

DEVELOPMENT OF A SCREENER FOR THE BEHAVIORAL ASSESSMENT OF
EXECUTIVE FUNCTIONS IN CHILDREN

by

MAURICIO ALEJANDRO GARCIA-BARRERA

(Under the Direction of Randy W. Kamphaus)

ABSTRACT

The problem of valid measurement of psychological constructs remains an impediment to scientific progress, and the measurement of executive function in children is not an exception. “*Executive Functions*” is a multidimensional construct that has been controversial since its first descriptions, and this lack of consensus resulted in the development of multiple evaluation models. Furthermore, increased awareness of the importance of executive functions in childhood has guided researchers to create more sophisticated measurement techniques of this complex function. However, questions about the ecological validity of traditional tests remain, and improvements have been observed with the introduction of executive functioning rating scales. The purpose of this dissertation was the development of a behavioral screener for the estimation of executive functions in children.

Therefore, a CFA model with 25 items loading into four latent factors representing four executive functions was developed. The factors were labeled as behavioral control, emotional control, attentional control, and problem solving, and its statistical properties were examined using Structural Equation Modeling. Each factor corresponds to a cortical representation in the prefrontal cortex and its connections. The sample was obtained from the original standardization

sample of the BASC Teachers for children aged 6-11 (N=2165). The items were derived from the original Reynolds and Kamphaus' Behavior Assessment System for Children ages 6-11. Reliability analysis demonstrated moderate to high factor internal consistency. Analysis of content validity (panel of experts) eliminated construct irrelevant indicators. Construct validity demonstrated the multidimensionality of the model and its adequate fit (CFI= 0.948). Measurement Invariance analysis across sex and age demonstrated that the model was invariant.

These results supported the hypothesis that a reliable and valid executive functions measure could be obtained from a behavioral rating scale, and that its properties can be reliably tested using CFA and SEM methods. Furthermore, screeners are useful clinical tools in assessment settings. If implemented, this screener will provide clinicians and researchers with an instrument for the estimation of executive functions in children. Finally, these results suggested that future investigations should be oriented toward achieving a better understanding of the behavioral indicators of executive functioning, especially in children and adolescent populations.

INDEX WORDS: Executive functions, Prefrontal cortex, Screener, Rating scale, Ecological validity, Modeling, Measurement development, Behavioral Assessment.

DEVELOPMENT OF A SCREENER FOR THE BEHAVIORAL ASSESSMENT OF
EXECUTIVE FUNCTIONS IN CHILDREN

by

MAURICIO ALEJANDRO GARCIA-BARRERA

B.A., Universidad de Antioquia, Colombia, 1999

M.Ed., The University of Georgia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2008

©2008

Mauricio Alejandro Garcia-Barrera

All Rights Reserved

DEVELOPMENT OF A SCREENER FOR THE BEHAVIORAL ASSESSMENT OF
EXECUTIVE FUNCTIONS IN CHILDREN

by

MAURICIO ALEJANDRO GARCIA-BARRERA

Major Professor: Randy W. Kamphaus

Committee: Jonathan Campbell
Morris Cohen
George W. Hynd
L. Stephen Miller

Electronic Version Approved

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2008

DEDICATION

To my family

ACKNOWLEDGEMENTS

I would like to express my gratitude to Randy W. Kamphaus for his extraordinary support to my career. His dedication to mentoring international students, his passion for meaningful and outstanding research, his knowledge of the field of school psychology, and his personal ethics and standards have provided me with a strong model to follow when facing the upcoming milestones of my career in academia. This dissertation was possible thanks to his close and dedicated mentorship and inspiration.

I would also like to thank George W. Hynd for his outstanding support and mentorship during my research fellowship in pediatric neuropsychology at the Center for Clinical and Developmental Neuropsychology. His passion for neuropsychology was inspirational, and his talent as a researcher remarkable. I appreciated very much his input in this work.

I especially thank Jonathan Campbell, Morris Cohen, and Stephen Miller, members of my doctoral committee whose guidance, time, and mentorship guided me throughout the prospectus, final defense, and the arduous career decision making process that followed.

This is an opportunity to thank my family, to whom I dedicate this work and my career. I am a product of their efforts and dreams, their unconditional support, and their ability to make decisions during my early years that guided me in the right direction. I would also to express my gratitude to my friends Jason Davidow, Colby Butzon, and Juliana Sanchez Bloom, whose support and friendship made everyday of work and recreation a very fulfilling one. I have special gratitude to Stephen Owen who not only took the time to proof read and edit my dissertation but also provided me with love, care, and support during the graduate school years.

The difference between the child and adult resides in the unfolding of executive functions.

Martha Bridge Denckla (1996b)

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES	ix
LIST OF FIGURES	xi
CHAPTER	
1 INTRODUCTION	1
2 LITERATURE REVIEW	7
Criteria for Inclusion and Exclusion	7
Historical Background	8
Executive Function: a Concept	11
The Unit versus Network or “ <i>Neuroanatomical Problem</i> ” of Executive Functions	13
The Unity versus Diversity or “ <i>Conceptualization Problem</i> ” of Executive Functions	20
The “ <i>Measurement Problem</i> ”	40
Conclusions	54
3 METHODS	57
Original Instrument	57
Participants and Data Collection	57
Screener Derivation	60

4	RESULTS	75
	Phases I to IV: Data and Item-Level Screening	75
	Phase V: Reliability and Validity Analysis	77
5	DISCUSSION	93
	Data and Item Screening	93
	Latent Construct Analysis and Operationalization	94
	Model Modification and Item Analysis	102
	Unidimensionality versus Multidimensionality	103
	Measurement Invariance and Construct Validity	105
6	CONCLUSIONS	107
	Summary of Findings	107
	Clinical Implications, Limitations, and Future Research	109
	REFERENCES	112

LIST OF TABLES

	Pages
Table 1: Summary of traditional neuropsychological tests used for the assessment of cognitive executive functions	44
Table 2: Comparison across three well known frontal lobe/executive functions rating scales ...	48
Table 3: Demographic representation of the general-norm sample (%)	59
Table 4: General norm sample by special education classification and gender (%)	59
Table 5: Demographic representation of the clinical-norm sample, $N=393$ (%)	59
Table 6: Representation of the clinical norm sample by primary diagnosis or classification (%)	59
Table 7: Representation of the aggregated sample by developmental age (%)	59
Table 8: Original distribution of items per scale	63
Table 9: Fit indexes and cut-off scores of goodness of fit examination	72
Table 10: Data distribution and correlation matrix	76
Table 11: Distribution of items per scale after content validity analysis	77
Table 12: Internal consistency reliability (Cronbach's alpha coefficients by scale)	78
Table 13: Fit indexes for the unidimensional versus multidimensional model of executive functions.....	80
Table 14: Exploratory analysis following modification index recommendations	82
Table 15: Factor loadings	84
Table 16: Model fit for each sub-sample (by Sex)	87

Table 17: Standardized factor loadings for aggregated, only female, and only male sub-samples	87
Table 18: Measurement invariance for sex groups	90
Table 19: Model fit for each sub-sample (by age)	90
Table 20: Standardized factor loadings for aggregated (6-11), young (6-8), and older (9-11) samples	91
Table 21: Measurement invariance for two age groups	92

LIST OF FIGURES

	Pages
Figure 1: Conceptual model of evidence-based cognitive taxonomy of executive functions	33
Figure 2: Conceptual map of a hypothetical/theoretical-based model of the behavioral taxonomy of executive functions.....	34
Figure 3: Conceptual map of an evidence-based model of self-regulatory components of executive functions	39
Figure 4: Screener derivation process model	60
Figure 5: Hypothetical CFA model for the BASC Executive Functions Screener	69
Figure 6: Model used by the WLSMV estimator under LVR formulation for ordinal data	71
Figure 7: Unidimensional model 1	80
Figure 8: Multidimensional model 2	80
Figure 9: Multidimensional model 3	81
Figure 10: Multidimensional model 4	81
Figure 11: Test for model modifications	83
Figure 12: Final CFA model for the BASC Executive Functions Screener.....	86

CHAPTER 1

INTRODUCTION

“Dr. James, you are not alone in having invested many hours in research producing uninterpretable results because of unreliable measures. In fact I would go so far as to say that difficulty in measuring the phenomena of interest in the behavioral sciences goes a long way toward explaining why we live in a world so technologically advanced and so behaviorally primitive.” (Strayhorn, 1993, p. 1303)

The problem of valid measurement of psychological constructs remains an impediment to scientific progress (Strayhorn, 1993), including the measurement of executive function in child psychology. This multidimensional ability has been controversial since its first descriptions in early neuropsychological literature, and multiple models have been developed in an attempt to understand the different components of such regulatory system (Elliott, 2003; Miyake, Friedman et al., 2000; Tirapu-Ustarroz, Munoz-Cespedes, & Pelegrin-Valero, 2002). The famous case of Phineas Gage created awareness of the importance of frontal lobe areas in the executive control of behavior (Harlow, 1868, 1999; Miyake, Emerson, & Friedman, 2000). A. R. Luria introduced the first clinical descriptions of brain injuries affecting the system responsible for planning, monitoring, and executing behaviors oriented towards a goal (Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Elliott, 2003; Lezak, Howieson, & Loring, 2004; Luria, 1973; Mahone, Zabel, Levey, Verda, & Kinsman, 2002; Royall et al., 2002). The initial interest in executive dysfunction, related to structural brain damage, has recently shifted from morphological to functional and from a focus on adult psychopathology to that of children's. For instance, systematic research originated in studies of executive functioning in adult

psychopathology including schizophrenia (Bell & Zito, 2005; Goldman, Axelrod, Tandon, & Berent, 1991; Morice & Delahunty, 1996; Zanello & Huguelet, 2001), obsessive-compulsive disorder (Moritz et al., 2002), bipolar disorder (Meyer et al., 2004), Alzheimer's disease, and other dementias (Grafman & Litvan, 1999; Watkins et al., 2000); however, there is an increasing number of studies on children, including analysis of executive impairments in Autism Spectrum Disorders (Gilotty, Kenworthy, Sirian, Black, & Wagner, 2002; Hughes, Russell, & Robbins, 1994), Attention Deficit Hyperactivity Disorder (Barkley, 1997b, 1998, 2003; Kibby, Cohen, & Hynd, 2002; Lawrence et al., 2002; Oades, 1998; Pineda et al., 1998; Shallice et al., 2002), Developmental Dyslexia (Brosnan et al., 2002; van der Schoot et al., 2004; van der Schoot, Licht, Horsley, & Sergeant, 2002), Conduct Disorder (Clark, Prior, & Kinsella, 2002; Stevens, Kaplan, & Hesselbrock, 2003), and Traumatic Brain injury (Ewing-Cobbs, Prasad, Landry, Kramer, & DeLeon, 2004).

As awareness of the importance of executive functions grows, several efforts have been made to improve the sophistication of measurement techniques of this complex function. Early models in neuropsychology considered the Wisconsin Card Sorting Test (WCST, Heaton, 1981), Towers of Hanoi and London and other variations (TOL, Anzai & Simon, 1979; Shallice, 1982), and Complex Figure of Rey-Osterrieth (Rey & Osterrieth, 1993) as gold-standard tools for assessing executive function (V. Anderson, 2001; Lezak, Howieson, & Loring, 2004). However, criticisms of the limitations on sensitivity and specificity of these and other instruments created the need for a more multidimensional approach (Anderson, 2001; Hughes & Graham, 2002; Manchester, Priestley, & Jackson, 2004). To compensate for this limitation, several models have been proposed. The neuropsychological approach recommends an assessment based on the analysis of lower-order processes believed to be related to components of executive function

(Anderson, 2001; Hughes & Graham, 2002). Another psychometric approach has been focused on the analysis of the everyday-behavioral components of executive functions, and it is recognized for the development of behavioral ratings of frontal and executive function; examples are the Behavior Rating Inventory of Executive Function (BRIEF, Gioia, Isquith, Guy, & Kenworthy, 2000), the Frontal Systems Behavior Scales (FrSBe, Grace & Malloy, 2001), and the rating included on the Behavioural Assessment of the Dysexecutive Syndrome, The Dysexecutive Questionnaire (DEX - BADS, Wilson, Evans, Emslie, Alderman, & Burgess, 1998). A different approach was recently introduced by Manchester and colleagues (2004), who reviewed the shortcomings of the classic methods and recommended focusing on naturalistic observations of behavior.

