that inhibition may be more closely related to academic achievement in non Head Start children compared to Head Start children. Environmental effects associated with Head Start and non Head Start children could contribute to differential development for aspects of executive function, which would likely also influence how the aspects of executive function relate to achievement. Simple inhibition may develop at similar rates in Head Start and non Head Start children, as suggested by the non-significant mean differences in Day-Night Stroop performance, but the other aspects of executive function (attentional shifting and working memory) might be more important for Head Start children when predicting achievement. These conclusions are speculative and need to be validated in a larger sample, presumably with multiple measures that assess different forms of inhibition.

The other task that had a large change between Head Start and non Head Start children in predictive validity was the DCCS, but only for predicting emergent literacy. The DCCS, however, did not vary in a systematic way across all achievement outcomes like the Day-Night Stroop, suggesting its difference was not related to the underlying ability of attentional shifting. Previous research has focused on how demographic risk factors relate to children doing worse on executive function tasks (Howse et al., 2003; Rhoades et al., 2011; Sektnan et al., 2010; Wanless et al., 2011), but no study to the author’s knowledge have explicitly explored if tasks have different predictability by
income status. If children with risk factors are doing worse on executive function tasks, it is imperative that researchers understand how the specific aspects of executive function may differentially relate to the global ability and to academic achievement. In understanding the connections between aspects of executive function and achievement, interventions can potentially be targeted towards skills that are more related to academic achievement, leading to greater effectiveness. Specifically, the current findings suggest that targeting attention related tasks, more so than inhibition, might be more effective in low-income children. The findings need to be interpreted cautiously because of a relatively small sample size, yet future studies should continue to explore if aspects of executive function are differentially related and predict achievement differently by income status.

**Implications**

The findings in this study have a number of important implications for research related to executive function, school readiness/interventions, and differences associated with income status. First, the dimensionality of executive function and relations between tasks has been a focus of recent research in early childhood (Allan & Lonigan, 2011; Hughes & Ensor, 2011; Monette et al., 2011; Wiebe et al., 2008). The current study found support for a single dimension in cognitive and behavior executive functions, and teacher ratings of child classroom behavior, suggesting executive functions are accurately modeled as a single latent variable. Although support for a single dimension was found, the current study also highlighted differences in convergent and predictive validity for the different tasks. The findings suggest that tasks are unique in their contributions toward an underlying construct and toward predicting achievement. Therefore, executive
functions may not always be best represented by a composite or latent variable when trying to understand how specific cognitive skills relate to early academic achievement. Moreover, the Day-Night Stroop, an inhibition measure, showed very different patterns in both its convergent and predictive validity across Head Start and non Head Start children. This finding suggests that there may be differences in executive function related to inhibition these two groups of children. Measures of inhibition may be more predictive of early academic achievement in children from non low-income households compared to low-income households. Furthermore, these differences might be because of environmental influences that cause differential development of inhibition related to other aspects of executive function (including working memory and attentional shifting).

Finally, understanding measurement of executive function becomes especially important as a greater number of studies stress the importance of its relation to school readiness and use it as tool for early intervention (Diamond & Blair, 2008; Morrison et al., 2010; McClelland et al., 2010; Raver et al., 2011). If researchers are aiming to improve executive function in early childhood, it is imperative that the specific components are well understood for the target populations. The current study found low-income children struggled more with attention related tasks (the HTKS and the DCCS) than an inhibition task (the Day-Night Stroop), and that the attention related tasks were more predictive of achievement. Therefore, interventions may be more successful if targeting attention related skills versus inhibition in low-income children. The current study contributes to the knowledge of three commonly used direct assessments and teacher ratings related to executive function literature in preschool aged children.

Limitations
The study has two important limitations worth noting. First, the analyses are all cross-sectional, and therefore, growth in achievement is not being predicted, only concurrent performance. It is unclear if the tasks used in the current study would predict growth of skills differently than when predicting achievement concurrently. For example, teacher ratings may be a strong predictor of concurrent achievement, but the direct assessments of executive function may be strong predictors of growth in achievement as previously found (Bull et al., 2008; McClelland et al., 2007; Ponitz et al., 2009). A second limitation is the relatively small sample size of Head Start and non Head Start children for comparing convergent and predictive validity. Both samples were roughly 50 children, and although sufficient for exploratory analyses of differences in convergent and predictive validity, true differences would be better estimated with a larger sample. Despite the limitations, the current study makes a number of contributions in understanding measurement of executive function in early childhood.
Chapter 6. Conclusion

Conclusion

The current study adds to the literature in several important ways. First, the study contributes to a greater understanding of the relations between commonly used direct assessments of cognitive and behavioral executive functions and teacher ratings. Although the importance of executive function has been well documented, understanding the measures of executive function should be of equal focus in research. Second, the current study contributes to the understanding of predictive validity of executive function tasks. The tasks that are more predictive of different domains of academic achievement could potentially inform researchers that use interventions of which specific aspects of executive function are most important for early achievement. Third, by exploring differences associated with income status this study finds that not only are low-income children scoring lower on executive function tasks, but that there may be some important differences in the convergent and predictive validity of executive function. Although these findings are predominantly exploratory, they warrant greater attention moving forward and may have implications in both addressing school readiness in low-income children and executive function interventions. In general, these findings contribute to the current literature on executive function, behavioral regulation, and academic achievement during the preschool years.
References


http://www.developingchild.harvard.edu


StataCorp. 2011. *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.