However, average and unimpaired executive systems at different developmental stages have been generally overlooked and only recently have been the focus of research (Anderson, 2001; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; MacPherson, Phillips, & Della Sala, 2002; Salthouse, Atkinson, & Berish, 2003; Tamm, Menon, & Reiss, 2002). Moreover, research on the assessment of executive function has been criticized for neglecting the inclusion of its analysis on “normal” children populations involve, and it is only recently that a burst of studies of executive functions in preschoolers and school-aged children has started to be published (V. Anderson, 2001; Bialystok & Martin, 2003; Espy, 2004; Espy, Kaufmann, & Glisky, 2001; Espy et al., 2004; Hughes & Graham, 2002; Hughes et al., 1994; Isquith, Gioia, & Espy, 2004; Senn, Espy, & Kaufmann, 2004).

Executive functions such as problem solving, behavioral control, attentional control, and emotional control are important in students’ everyday routine; however, research in the area of executive function has neglected the analysis of these powerful meta-cognitive skills in the

learning and the school contexts (C. Clark, Prior, & Kinsella, 2002). It appears that one limitation of such type of research is the difficulty finding an assessment tool that would serve as a screener of executive function, without the rigor of the neuropsychological testing or the costs of executive function rating systems.

The Behavior Assessment System for Children (BASC, Reynolds & Kamphaus, 1992) is a multidimensional rating scale of externalizing, internalizing, and adaptive skills that includes questionnaires for parents and teachers of children and adolescents between the ages of 2 and 18. The BASC system is among the most widely used measures of child behavior in the United States, where it is ranked as either first or second in usage surveys (Reynolds & Kamphaus, 1992; Reynolds & Kamphaus, 2002; Sandoval & Echandia, 1994). A Colombian version was validated and standardized by Pineda and colleagues (Pineda, Henao et al., 1999; Pineda, Kamphaus, Mora, Puerta et al., 1999; Pineda, Kamphaus, Mora, Restrepo et al., 1999); another has recently been developed and published in Spain, and several versions are under development in Croatia, Mexico, Iceland, and other countries. The validity of the BASC as an assessment tool for frontal lobe/executive function has been previously studied (Jarratt, Riccio, & Siekierski, 2005; Mahone, Zabel, Levey, Verda, & Kinsman, 2002; Riccio et al., 1994). However, only one recent study utilized the Frontal Lobe Functioning/Executive Control scale (FLEC) recently included in the BASC-2 extended software (Sullivan & Riccio, 2006).

This scale was originally developed by Barringer and Reynolds and presented as a paper at the annual meeting of the National Academy of Neuropsychology (1995, presented in Reynolds & Kamphaus, 2002). Barringer and Reynolds identified a subset of items from the BASC. These items were hypothesized as being associated with frontal lobe/executive function using an expert approach by surveying editorial board members of three leading clinical journals

in neuropsychology. A final set of 18 items produced a coefficient Alpha of .84, and comparison of various clinical groups showed high levels of discrimination (Reynolds & Kamphaus, 2002). Although Barringer and Reynolds' analysis included only the Parent Rating Scales, parallel forms of this set of items can be identified across different rating scales from the BASC, including the Teacher Rating Scales for both age groups, children (6-11 years old) and adolescents (12-18 years old) (Reynolds & Kamphaus, 2004).

This project includes the development of a screener for the assessment of executive function, guided by the question: *Is it possible to obtain a reliable and valid measure of children's executive functions with behavioral rating scales?*

In order to answer this question, a screener of executive function was created from a group of items from the BASC. An initial set of 28 items believed to assess executive functions such as problem solving, attentional control, behavioral control, and emotional control was identified and examined for content validity using a panel of experts approach. Pearson Assessment authorized the use of the BASC standardization data for the validity analysis. Reliability analysis was examined using Cronbach's Alpha coefficient for internal consistency. Finally, construct validity analyses included a multidimensional Confirmatory Factor Analysis model created in a Structural Equation Modeling framework. The baseline model included four factors and 25 indicators, and its goodness of fit was examined using Mplus 4.0 (L. K. Muthén & Muthén, 1998-2006). Further construct validity analysis included measurement invariance analysis across different groups (gender and developmental ages).

Results of the analysis, further discussion, conclusions, and results implications are included. Finally, the potential impact of the results of this research includes the development of a reliable and valid screener for the estimation of executive functions that can potentially be used

for the early identification of children in need of academic and/or behavioral interventions, and that can be cross-validated in different cultures and languages, as it would serve as a guideline for the creation of parallel forms of the screener with different BASC datasets.

CHAPTER 2

LITERATURE REVIEW

Criteria for Inclusion and Exclusion

The core of this review is based on a literature search conducted between the spring of 2005 and the spring of 2006. To frame the search, the aim was to examine theoretical, pseudo-experimental, and experimental papers on executive functioning and its components, with emphases on their relationship with academic achievement, intelligence, and specifically, with the assessment in children. The focus was on understanding the state-of-the-art of research in this area, limitations and frequent problems encountered, and areas in need of clarification and new contributions. Additionally, theoretical and methodological guidelines, and baselines for comparisons of research result were also included.

Frequent key words used were “executive function,” “executive functioning,” “executive control,” “prefrontal cortex,” “frontal lobes,” “executive dysfunction,” in combination with “pediatrics,” “children,” “latent variable analysis,” “cognitive development,” “assessment,” “rating scales,” and “neuropsychology,” among others. English and Spanish-language peer-reviewed articles and reviews were considered. Engine search tools were used as well, such as PubMed-MEDLINE, PsycINFO, EBSCOhost, and other electronic databases available at the University System of Georgia (GALILEO). Some articles and books were collected from the Main and Science Libraries at the University of Georgia. Articles were identified and accessed if the topic was of relevance for this review. Publication identification was facilitated by advisors’ recommendations, ground-breaking contributions, reference lists from scholarly manuscripts,

and others. Major contributing authors to the literature on executive function were reviewed as well. Finally, key historical documents were analyzed and included in this literature review.

Given the extent of the literature on executive functions, irrelevant or unrelated material was excluded from this review, including some studies of assessment of executive functions in adults, studies in psychopathology such as schizophrenia, autism, Attention Deficit Hyperactivity Disorder (ADHD), and others. Pilot studies with small samples sizes or ambiguity in their results were excluded. Newspaper documents or communications were generally excluded. Some articles were not found in electronic or paper format; in some cases, if relevant, authors were contacted to obtain these articles. Some articles had promising topics but were published in non-indexed journals, and thus were excluded on the basis of potential poor editorial reviews. Those manuscripts from which only an abstract was obtained were excluded as well. Finally, a few unpublished manuscripts were reviewed but only for guidelines, and therefore were not included on the review.

Historical Background

Interest in studying higher cortical processing, its relation to a specific brain area, and its impairments after neurological damage, can be traced back to the Egyptians (circa 3,500 A.C.) and their early yet sophisticated studies about human behavior and illness. However, the development of a theory of frontal lobe relations to human “higher” cognitive behavior is more recent. Early descriptions of this neuroanatomical and functional relationship include Harlow’s famous case of Phineas Gage and the passage of an iron bar through his head, originally presented in 1848 as a letter to the editor of the *Boston Medical and Surgical Journal*, and recently included in *The Journal of Neuropsychiatry and Clinical Neurosciences* (Harlow, 1999; Neylan, 1999). Mr. Gage was severely injured on his left, and probably right, prefrontal areas,

but was able to walk and talk immediately after the accident, failing to demonstrate signs of behavioral change in the eyes of Harvard's Medical School examiners. Harlow's case was commonly used as an example against Franz Gall's phrenology, which was a mainstream theory at the time. However, Harlow reported the behavioral changes observed in Mr. Gage 20 years later, such as impulsivity, impatience, and irreverence, among other symptoms that interfered with Gage's capacity to perform at work and in personal relationships. These historical notes constituted the earliest description of what became known as "a frontal lobe syndrome."

Furthermore, and as was presented by Royall et al (2002), it was not until World War II and the seminal work of Alexander Romanovich Luria on *The Working Brain* (1973), that clinical observations of behavioral and cognitive changes following frontal lobe injuries were systematically described and analyzed. Luria had the unique opportunity to examine and follow war veterans and injured military personnel. Luria's reports included the description of impaired ability to engage in the "programming, regulation, and verification of behavior" (Royall et al., 2002). Even earlier, Luria's records seemed to include observations about the existence of such a specialized system (Shallice, 1982), which is documented to be the earliest definition of executive function. Tirapu-Ustárroz, Muñoz-Cespedes, and Pelegrin-Valero (2002) noted that Luria was the first to use this concept, but that the term was actually coined by Lezak (1982).

Around the same period, several competing models of memory were being developed under cognitive sciences and information processing perspectives, and theorists agreed upon a clear dichotomy between long-term and short-term memory, which was recognized as the "modal model" given the emphasis on the time mode (i.e., short-long). Examples were Broadbent's (1958, as cited by Baddeley, 1983) and Atkinson and Shiffrin's classical memory models (1968, as cited by Baddeley, 1983). Baddeley and Hitch (1974) created a model of

working memory as an alternative to the “modal model” with an emphasis on processing. This new conceptualization of memory was important for the later development of theories of executive functioning. Their model included three components: first, the articulatory loop (primarily an input store for speech perception, activated through subvocal rehearsal); second, the visuo-spatial scratch-pad (input store responsible for maintenance and manipulation of visuo-spatial images); and third, the *Central Executive*, which functioned as an attentional control for working memory. Although initially vague, the Central Executive was better defined in later publications about the model, given the new theoretical perspectives offered by Norman and Shallice (Baddeley, 1986, 1996). In the early 1980’s, Norman and Shallice published a technical report explaining their model of attentional control over information and execution of behavior. It included a presentation of the role of what they called the *Supervisory Activating System*, an attentional controller that has the ability to override automatic/habitual response patterns when a new behavior is needed. SAS was seen as a specification into an information processing model of Luria’s early ideas of an executive control system (Shallice, 1982). Baddeley stated that the central executive is his version of SAS, and referred to the central executive as a “homunculus-a little man who sits in the head and in some mysterious way makes the important decisions.” This definition, however, created misunderstandings about the existence, pertinence, and neurological correlates of this concept. Examples of the controversial propositions about the “central executive” can be found in Parkin’s communication, “*The central executive does not exist*,” and its subsequent response from Baddeley, “*The central executive: A concept and some misconceptions*” (Baddeley, 1998; Parkin, 1998).

Baddeley (1996) defined some of the “central executive” functions in terms of its capacity to coordinate dual-tasks (such as digits backwards and verbal fluency tasks), ability to

mediate between other slave systems of working memory (e.g., random generation tasks and trail making tests), serving as an attentional controller filtering relevant information from irrelevant (selective attention tasks, such as Continuous Performance Tests), and the ability to select and manipulate information in long-term memory (measures of working memory span such as word-pairs). Some of these functions have been analyzed and replicated by others (e.g., Duff, 2000).

Taken as a group, these early conceptualizations demonstrated the lack of a bridge between the clinical and more neurological-oriented model of “frontal lobe function” and the theoretical and more cognitive-oriented model of “the central executive.” Although the dichotomy remains, newer models of “executive function” have served as an attempt to bridge the gap between theory and practice, and of course, newer debates about its definition have been unavoidable.

Executive Function: a Concept

All the so-called executive functions are phenomena of the processing in neural networks of the frontal lobe. [...] Nonetheless, such constructs are useful heuristically for as long as we do not fully understand the neural mechanisms of frontal function. (Fuster, 1997, p. 218)

One of the obstacles that researchers encounter when performing studies on executive functions is how to define this construct. A product of the diversity of theories is the plethora of terms that have been associated with this complex cognitive function, and an examination of them illustrates the difficulty in accurately label it. For instance, the term “executive function” seems to imply the idea of a singled-unit system, whereas the term “executive functions” would have the opposite effect. The term “executive functioning” emphasizes “processing” more than the function itself; “central executive” places it on a hierarchy-like system, in which there is a central unit and subordinate systems; “executive control” undermines some cognitive

components of the system (e.g., working memory, sequencing, categorization) while focusing on its ability to coordinate or control other systems (e.g., attention, behavioral inhibition, execution of motor programs). It has also received the names “metacognition,” “metacognitive function,” and “metacognitive ability” from the cognitive and educational psychology fields, with an emphasis on its relationship to problem solving, decision making, and learning skills. Finally, it has been historically recognized as “frontal lobe function” or “frontal functioning” in neuropsychology, neurology, and other medical disciplines; however, such names appear localizationist in nature, deemphasizing one of its most important characteristics: its extraordinary connectivity and coordinated networking with the rest of the brain; moreover, this label seems oblivious to the bulk of research on the prefrontal cortex and executive functions, raising the question whether it would be more accurate to call it “prefrontal function” instead.

As it can be observed, choosing an adequate label for this cognitive “capacity” is challenging. The term “executive system” appears to be the most adequate term for theoretical conceptualization, but its operationalization in terms of behavior is difficult. For the purposes of this review and the following study, the term “*executive functions*” seems to be more appropriate, given three major reasons: first, it emphasizes the idea of “diversity” rather than “unity” as a characteristic of the system; second, it stresses its “functional” dimension more than its “morphological” association, allowing for the idea of “control” or “coordination” as one of its functions; and third, its behavioral operationalization appears more plausible, which is the ultimate objective of this validation study.

In association with what one could call “the labeling issue” in executive functions research, it is important to consider here three major current debates regarding the construct: firstly, the divergent localizationist versus connectionist (antilocalizationist) views of executive

functions, what will be called “the unit versus network” or neuroanatomical problem; secondly, and derived from the first topic, the discussion about the “unitary” versus “diverse” structure of this executive “system” of functions; and lastly, a discussion about what is the best way to assess it, what one could call “the measurement problem”, which is of great importance for the purposes of this study.

The Unit versus Network or “*Neuroanatomical Problem*” of Executive Functions

Since the early publication of H. L. Teuber’s (1972) article, “*Unity and diversity of frontal lobe functions*,” the question has been raised about the existence of a single system or a complex network of multiple subsystems in charge of these functions. That is, is it possible to define frontal lobe function in terms of a single unitary function, eventually, an *Executive Function*? Teuber reviewed several research discussions that took place at the symposium “Frontal granular cortex and behavior” in August of 1971, and compared them to earlier research in the area. His manuscript was organized under four questions: the “where,” “when,” “what,” and “how” of the granular prefrontal cortex functions. The “when” referred to developmental trajectories of the functions and other time-related issues. The “what” and “how” are presented as unresolved issues related to the actions and interactions of prefrontal neurons.

Important for this discussion is Teuber’s analysis of the “where.” Teuber asserted that there was a large amount of research demonstrating the need for a functional parcellation of the frontal lobes into several regions, given evident performance differentiation in response to diverse types of frontal lesions. In this way, he remarked that the functional and morphological double gradient “up-down” (dorsolateral frontal surface associated to delayed-response versus orbito-frontal convexity associated to emotional and behavioral control) and “back to front” (periarculate backward projection to interparietal sulcus related to spatial abilities versus sulcus

principalis projections to cingulate and hippocampus associated with memory), observed originally in monkeys and later reported by Brenda Milner in humans with frontal-lobe removal as treatment for epilepsy. In addition, Teuber presented what he called “a parcellation unique to man,” a differentiation between right and left frontal-lobe functioning, also found in the works of B. Milner with sorting tasks and fluency tests. It was observed that the difference between performances on the sorting task on fronto-dorsolateral-injured subjects versus orbito-frontal-injured subjects was not as significant as the difference on performances between left-sided versus right-sided frontal removal subjects. Right-sided removals had more difficulty with sorting whereas left-sided had more difficulty with verbal fluency.

These early conclusions have been consistently replicated in later studies of the frontal lobes. For several examples it is recommended to follow up Joaquin M. Fuster’s research on frontal lobes and prefrontal cortex (e.g., Fuster, 1985, 1991, 1997, 2000a, 2000b, 2001, 2002). Finally, an important point in Teuber’s work is his recognition of the connectivity of prefrontal cortex, in terms of its multiple reciprocal interactions with other cortical and subcortical systems (Teuber, 1972).

Baddeley (1996) has also criticized the idea of identifying executive function as a pure frontal lobe function, stating that such a complex function should not be defined in terms of a specific anatomical location, given the risks of either excluding important processes because they cannot be associated with frontal function, or including frontal lobe processes that are not executive in nature. Clinically, several cases suffering from impairments of specific executive functions (e.g. inhibition vs. working memory) after neurological (morphological and/or functional) frontal lobe damage have been commonly reported, as well as studies using observation of large clinical samples and performance of normal individuals on testing, which

have demonstrated that frontal lobe-lesion patients display different performance patterns during testing (e.g., Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999; Miyake, Friedman et al., 2000). These findings, however, do not offer a convincing answer to Baddeley's concerns.

It will be assumed in this study that executive functioning is associated with the frontal lobes, specifically prefrontal granular cortex, including its vast connectivity with the rest of the brain areas that allows it to monitor and execute goal-oriented activity. However, it is not assumed that the term "executive functions" is interchangeable with "frontal lobe functions." The analogy could be made with a sizeable factory, in which the headquarters are centrally located but the bulk of the execution of tasks is broadly dispersed. There is a large amount of research on the prefrontal cortex and its functions. The following section is a summary of key points that are important to recognize when performing measurement research involving executive functions.

Some key points on the neuroanatomy of executive functions

As established earlier, executive function has traditionally been associated in neuropsychology with frontal lobe functioning, specifically with prefrontal cortex activity. The prefrontal cortex corresponds to the most anterior area of the brain, and is one of the cortical areas to develop most and last in the course of individual development. According to Fuster (2002), executive function is a process where "ontogeny recapitulates phylogeny," meaning that during the course of human evolution, prefrontal areas were the last areas to develop (phylogeny) and the cycle repeats itself in each human-being's development (ontogeny) in which the prefrontal cortex is the last area to fully develop. Fuster also documented differences between gray and white matter ratios across the lifespan. Prefrontal gray matter appears to increase volumetrically after birth, reaching a "maximum" between ages 4 and 12, and then seems to

decrease gradually with aging. In contrast, the volume of the prefrontal white matter continues to increase during childhood and early adolescence (Fuster, 1993, 2002).

Around birth, as in later life, the development of axons and dendrites of frontal areas appears to be delayed with respect to other cortical areas. Moreover, cognitive development through the first years of the lifespan appears to correlate with the development of the prefrontal cortex. Fuster (2002) asserts that “this correlation is most obvious as we consider the evolution – with chronological age- of those cognitive functions of the prefrontal cortex that most contribute to intellectual maturation: attention, language, and creativity. All depend on the ability to organize behavior and cognition into goal-directed structures of action,” (p. 377).

Connectivity is the most relevant characteristic of the prefrontal cortex, and also the most difficult to examine. Prefrontal neurons have connections with other brain structures, both cortical and subcortical. Fuster (1997; 2002) asserts that the prefrontal cortex is possibly the best connected of all cortical structures. There is a consensus about prefrontal cortex specialization in the temporal structuring of new and complex goal-directed series of actions, applying to all forms of behavior (e.g., motor, speech, and reasoning). Fuster points out that its ability to deal with novel situations has given it the label “organ of creativity.” He also asserted that “the participation of the prefrontal cortex in the choice between alternatives, in decision making, and in executing temporally structured actions are the reasons that this cortex has been considered the ‘central executive’” (Fuster, 2002, p. 378)

Consistent with the concept of “connectivity” of the prefrontal cortex and frontal lobes, and supported by an extensive list of studies, Lezak (2004) asserted that executive functions are sensitive not only to frontal lobe damage but also to also to other brain areas (e.g., subcortical

structures and limbic systems), which are frequently affected by alcohol abuse, inhalation of organic solvents, and anoxic conditions, among others.

There is supporting research demonstrating differentiated brain pathways from the prefrontal cortex and other anterior, posterior, and subcortical brain structures, in relation to initiation of behavior, behavioral self-regulation, and emotional control. There are three main frontal circuits that are involved in executive functioning, dorsolateral, lateral orbitofrontal, and anterior cingulate circuit (Royall et al., 2002). The dorsolateral prefrontal circuit is related to Brodmann areas 8-12, 46, and 47, which are irrigated by the middle cerebral artery. This circuit includes the connections between the dorsolateral caudate nucleus (receiving input from the posterior parietal cortex and premotor areas), the dorsolateral portion of the globus pallidus and the rostral substantia nigra reticulate, and the parvocellular region of the medial dorsal and ventral anterior thalamic nuclei. Feedback connections to the frontal regions from the thalamus complete this circuit (Royall et al, 2002). This circuit has been involved in planning, self-monitoring, and other higher cognitive functions.

The lateral orbitofrontal circuit includes Brodmann areas 10-15 and 47, and is irrigated by the anterior cerebral artery (medial areas) and the middle cerebral arteries (lateral regions). This circuit includes projections from cortical areas to the ventromedial caudate nucleus, continuing to the dorsomedial area of the internal globus pallidus, then to the rostromedial portion of the substantia nigra reticulate; from there to the magnocellular region of the medial dorsal and ventral anterior thalamic nuclei, and back again to the lateral orbitofrontal areas (Royall et al, 2002). According to Royall et al., (2002; p. 381) this circuit may be “involved in the initiation of social and internally driven behaviors and the inhibition of inappropriate behavioral responses.” Barkley (1997) supported this neuroanatomical correlation and presents

evidence of research demonstrating the relationship between right prefrontal region and ventromedial regions of the striatum in interference control and response inhibition. It is possible that this regulatory function is mediated by the involvement of orbitofrontal cortex in the decoding and representation of primary reinforcers, which in turn regulate reward-related and punishment-related behavior (Rolls, 2002). Another circuit that appears to be involved in behavioral self-regulation and monitoring is the anterior cingulate circuit. This circuit involves Brodmann areas 9-13, 24, and 32, and is irrigated by the anterior cerebral artery. It includes pathways from the ventral striatum to ventral pallidum, rostromedial substantia nigra, and to the dorsomedial thalamic nuclei. The circuit ends with a pathway from the thalamus to the anterior cingulate (Royall et al, 2002).

Fuster (1997) affirmed that the orbitomedial frontal cortex has a close connection to limbic structures (i.e., amygdala), through cellular networks with a specific layer pattern which has been recognized as the “paralimbic” cortex. The ventromedial frontal cortex receives information from the autonomic system via projections from the amygdala, hypothalamus, and the magnocellular dorsomedial nucleus of the thalamus, and reciprocates this input with outputs to orbitofrontal areas that influence emotional behaviors.

An important line of research that has shed light on our understanding of the relationship between the frontal lobe functioning and executive functions has arisen from developmental neuropsychology research. Early work by Passler, Isaac, and Hynd (1985), demonstrated developmental differences in the performance on tasks involving the ability to self-regulate behavior and inhibit motor action in typically-developed (normal) children aged 6 to 12 years. Their experimental protocol included verbal (sequencing words) and nonverbal proactive inhibition tasks (sequenced tapping), verbal (two three-word series) and nonverbal (two three-

taps series) retroactive inhibition tasks. Four more tasks involved verbal and nonverbal conflict, and perseveration. Analysis of Variance (ANOVA) demonstrated significant differences across different developmental ages for all the tasks, yet differences between male and female groups were found no stable across tasks. Further analysis revealed that although all children aged 6 to 12 years were able to perform the tasks, performance on the younger group (6 years old) was poorer than performance in older groups (8 – 9, 10 -11, and 12 years old). Authors reported that these results demonstrated that the greatest period of executive development occurs at the 6 to 8 year-old level, but full mastery is acquired at age 12 (Passler et al., 1985).

This frontal functioning multistage process has been analyzed in other studies, and results appeared consistent. For instance, Becker, Isaac, and Hynd (1987) administered five specific tasks believe to measure frontal functioning to a sample of children aged 5 -12 years. The experimental tests included a simple Go/no-Go task, a Go/no-Go task with distractor stimuli, a nonverbal conflict task using auditory-sequential stimuli, a nonverbal conflict task with visual-simultaneous stimuli, and a temporal ordering task. Significant differences across ages but not across gender were reported. Similar to the Passler et al's (1987) results, these results demonstrated a developmental burst of executive/frontal lobe functioning at age 6 to 8 years old, with increasing developmental ability as children become older. Some stable race differences were observed across groups and genders as well (Passler et al., 1985).

An extensive and comprehensive study on the development of executive functions was recently published by Davidson, Amso, Anderson, and Diamond (2006). They explored working memory, inhibition, and task switching in a group of 325 participants aged 4 to 13 years utilizing a computerized battery that included a traditional Simon task, arrows congruency and incongruency task, a memory for dots task, and an abstract shapes test for memory. Results in

this study were significantly consistent with previous research (e.g., Passler et al., 1985, Becker et al., 1987). Inhibitory control was more difficult for the younger children and demonstrated to improve with age. By age 10, inhibitory tasks demands are compensated with working memory skills, whereas the opposite pattern was observed with younger kids (Davidson et al., 2006). These authors reported that working memory and the inhibitory control may not be independent functions in a developmental progression context. The ability to shift (cognitive flexibility) was lower in all groups respect to adult performance observed in other studies, and demonstrated to be less related to memory skills (Davidson et al., 2006).

The Unity versus Diversity or “*Conceptualization Problem*” of Executive Functions

As was pointed out earlier, there is not a more complex and yet ill-defined concept in neuropsychology than that of “executive functions.” Although understanding this concept has been of great importance when performing research on cognitive abilities such as working memory, attention, and intelligence (Miyake, Emerson, & Friedman, 2000), little is known about the relationships among these cognitive abilities, and research findings seem to be divided in two poles: those that support the hypothesis that they are all components of an underlying complex system (e.g., Duncan, Emslie, Williams, Johnson, & Freer, 1996) as opposed to those that demonstrate that they rely on independent systems (Friedman et al., 2006).

For instance, Fuster (1997) introduced the idea of the existence of an amount of neural energy, or *drive*, that is the source of alertness or general attention, and seems to be related to activation of dorsolateral and medial frontal cortex, after it is provided to the frontal cortex from the reticular formation and other subcortical and limbic formations via the mediodorsal nucleus of the thalamus. Luria (1973) defined it earlier as *arousal*, and linked its function to volition and initiation of behavior. Fuster associated this *drive* with the ability of the prefrontal cortex to

regulate selective attention, monitor and update relevant information, intentional elaboration and coordination of movement, and activation of specific schemas of the action and its goal. He added that these are “executive functions” and attention being one of them. Fuster divides attention into two components: a selective component (equivalent to Shallice’s SAS and related to dorsolateral frontal cortex), and an exclusionary component equivalent to an inhibitory control of interference.

Fuster’s early emphasis of the role of prefrontal cortex on the conceptualization of attention has resulted, in part, in the common association between executive functions and general attention. In March of 2001, Posner presented an important historical paper to the audience of the fourth Delange Conference, “Neurobiology of Perception and Communication: From Synapse to Society.” His paper was entitled, “Attention as an organ system” (Posner, *in press*). In this work, as it was shared by the author at the past 35th International Neuropsychological Society meeting, Posner collected evidence from cognitive science, clinical and applied neuropsychology, genetics, and neuroimaging, about the existence of an overall regulatory system, the attention, in human and other animal species. In brief, his theory asserts that there are three main attentional networks in charge of alerting, orienting, and the executive control of behavior. Alerting refers to wakefulness and arousal of an organism, orienting refers to the selection of information from the sensory registrar, and the executive control is related to the mechanisms to resolve conflict among responses, feelings, and thoughts (Rueda, Posner, & Rothbart, 2005). Posner and colleagues affirmed that this last element of attention, the executive control, is responsible for controlled and intentional aspects of attentional systems.

Other models of attention (e.g., Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991) separated its taxonomy into at least three components: focus, the attentional ability to select

relevant target information, which is associated with cortical areas and subcortical areas (thalamus and corpus striatum); sustain, the capacity to maintain focus and vigilance, related to frontal systems; and shift, the ability to change attentive focus when adaptive. Mirsky et al (1991) asserted that the term *attention for action* was commonly associated with the frontal executive systems but that was specifically involved in selecting relevant input and linking it to relevant output.

In other words, executive systems appear to perform a coordinating or controlling function over attentional systems, but general attention may not be necessarily an executive function itself. Baddeley presented the same case for working memory processes (as explained above), and recent research has isolated some of the executive control functions over working memory beyond the coordination between slave systems, introducing the role of executive control on working memory as the capacity to “update” working memory representations (Friedman et al., 2006; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake, Friedman et al., 2000).

There are some studies in which the hypothetical association between intelligence and executive function has been tested. Some of these studies have been designed under the assumption that executive functions and intelligence are part of the same latent construct, Spearman’s *g*, thus supporting the “unitary” hypothesis of executive function. In most cases, only low associations have been reported (Ardila, Galeano, & Rosselli, 1998) or none (e.g., (Welsh, Pennington, & Groisser, 1991); however, authors in favor of the “unitary” hypothesis have seen these results as illustrations of poor intelligence test construction. For instance, Ardila, Pineda, and Roselli (2000) performed executive functions and intelligence testing on a sample of fifty 13-to 16-year-old male children from Medellin-Colombia. Their purpose was to analyze the

relationships between executive function tests and IQ. They administered the WCST, the Verbal Fluency Test (phonological and semantic), and the Trail Making Test (TMT, A & B), for the assessment of executive function, and the Wechsler Intelligence Scale for Children-Revised (WISC-R), Spanish Version for the assessment of intelligence. Their results demonstrated significant correlations between some of the tests; specifically, verbal fluency tests significantly correlated with Verbal IQ (phonological fluency, $r = 0.32$, $p < 0.05$; semantic fluency, $r = 0.39$, $p < 0.001$), WISC-R Vocabulary (phonological fluency, $r = 0.40$, $p < 0.01$; semantic fluency, $r = 0.38$, $p < 0.001$), WISC-R Information (phonological fluency, $r = 0.34$, $p < 0.05$; semantic fluency, $r = 0.51$, $p < 0.001$), and WISC-R Similarities (phonological fluency, $r = 0.32$, $p < 0.05$; semantic fluency, $r = 0.44$, $p < 0.001$) perseverative errors from the WCST negatively correlated with WISC-R Arithmetic ($r = -0.36$, $p < 0.01$), WISC-R Block Design ($r = -.037$, $p < 0.01$), and WISC-R Verbal IQ ($r = -0.37$, $p < 0.01$); TMT-A Time negatively correlate with Performance IQ ($r = -0.31$, $p < 0.05$); and TMT-B Errors negatively correlated with WISC-R Vocabulary ($r = -0.34$, $p < 0.05$). Although significant, the sizes of these correlations are only low to modest, and the associations related mainly to Verbal-IQ tasks. These results serve as evidence of an association with cognitive abilities, yet not strong enough to consider them as parts of the same ability construct. However, given the original assumption, the authors concluded that “psychometric intelligence tests do not appropriately appraise intelligence,” (p. 35), adding that the most important aspects of intelligence relate to executive functions and are not included in the test structure (Ardila et al., 2000).

Duncan and colleagues (1996) hypothesized that Spearman’s g reflected the controlling functions of the frontal lobes, which they called “the conventional view,” suggesting overlap between functions related to g and impairments due to frontal lobe damage. They disagree with

Teuber's opinion about the unsuitability of IQ measures for demonstrating frontal functions, stating that this misconception occurs when IQ is taken as the only estimation of intelligence. IQ represents a composite of tasks that correlate with each other, but research has shown that tests of "fluid intelligence" load the highest on g (Kamphaus, 2001), with Progressive Matrices being a good example of novel problem-solving task with spatial material, and one of the most closely related to g . In addition, they consider the IQ tests dependence to knowledge as a limitation more than strength, especially if used to measure changes of g following brain damage.

In this regard, and in support of Duncan et al.'s (1996) argument against using FSIQ as a measure of brain changes, Aylward (2002) presented a review of common developmental outcomes observed in follow-up studies with preterm infants that were not predicted by IQ measures. He states that although several studies reported group FSIQ differences of 5-7 points, preterm infants continue to classify as "average" or "low average" on intelligence, whereas their school performance is far from being average; therefore, he hypothesizes that a major constraint is placed when composites are used, allowing subtle impairments to be masked. Earlier, Duncan, Burgess, and Emslie (1995) demonstrated that when a test of fluid intelligence is administered, such as the Cattell's Culture Fair test of g , a test with a high correlation with g (.81) that assesses spatial problem-solving, clinically significant deficits on IQ (23-60 points) become evident, assuming integrity of other functions required for performing the test (e.g., visual perception).

Furthermore, Duncan and his colleagues suggested that a diversity of functions reflected in g could be associated with specific frontal regions (e.g., dorsolateral prefrontal cortex, anterior cingulate) given findings on neuroimaging research (Duncan et al., 1996). Duncan and colleagues attempted to demonstrate this hypothesis using a simple task in which subjects were

shown to consistently neglect a component of the instruction, and this goal negligence was associated with *g*.

Duncan and colleagues' (1996) study was proven to be limited, a year later. Duncan, Johnson, Swales, and Freer (1997) ran a series of executive functions tasks on a larger sample of head-injured subjects and found no significant differences between performances on diffuse cortical damaged subjects versus focal frontal lesions. He attributed to connectivity between frontal lobes and the rest of the brain, and the potential damage to functional paths, which might explain the relationship found between *g* and frontal lobe impairments that he neglected earlier. They also reported that the instruction negligence phenomena was indistinguishable observed in both frontal-lobe injured and healthy subject.

Similarly, Friedman and colleagues (2006) recently reported a latent variable analysis of a sample of adolescents (16-18 years old), of the relationships of fluid and crystallized intelligence (as measured by the Wechsler Adult Intelligence Scale) to three executive functions: inhibiting prepotent responses, shifting mental sets, and updating working memory. They asserted that neuropsychological research has neglected the inclusion of crystallized intelligence in their analysis of the relationships between executive functions and intelligence, based on evidence of the relationships between fluid reasoning impairments and frontal lobe damage. However, and as it has been suggested in Carroll's model (Kamphaus, 2001), fluid intelligence might be necessary to acquire knowledge -the foundation of crystallized intelligence; therefore, their analysis rested on the idea that different executive functions may have a unique relationship with intelligence.

Subjects were recruited from the Colorado Longitudinal Twin Study, and a battery of computerized executive function tasks in addition to paper and pencil intelligence measures was

administered to 234 twins. Earlier, Miyake and colleagues (2000) performed a latent variable analysis using structural equation modeling (SEM) techniques on a battery of executive functions tasks, to assess their individual contributions to performance on complex executive tasks (e.g., WCST). Three major executive functions, “inhibiting” of automatic responses, “updating” working memory, and “shifting” mental sets were isolated, and a confirmatory factor analysis indicated that they were related but separable. Details of this study are described later in this review. Friedman et al’s (2006) findings guided them to conclude that both fluid and crystallized intelligence measures relate closely to “updating” working memory, sharing 41% to 48% of their variances with this executive function, whereas “inhibiting” of automatic responses and “shifting” mental sets, only shared 2% to 14% of their variances. They pointed out that correlations found between inhibiting and shifting with intelligence measures were derived from their own shared variance with updating, since the three functions moderately correlated.

Although Friedman et al’s study offers an important contribution to the debate about the unity or diversity of executive functions, in that it demonstrates that there are strong relationships between at least one specific executive function, *updating working memory*, and both fluid and crystallized intelligence, some limitations could be pointed out. First, although the authors make an effort to differentiate “updating of working memory representations” from simply “working memory,” it appears that such a differentiation only applies to their theoretical frame, but not to their methodological design. Their “updating” tasks are defined as requiring adding and deleting information in working memory, but it is unclear how that process is not working memory itself. Failing to clarify this distinction raises questions, specially if they neglect the fact that Carroll’s three-stratum model of intelligence included a significant factor loading of “memory” itself (including working memory) in *g* (Kamphaus, 2001).

Second, it is confusing if the authors support either ideas of “unity” or “diversity” of executive functions, in that on one hand, their results appear to point out that intelligence only relates to a specific executive function, which would underlie the idea of diversity (if only one function relates to intelligence, thus there are other “functions” that function as independent systems). On the other hand, their last conclusive idea on the manuscript asserts that “traditional measures of intelligence are missing some fundamental supervisory functions,” a statement that underlies the idea of “unity” in that they blame the intelligence measure of lacking of tasks that would demonstrate intelligence association to other executive functions. This would be only plausible if the idea of a equality between executive functions and intelligence were to be assumed, which does not seem to be the case as the article title states “*Not all executive functions are related to intelligence*” (Friedman et al., 2006).

Finally, Friedman et al’s contribution appears to undermine another ongoing issue that is part of this debate, that is, if “diversity” were to be assumed, which are the “functions” involved in the executive system? In this regard, the authors failed to clarify under what theoretical background (other than Miyake et al’s earlier study) it was assumed, for instance, that *updating working memory* could be considered or not an executive function. As it was presented earlier in this chapter, Baddeley’s model of working memory included the central executive as a coordinating component of the working memory model, not vice versa (Baddeley, 1983). Therefore, is their intention to rename Baddeley’s “central executive” for what they called “updating working memory representations?” Since this is an empirically unresolved situation, it would have been important to clarify their stand point, at least theoretically. These types of limitations are more evident when there is a lack of a clear consensual definition of a construct, its components, and its operationalization.

Towards a unitary yet diverse definition of executive functions

If an assumption of the diversity of the executive functions were to be made, questions that arise include: what are its components? and further, can we define “executive functions” in terms of its components? This ongoing controversy has made it difficult to confine the construct of “executive functions” into a one statement-definition. For instance, Elliot (2003) asserted that “there is no intuitive lay concept that incorporates the essence of executive function” (p. 50), and that the term has been used to define complex cognitive processing requiring the coordination of several sub-processes to achieve a particular goal. Royall and colleagues (2002) defined executive functions as a set of cognitive skills responsible for the planning, initiation, monitoring, and evaluation of complex goal-directed behavior. Tirapu-Ustárriz and colleagues (2002) presented executive functions as a series of processes that associate ideas, movements, and actions geared towards complex behaviors. Lezak (1982) affirms that, “The executive functions comprise those mental capacities necessary for formulating goals, planning how to achieve them, and carrying out the plans effectively” (p 281).

According to Gioia and colleagues (2000), executive functions in children are “mental processes that direct a child’s thought, action, and emotion, particularly during active problem solving. Specific skills include (a) selecting appropriate goals for a particular task, (b) planning and organizing an approach to problem solving, (c) initiating a plan, (d) inhibiting (blocking out) distractions, (e) holding a goal and plan in mind, (f) flexibly trying a new approach when necessary, and (g) checking to see that the goal is achieved. The executive functions are also responsible for controlling a child’s emotional responses, thereby allowing for more effective problem solving.” Gioia and colleagues (2000b) also agree that executive functions not only control cognitive behavior but also emotional and motor-action behavior.

As Miyake, Emerson, and Friedman pointed out (2000), many authors attempted to define this construct by enumerating its components rather than actually developing a precise definition of it. This “appealing” listing has, as these authors assert, several problems. Among them are the overlapping subsystems (e.g., “planning” and “sequencing”, “inhibition” and “resistance to interference”), the difference between complexity levels (e.g., “problem solving” versus “inhibition”), different classifications across authors and models; and moreover, misspecification of definitions (e.g., “cognitive flexibility” defined as both “attention switching” and “task switching”).

At the same time, it can be observed that some of those definitions imply both “unity” and “diversity” of executive functions. The tendency is to group, but this unitary representation is more a theoretical artifact that helps us to approach the concept, rather than an actual one-single system. In those lines, “executive functions” would be the term used to name the “network” as shorthand for the complexity of levels, systems, and behavioral and cognitive abilities that this network includes; moreover, it seems plausible to say that there is a consensus that such self-regulatory yet complex executive system encompasses a set of coordinated cognitive abilities required in the planning, monitoring, and execution of behaviors oriented towards a goal (Anderson et al., 2002; Elliott, 2003; Lezak et al., 2004; Mahone, Cirino et al., 2002; Royall et al., 2002). Therefore, and as it was recommended by Baddeley (1996), the first step in understanding this complex function, consists in identifying, defining, and analyzing its processes, before being able to decide if they should be studied as a unit, or as a set of “individual and separable functions.”

Executive “functions”

If diversity of executive “functions” is to be assumed here, it becomes important (as Baddeley recommended) to identify and differentiate its taxonomy. There is no agreement about what sort of functions are “executive” in nature, and what are not (Rabbitt, 1997). Debatable executive functions include planning and organizing, initiating action plans (volition), inhibiting automatic distractions and interference, inhibiting impulses and automatic responses, process and activity monitoring, decision-making, shifting between actions, flexibility in novel situations, emotional control, working memory, and attentional control.

Royall and colleagues (2002) presented in their extensive report from the Committee on Research of the American Neuropsychiatric Association, the existence of two main themes on executive functions considered on the literature: first, the “executive cognitive functions” involved in functions such as insight, abstraction, and judgment (Royall et al., 2002); and second, the “cybernetic” or “executive” aspects of executive functioning (e.g., cognitive control); in other words, the interactions between executive functions and nonexecutive processes to control the execution of complex activities. This second theme considers that executive impairments would be only observed if disorganization in operations of the nonexecutive domains is observed, and is also recognized for its emphasis on the networking associations between frontal lobes and other brain areas (Royall et al., 2002).

There are some research studies on the dimensions or components of executive functions in both pediatric (e.g., Lehto et al., 2003) and adult populations (Miyake, Friedman et al., 2000). For instance, Miyake, Friedman et al (2000) performed a latent variable analysis of executive functions using SEM and CFA (with the CALIS procedure from SAS) on a group of 137 undergraduates from the University of Colorado at Boulder. They pragmatically identified three

basic executive functions based on the literature and previous research, possibility of an operational definition, and their potential implications in the performance of complex executive tasks. These functions were: mental set shifting (“Shifting”), information updating and monitoring (“Updating”), and inhibition of prepotent responses (“Inhibition”). Each latent construct had three indicators, and each indicator, (observed variables) corresponded to a basic executive function tasks (e.g., plus-minus, number-letter, antisaccade eye movement, stop-signal, letter memory, and tone monitoring).

One of the aims of their study was to provide a complex analysis of the unity and diversity of executive functions based on two questions: to what extent are the different executive functions separable and to what extent are they a one-common function. A full CFA model and multiple nested models were created and analyzed. The three-factor model produced the best fit indexes, and $\chi^2 (24, N=137) = 20.29, p > .65, AIC = -27.71, SRMR = .047, CFI = 1.00, IFI = 1.04$. These results suggested that the three executive functions postulated are indeed separable but moderately correlated constructs, which is consistent with the idea of the unity and diversity of executive functions. In this regard, the correlations among tasks and latent constructs demonstrate that there is a shared common task requirement important for the attainment of a goal. Authors proposed that this common agent could be “controlled attention,” a domain-free attentional capacity related to goal maintenance, conflict resolution, and error monitoring, among other controlling aspects needed on the execution of any goal. In contrast, the goodness of fit of their CFA model demonstrates that including several yet different latent constructs is very valuable, and counts as support for the “diversity” component of the executive functions. In fact, the authors concluded that future research should include other functions (e.g., dual task

performance), and also intermediate levels of analysis such adding complex executive functions to the model (e.g., “planning”) (Miyake, Friedman et al., 2000).

Lehto and colleagues (2003), investigated the dimensions of executive functioning in a group of 108 normal children (8-to 13-year-olds), using a set of measures of executive functions for children. The instruments included the Trail Making Test (TMT) parts A and B, and an additional part C (in which the participants were asked to connect a set of 15 letters instead of numbers); the Auditory Attention and Response Set (AARS) from the Developmental Neuropsychological Assessment (NEPSY), the Word Fluency task from the NEPSY, Matching Familiar Figures, Mazes from the WISC-R, and four tasks from the Cambridge Neuropsychological Test Automated Battery (CANTAB) Working Memory and Planning Battery, including a computerized version of the Tower of London. Authors performed a preliminary EFA (principal components with Direct Oblimin Rotation) on SPSS, and a CFA using LISREL modeling (LISREL 8.30, Jöreskog & Sörbom, 1993) and Maximum Likelihood as estimation method. The EFA resulted on a three-factor solution that counted for 40.5% of the total variance. These factors were labeled working memory, inhibition (which included the Tower of London), and shifting (which included verbal fluency). The CFA model confirmed the EFA model. The best fit indexes were obtained when a three factor structure was used. To test the hypothesis of “unity” they performed CFA using only one factor (all variables loading in one factor) but poor fit indexes suggested rejection of the model; however, the three-factor CFA model fit the best when correlations among factors were allowed (which is set by default in LISREL). These results support the hypothesis of a “unitary” yet “diverse” structure of executive functions in children.

It is important to note here that both studies (Miyake's and Lehto's) assessed two developmentally and culturally different and independent samples, with different instruments, yet both studies seem to be consistent in demonstrating that inhibition, updating working memory, and shifting are important "cognitive" components of executive function. More cross-validation analyses are needed to provide with evidence for the inclusion of different components on the structural models. Figure 1 illustrates a conceptual map summarizing the empirical research results of the studies presented. This hypothetical model functions as a conceptual map, not as a structural model; therefore, directionality is not assumed. In this case, the ellipses represent three latent variables believe to group as "cognitive" components of executive functions. The curved-double head arrows represent relationships among all the latent constructs. Straight single-headed arrows represent membership into a category, not causality.

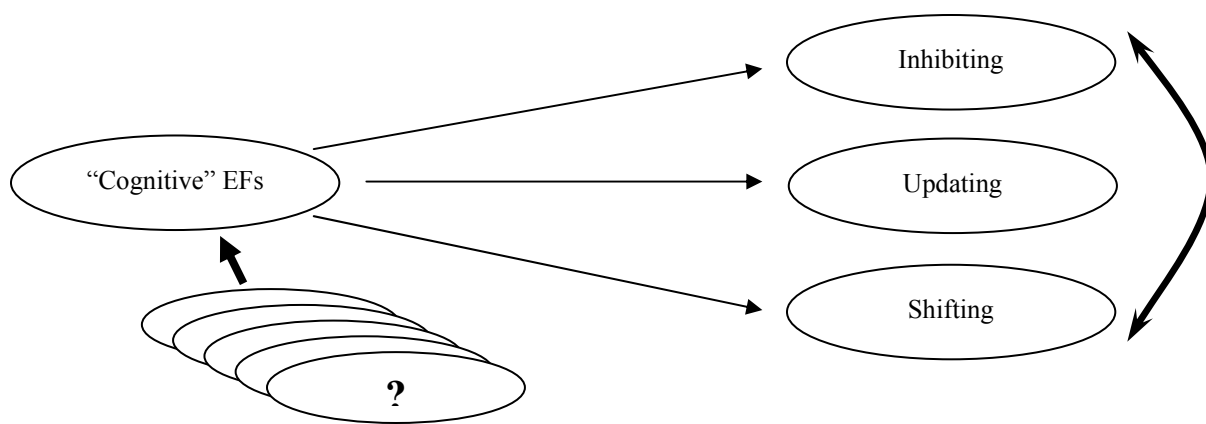


Figure 1. Conceptual model of evidence-based cognitive taxonomy of executive functions

In contrast to the more cognitive-oriented models of executive functions, some other definitions of this construct have a behavioral orientation. For instance, Lezak (2004, page 611) referred to executive functions as "the most complex of behaviors" and from that perspective she includes the following components: (1) volition, defined as the capacity for intentional behavior,

including the formulation of a goal, conceptualization of what is needed, motivation, and ability to initiate an activity or program of actions; (2) planning, defined as the ability to identify and organize the steps and elements needed to carry out an intention or achieve a goal; (3) purposive action, or “programming of activity” refers to the translation of a plan into a productive activity, including maintain, switch, and stop sequences of complex behavior; and (4) effective performance, related to self-regulation and monitoring during the performance of the activity.

Several authors have integrated Lezak’s taxonomy of executive functions into their research (V. Anderson, 2001; Tirapu-Ustarroz et al., 2002), and it has been utilized as a clinical-framed tool for the assessment of executive dysfunction. However, this model has also been criticized for difficulty in its operationalization for research purposes (e.g., Miyake, Friedman et al., 2000). The following conceptual map illustrates a hypothetical model of the behavioral components of executive functions.

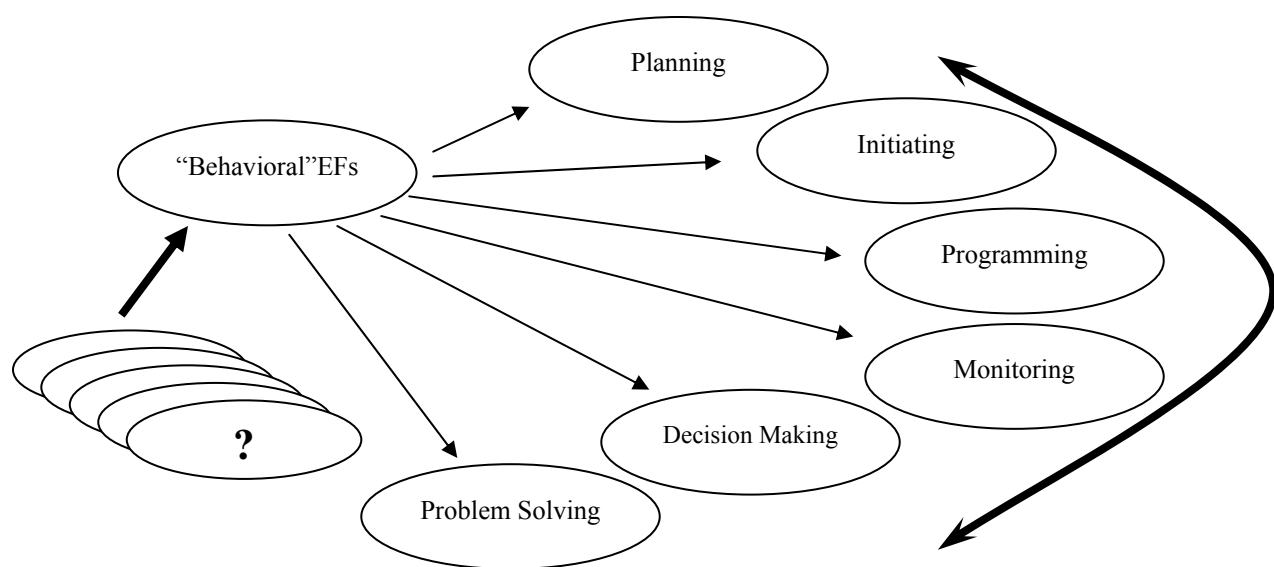


Figure 2. Conceptual map of a hypothetical/theoretical-based model of the behavioral taxonomy of executive functions

Although not originally included in Lezak's model as independent components, there is an extensive amount of research in the executive aspects of decision making and problem solving (e.g., Burgess, 2000; L. Clark & Manes, 2004), including their relationship with other cognitive abilities (Bechara & Martin, 2004; L. Clark & Manes, 2004; Fuster, 1995; Kerr & Zelazo, 2004), their association with prefrontal cortex (S. W. Anderson, Damasio, Tranel, & Damasio, 2000; Dalley, Cardinal, & Robbins, 2004; Ernst et al., 2003; Fuster, 1997; Manes et al., 2002; Robbins, 1996; Scott, Holmes, Friston, & Wise, 2000), and their impairing effects on diverse psychopathology (Barkley, Murphy, Dupaul, & Bush, 2002; Bechara et al., 2001; Bechara & Martin, 2004; Bolla, Cadet, & London, 1998; Brand et al., 2005; Brand et al., 2004; London, Ernst, Grant, Bonson, & Weinstein, 2000; Rahman, Sahakian, Hodges, Rogers, & Robbins, 1999; Watkins et al., 2000).

Finally, and although two models of the components of executive functions have been presented, there is yet one more to introduce here. Historically, advances on the study of human behavior have led to a better understanding of cognitive functioning; furthermore, the need for integrating concepts to explain the complexity of the relationships between cognitive systems gave rise to the creation of hypothetical models that included a coordinating function. Due to its "cybernetic" capacity, this function was recognized as "executive" in nature. This third hypothetical paradigm seems to have emanated from the earlier works of Norman and Shallice (Shallice, 1982; Shallice & Burgess, 1996) and Baddeley (Baddeley, 1983, 1996) on the *Supervisory Attentional System* (SAS) and the *Central Executive* explained earlier. The focus of this orientation relates to the "controlling" or "coordinating" aspects of executive functions over other cognitive abilities (e.g., attention, working memory), emotion (e.g., lability, mood changes), and behavior (e.g., impulse control, behavioral regulation).

As Royall et al (2002) asserted, "...executive functions control the execution of complex activities. This view implies, first that ECF [executive control function] interacts with nonexecutive processes, and second that ECF impairment is made visible only via the disorganized operations of nonexecutive domains" (p. 378). This conceptualization of executive functions as a control system has been also recognized as a "self-regulatory" capacity, highly related to frontal and prefrontal activity (Stuss & Benson, 1986). This model includes a differentiation between two major concepts, capacity and control. Royall et al (2002) presents a good example of this conceptualization using the clock-drawing task as an illustration. Elderly retirees performed significantly poorer when asked to draw the clock in the absence of visual clues. Once the model is provided and they are to copy it, they significantly improved their performance. This pattern of performance is a demonstration of the role of executive control in visual-spatial tasks, and has been observed in normal children (Cohen, Riccio, Kibby, & Edmonds, 2000) and in children with ADHD (Kibby, Cohen, & Hynd, 1999, 2002).

Similar roles of the executive control have been reported for memory (Fletcher, Shallice, & Dolan, 1998; Fletcher, Shallice, Frith, Frackowiak, & Dolan, 1998) and attention (M. B. Denckla, 1996b; Luria, 1973). Posner and Rothbart (1998) called it "*executive attention*" and demonstrated that it was observable -yet immature- not only in adults but also in children as young as 18 months (with a context-sensitive-learning of sequences task) and in toddlers of 2-3 years of age (with a Stroop effect task). This developmental ability of *executive attention* appears to positively correlate with the development of the anterior cingulate (Posner & Rothbart, 1998). This executive attention or "attention control" executive function also appears to be associated with performance on attention tasks, such as the Continuous Performance Tests or "CPTs" (Riccio, Cohen, Hynd, & Keith, 1996; Riccio & Reynolds, 2001; Riccio, Reynolds, Lowe, &

Moore, 2002; Riccio, Reynolds, & Lowe, 2001). Recently, Mani, Bedwell, and Miller (2005) documented developmental differences in CPT performances in a sample of normal young and older adults. Significant age-related differences were found, specifically in the number of commissions and false alarm errors, in which the older sample performed worse. These findings contribute to the evidence of age differences in executive functioning across the lifespan (Kray, Eber, & Lindenberger, 2004), and furthermore, to the effects of age on the attentional control and selective response inhibition (Posner & Rothbart, 1998; Tamm, Menon, & Reiss, 2002).

The role of executive *control* functions over emotional behavior has been reported as well. For instance, Fuster (1997) asserts that the regulatory function over emotion is associated with the close connection between orbital and medial frontal-lobe cortex with limbic structures. This area has been recognized as the “paralimbic cortex” for that matter. The frontal-limbic pathway is reciprocal, receiving inputs from limbic structures (e.g., amygdala and hypothalamus) and autonomic systems via projections from the magnocellular portion of the dorsomedial nucleus of the thalamus to the ventromedial frontal cortex, and sending output to the same structures. This mechanism contributes to the regulation of emotional behavior. Furthermore, this regulatory loop has also been associated with the frontal initiation and integration of emotional behavior, as it has been clinically observed in patients with ventromedial lesions and subsequent emotional disturbances, and also identified with neuroimaging techniques (Fuster, 1997).

Based on Baddeley’s models and following research on the prefrontal cortex connectivity, Patricia Goldman-Rakic and colleagues (1992), presented a model of the central executive system and its relationship with the regulation of motor behaviors such as initiation, facilitation, and inhibition. In their words, this framework “isolates internally memory-based functions from externally guided sensory-functions, identifies the internally-based functions with

prefrontal circuits, the sensory-based functions with premotor circuits, and integrates the specializations of prefrontal, premotor, and subcortical structures in the control of motor acts” (Goldman-Rakic, Bates, & Chafee, 1992). A very interesting point of this model is the association between dopamine and the cognitive control over motor programs. In their model, prefrontal control is necessary for the direction of motor programs over a specific goal or action, and the executive central will be the system responsible for the coordination of the three independent control systems: motor, sensorial, and mnemonic (Tirapu-Ustarroz et al., 2002).

In this direction, Fuster (1997) asserted that another function of the prefrontal cortex, specifically the orbitofrontal area related to behavioral regulation, is the inhibitory control of interference, which protects behavior, language, and thinking processing from internal and external influences. This protection from distracters appears to be of evolutionary importance, and relates to the capacity to direct attention to motor programs and abort impulses toward immediate gratification that threaten the structure of adaptive and goal-oriented behavior. Impaired inhibitory control can result in the presence of perseverative behavior.

Based on this conceptualization of executive behavioral control and self-regulation, Grigsby and Kaye (as cited in Suchy, Leahy, Sweet, & Lam, 2003) developed the Behavioral Dyscontrol Scale (BDS) to measure executive processes such as impulse control and initiation of behavior. Later validity studies using factor analysis demonstrated that this 10 minute measure with 9 items had a three-factor structure: (1) Motor Programming Factor, including tasks believed to tap abilities to volitionally generate and sustain motor responses; (2) Environmental Independence Factor, with go/no go tasks for the assessment of impulsivity and environmental dependency; and (3) Fluid Intelligence Factor, including working memory tasks and insight items believed to measure the ability to use feedback and to reason (Suchy et al., 2003).

Convergent and discriminant analysis (using binary logistic regression models) of the scale's capacity to accurately classify cases demonstrated that the scale was more successful in the rates of correct classification than traditional measures of executive functions (Trail Making Test -Part B, Control Oral Word Association, and Stroop-Color Word Position).

These findings serve as support for the relevance in considering the *control* components of executive functions in assessment. Although BDS has been validated only in adult samples, evidence of the self-regulatory executive functions has been reported in children as young as 3-years old (Rueda et al., 2005). The following conceptual map (Figure 3) summarizes the control and self-regulatory executive functions model.

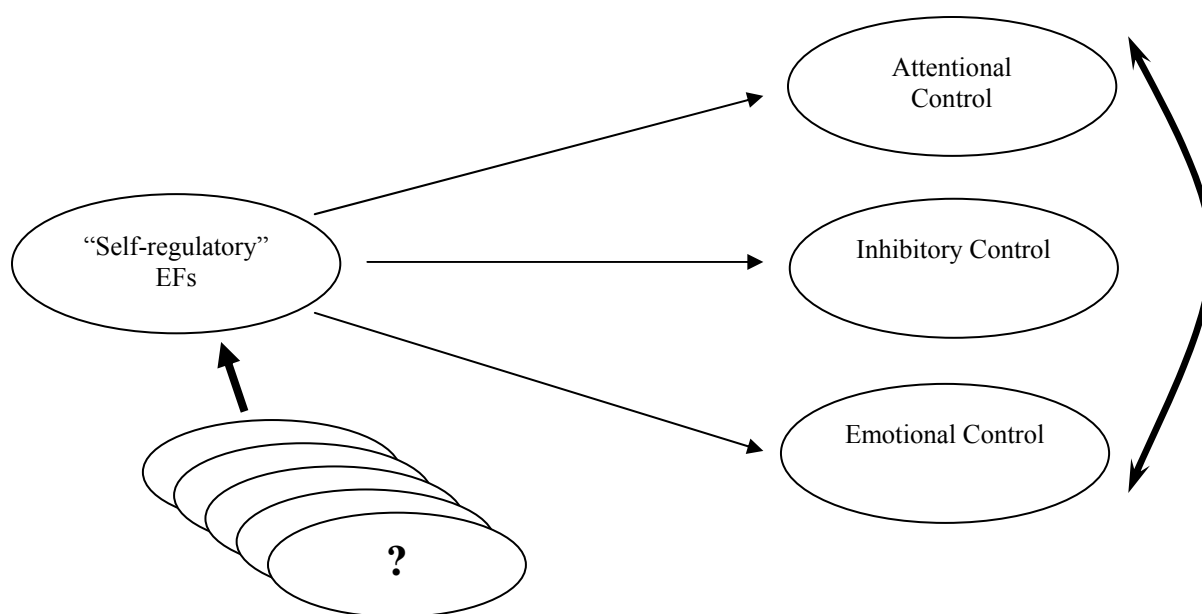


Figure 3. Conceptual map of an evidence-based model of self-regulatory components of executive functions

The “Measurement Problem”

“The inherent weakness within the assessment of frontal lobe development is the lack of consensus of what the frontal lobe skills are exactly and how they are best assessed.” (Samago-Sprouse, 1999, p. 587)

As Strayhorn (1993) presented, the problem of finding valid and reliable measurements of psychological constructs is more common than is expected, especially in the case of theoretically complex constructs such as *executive functions*, whose complexity creates even more measurement challenges (Gioia, Isquith, Retzlaff, & Espy, 2002). However, several efforts have been made to improve the sophistication of measurement techniques of executive functions.

Lezak (2004) points out that an important first step in the assessment of executive functions is to differentiate them from other cognitive abilities. She offers a criterion to define if a measure is “executive” or not by stating that, “Questions about executive functions ask *how* or *whether* a person goes about doing something (e.g., Will you do it and, if so, how and when?); questions about cognitive functions are generally phrased in terms of *what* or *how much* (e.g., How much do you know? What can you do?)” (p.35). In their review, Royall et al (2002) presents more examples of questions to assess executive and nonexecutive capacities (p. 384).

Traditionally in neuropsychology, it is believed that complex tasks such as the Wisconsin Card Sorting Test (WCST) (Heaton, 1981), Towers of Hanoi, London, and variations (TOL) (Anzai & Simon, 1979; Shallice, 1982), and Complex Figure of Rey (ROCF, Rey & Osterrieth, 1993), serve as gold-standard tools for assessing executive function (Lezak et al., 2004). However, there is an accumulated amount of research including criticisms of the limitations of sensitivity and specificity of these and other instruments, which in turn has created the need for a more multidimensional approach (V. Anderson, 2001; Hughes & Graham, 2002; Manchester et

al., 2004). A closer look at these instruments allows a better understanding of the source of these criticisms.

The Wisconsin Card Sorting Test- WCST (Heaton, 1981)

Heaton and colleagues described the WCST as a measure of “abstract reasoning ability to shift cognitive strategies in response to changing environmental contingencies. As such, the WCST can be considered a measure of ‘executive function’, requiring the ability to develop and maintain an appropriate problem-solving strategy across changing stimulus conditions in order to achieve a future goal” and added that “the WCST requires strategic planning, organized searching, utilizing environmental feedback to shift cognitive sets, directing behavior toward achieving a goal, and modulating impulsive responding.” Given its sensitivity to frontal lobe damage, results on the WCST have been associated with frontal lobe and prefrontal cortex functioning.

The WCST has been the product of several years of modifications and research on its validity and reliability as a measure of executive function. In 1948, Esta Berg, working with previous research by Harlow (1946, cited in Heaton, 1981) in primates, described a procedure for assessing flexibility of thinking in humans. Grant and Berg modified the original procedure, and a product of these changes was the creation of the first version of the Wisconsin Card Sorting Test. A series of studies using this test followed and guided later modifications of the tests. Given the absence of a standard procedure for administration and scoring of the test, variations in materials and procedures made it difficult to compare findings across studies. In 1981, Heaton and colleagues provided a standard method and some normative data to assist clinical interpretation. This latter version of the WCST has been extensively used; however, it is important to note that one of the limitations of the test manual is that it does not provide

sufficient information on the process of development of the different test components (such as “number of trials”, “perseverative responses”, “learning to learn”, and “failure to maintain set”).

The WCST is the most studied executive function test in history; there are hundreds of studies of its validity for the diagnosis of frontal lobe dysfunctions since its earlier creation in 1981, and since Brenda Milner published the historical report of patients undergoing unilateral frontal lobe damage, providing evidence of its use for diagnosis of frontal lobe functional impairments (Milner, 1982). The WCST has a longer tradition of being included as an assessment tool in research, which makes it surprising to discover how little is dedicated in the manual to the explanation of test development. Its strength has come from multiples studies of its validity and uses in diagnosing executive dysfunction after lesions on the frontal lobes.

However, external studies about its psychometric properties have raised questions about its reliability (Bowden et al., 1998); construct validity using EFA (Greve, Brooks, Crouch, Williams, & Rice, 1997) and latent variable techniques (Greve, Stickley, Love, Bianchini, & Stanford, 2005) have questioned its capacity to differentiate among executive functions; sensitivity and specificity have been reported to be inconsistent (Demakis, 2003), and it has been suggested that the WCST is not specific to frontal lobe damage per se (Miyake, Emerson et al., 2000; Shallice & Burgess, 1991). Moreover, it has been reported that the appropriateness of this test for children has been questioned as well, given the developmental variability of younger children (Riccio et al., 1994).

Towers of Hanoi, London, and variations- TOL (Anzai & Simon, 1979; Shallice, 1982)

This task requires the subject to move disks from one peg to another, one at a time, with the aim of obtaining a goal-model configuration from the original presented, in the minimum number of moves. The Tower of Hanoi uses same size pegs but different size disks, and the

subject has to follow a set of rules, including one requiring that bigger disks not be placed on top of smaller disks. The Tower of London uses same size disks but different size pegs, so the number of disks on a peg is constrained. The Developmental Neuropsychological Assessment, NEPSY (Korkman, Kirk, & Kemp, 1998) includes a tower task for children that combines both original towers' characteristics, different size pegs, and different color disks.

The tower task has been traditionally considered as an example of the executive function of *planning* (Lezak et al., 2004). However, recent studies using SEM techniques for confirmatory factor analysis and latent construct analysis have consistently demonstrated that the tower task loads higher on factors associated with inhibitory control (Friedman et al., 2006; Kafer & Hunter, 1997; Lehto et al., 2003; Miyake, Friedman et al., 2000) thus questioning its validity. Moreover, analysis of its reliability have also been demonstrated to be poor (Miyake, Emerson et al., 2000).

Rey-Osterrieth Complex Figure – ROCF (Rey & Osterrieth, 1993)

Created in 1941 by Andre Rey, this geometrical complex figure was designed to assess visuoconstructional ability and visual memory in brain-injured patients. Later, Paul Osterrieth designed a systematic scoring system for its standardization. Given the complexity of the figure, the test has been used to examine other abilities, including executive functions such as organization and planning. Qualitative methods of interpretation of the drawings have been introduced for this purpose (i.e., using different colors during construction for further sequencing analysis). Although originally created for adults, this test has been standardized for pediatric assessment as well, and research results have demonstrated that if the qualitative methods are used, it is a useful and valid tool for the assessment of executive functions in children (Watanabe et al., 2005).

However, poor performance on the ROCF could be explained by impaired visual-spatial abilities or constructional praxis malfunctioning, compromising the potential to isolate the executive component. This potential overlapping of the functions that the task is supposed to tap has been recognized as the “task impurity problem” generally observed in complex executive tasks (Miyake, Emerson et al., 2000).

Taken as a group, executive tests appear to tap many functions at once (Miyake, Emerson et al., 2000), are difficult to factor analyze (M. B. Denckla, 1996a), and correlate poorly within and between themselves (Ardila et al., 1998; Ardila et al., 2000; Pineda & Merchan, 2003). Denckla (1996a) asserts that this low correlation between measures believed to tap executive functions may be explained by a “method variance rather than any clear central construct of EF [executive functions]” (p. 6). Table 1 presents some traditional tests used for the assessment of cognitive executive functions.

Table 1

Summary of traditional neuropsychological tests used for the assessment of cognitive executive functions

Test or task	Task description	Construct measured or Cognitive ability
Porteus Mazes, WISC- Mazes, Rey-Osterreith Complex Figure Drawing Test, Clock face.	The child is asked to perform tasks where planning of motor responses is required to succeed (e.g., drawing, finding the way out from a maze, finding a sequence of movements in a goal-oriented task).	Control of motor responses –Planning
Go/No Go -SST, CPT, Match-to-Sample tests, Stroop Color-Word Test, Wisconsin Card Sorting Task, Tower of Hanoi and London.	Motor inhibition tests where the child is tested for errors of commission, decreased accuracy or shorter latency of response, interference effects from conflicting stimuli, and set-shifting difficulties.	Control of motor responses –Execution and inhibition
Backwards digits span, self-ordered pointing, paced	The child is asked to retain in operational memory increasing	Working Memory

auditory serial addition, Simon Tone/Color Game and Dot Test of Visuospatial Working Memory	amounts of information (letters, numbers, colors, tones, dots)	
Continuous Performance Test –CPT (Auditory or Verbal). There are several versions, e.g.:	The child is instructed to press a button each time a predetermined stimulus is presented.	<p>Attention: <i>Omission errors</i>, number of times the designated stimulus is presented and the child fail to press the button.</p> <p>Impulsivity: <i>Commission errors</i>, number of times the child incorrectly presses the button when the designated stimulus is not presented.</p> <p>Response inhibition: From a variation of the original version in which the child is asked to respond when they detect a predetermined stimulus after seeing another predetermined letter.</p>
<ul style="list-style-type: none"> - (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956) - (Gordon, 1983) - (Conners, 1995) 		
Gordon Diagnostic System (Gordon, 1991)	<p>Three tasks:</p> <ol style="list-style-type: none"> 1. Standard Delay Task (delay responding for six seconds) 2. Standard Vigilance Task (respond to 1s followed by 9s) 3. Distractibility Task (ignore distracting stimuli) 	<p>Attention: Omission errors</p> <ul style="list-style-type: none"> - Inhibition of responding: Delay task - Sustained attention: Vigilance task - Selective attention: Distractibility task <p>Impulsivity: Commission errors</p>
The Stop Signal Task -SST (Go/no-go tasks) (Logan, 1994)	The child is asked to respond as quickly as possible to visual stimuli and try to inhibit their responses when an auditory tone is heard (a forced-choice reaction time task)	<p>Behavioral inhibition/Response time: Ability to stop –suddenly and completely a planned or ongoing thought or action. This ability is measured principally by the Stop Signal Reaction Time (SSRT), an estimation of the time the child needs to stop their usual behavior in response to a trained stop signal.</p>

Note: From Garcia-Barrera, M. A., and Kamphaus, R. W.,(2006) Attention Deficit Hyperactivity Disorders (Chapter 12). In R. W. Kamphaus, and J. M. Campbell, Eds. *Psychodiagnostic assessment of Children. Dimensional and Categorical Approaches*. Wiley. Authorization for reproduction in process.

Alternative approaches

Given the number of limitations in the assessment of executive functions (Lezak, 1982; Mahone et al., 2002), new alternative measurement paradigms have been appearing in the literature. For instance, a different approach was recently introduced by Manchester and colleagues (2004), who reviewed the shortcomings of the classic methods and recommended focusing on naturalistic observations of behavior. Another paradigm, more neuropsychological-oriented, recommends an assessment based on the analysis of lower-order processes believed to

be related to components of executive function (V. Anderson, 2001; Hughes & Graham, 2002; Miyake, Emerson et al., 2000; Miyake, Friedman et al., 2000). In this approach, theoretically and empirically known executive “functions” are assumed as latent constructs (e.g., inhibition), and thus broken down into components. These “components” are assessed via “pure” tasks that taken together serve as indicators of the latent construct (e.g., antisaccade task, stop-signal task, and Stroop task). This approach has been demonstrated to be useful in latent variable analysis of behavior, given that SEM method controls for the introduction of cumulative effect of variances into the latent construct, which takes care of the potential harm of method variances. These types of constraints are necessary given the low correlation among variables as noted earlier, differentiating the way that executive function is psychometrically analyzed, unlike other higher functions (e.g., intelligence) in which stronger correlations among measures allow for the creation of factors, which can be introduced as latent constructs into a theoretical model (e.g., Spearman’s *g* in Carroll’s model).

However, the most interesting new approach has arisen from researchers’ concerns about the ecological validity of executive function tasks. Ecological validity refers to the extent to which results obtained on a controlled standardized test generalize onto performance in naturalistic settings (Chamberlaine, 2003). In this regard, Burgess, Alderman, Evans, Emslie, and Wilson (1998) studied the ecological and discriminant validity of a set of tests believed to assess executive functions. This group of authors recognized that the concern is not new, and they listed some studies that have examined the issue of transferring test results into the real world (e.g., maze performance and driving ability –Sivak et al., 1981; trail making and matching numbers tasks and academic achievement –Naglieri and Das, 1987; among others cited by Burgess et al, 1998). Some executive tests have been created to be inherently ecologically valid, such as Boyd

and Sautter's route finding test (1993, cited by Burgess et al, 1998) or Shallice and Burgess's Multiple Errands Test (1991, cited by Burgess et al, 1998).

In this last *ecological* direction, yet another psychometric approach has been focused on the analysis of the everyday-behavioral components of executive functions, and is recognized for the development of behavioral ratings of frontal and executive functions; examples are the Behavior Rating Inventory of Executive Function –BRIEF (Gioia et al., 2000), the Frontal Systems Behavior Scales –FrSBe (Grace & Malloy, 2001), the Frontal Behavioral Inventory –FBI (Kertesz, Davidson, & Fox, 1997), and the rating included on the Behavioural Assessment of the Dysexecutive Syndrome-BADS, The Dysexecutive Questionnaire –DEX (Wilson et al., 1998).

Behavioral Assessment of Executive Functions in Children

Table 2 presents a comprehensive summary of a recently published review of rating scales for measuring frontal lobes/executive functioning by Malloy and Grace (2005) and other sources (Chamberlaine, 2003; Gioia et al., 2000; Wilson et al., 1998). As can be observed in Table 2, the BRIEF is the only rating measure reported in the literature, for the ecological and behavioral assessment of executive functions in children and adolescents. For this reason it is the only scale that will be presented in detail.

Behavior Rating Inventory of Executive Function –BRIEF (Gioia et al., 2000)

The BRIEF is an 86-item questionnaire for parents and teachers of school age children (5-18 years old). It was designed to provide a better understanding of a child's self-control and problem-solving skills by measuring eight aspects of executive functioning related to behaviors in both home and school settings. It is comprised of eight clinical scales (Inhibit, Shift,

Table 2

Comparison across three well known frontal lobe/executive functions rating scales

Scale	Description	Reliability	Validity	Norms
BRIEF	An 86-item questionnaire for parents and teachers of children 5-18 years old. It was designed to provide a better understanding of a child's self-control and problem-solving skills by measuring 8 aspects of EF related to behaviors in both home and school settings. It is comprised of 8 clinical scales (Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor) believed to represent different components of EF. Factor Analysis yielded 8 scales that combine to form two indexes (Behavioral Regulation and Metacognition) and a total composite (Global Executive Composite). In addition, this rating scale offers two validity scores: Inconsistency Scale and Negativity Scale	High internal consistency Alpha = 0.80-0.98. High test-retest reliability, $r = 0.81$ for parents and $r = 0.87$ for teachers. Low interrater reliabilities (mean $r = 0.32$)	Content: a panel of experts judged the items. Concurrent: demonstrated correlation with other rating scales for children. Factor analysis was used to support the two-factor model of executive functions proposed.	Child ratings from 1419 parents and 720 teachers from rural, suburban, and urban areas.
DEX	This rating scale is part of the BADS, an assessment battery designed to evaluate everyday problems presented as part of the "Dysexecutive Syndrome." The DEX includes 20 items, used as a supplement to the BADS, in an effort to add ecological validity to the instrument. Each item is scored on a five-point (0-4) Likert scale. It comes with two forms, a self-report for the patient and one for the caregiver.	Authors do not report reliability coefficients for the DEX in the Manual. Other sources were consulted (Chamberlaine, 2003; Wilson et al., 1998) but with unsuccessful results.	Factor analysis is presented to support three areas: Behavior (8 items), Cognition (5 items), and Emotion (3 items). BADS' measures correlated with all DEX factors except for Emotion. Some studies have utilized the DEX for diagnostic purposes. Patients were rated at clinical levels on the questionnaire, as expected. However, the instrument did not correlate well with other frontal lobe ratings.	216 normal subjects and 78 patients (ages 16-64) with diverse clinical diagnoses (e.g., closed head injury, encephalitis, dementia, and stroke).
FBI	Designed to optimize diagnostic accuracy for frontal lobe dementia, rating changes in behavior via caregiver interview. 24-items represent both negative and positive	Interrater reliability (Cohen $\kappa = 0.90$) Item-consistency (Cronbach Alpha = 0.89)	Two validation studies. One is a comparative analysis of the scale on groups of patients with FTD (n=12), Alzheimer's (n=16), and depressive dementia (n=11). Good discrimination was found.	Norms are not available.

FrSBe	<p>symptoms associated with frontotemporal dementia (e.g., apathy, inflexibility, inattention, disorganization, and aggression). A 46-item behavior rating scale, aimed to be a brief, reliable, and valid measure of the effects of frontal systems damage in adults, before (premorbid baseline) and after they occur. It includes two forms (self-rating and family/caregiver rating), with 3 scale scores (Apathy, Disinhibition, and Executive Dysfunction, and a total composite. T-scores are available for results analysis.</p>	<p>Authors report high internal consistency for each scale and total as measured by Cronbach's Alpha (Total, 0.92; Apathy, 0.78; Disinhibition, 0.80; and Executive Dysfunction, 0.87) for the family form, and similar results on the self-report (Total, 0.88; Apathy, 0.72; Disinhibition, 0.75; and Executive Dysfunction, 0.79). External studies results were consistent with those presented in the manual.</p>	<p>On a second study, the scale was administered to a larger sample (N=108) and a discriminant function of 92.7% for FTD was reported.</p> <p>Construct validity: A study by Grace et al (as cited by Maloy & Grace, 2005) with 24 FL injured patients, 15 without FL damage, and 48 healthy controls, in which significant differences across scores in the groups were found, and significant scores of behavior change before-after damage on the FL group were observed as well.</p> <p>Convergent validity: Norton et al (as cited by Maloy & Grace, 2005) found significant correlations between NPI and FrSBe in 30 dementia patients and their caregivers.</p> <p>Discriminant validity, ecological validity and other analysis are reported in Malloy and Grace (2005) and in the manual.</p>	<p>436 men and women (18-95 years of age and education level from 10 to doctoral level). T-scores provided by age, gender, and level of education.</p>
-------	--	--	--	--

Abbreviations notes: BRIEF = Behavioral Rating Inventory of Executive Functions (Gioia et al., 2000), DEX = Dysexecutive Questionnaire (Wilson et al., 1998), BADS = Behavioural Assessment Battery of the Dysexecutive Syndrome (Evans, Chua, McKenna, & Wilson, 1997; Wilson et al., 1998), FBI = Frontal Behavioral Inventory (Kertesz et al., 1997), FrSBe = Frontal Systems Behavior Scale (Grace & Malloy, 2001), NPI = Neuropsychiatric Inventory, EF= Executive functions, FL = Frontal lobe.

Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor) believed to represent different components of executive functioning. Products of a factor analysis, the eight scales combine to form two indexes (Behavioral Regulation and Metacognition) and a total composite (Global Executive Composite). In addition, this rating scale offers two validity scores: Inconsistency Scale (for inconsistencies in patterns of response) and Negativity Scale (measures the extent to which the respondent answers selected BRIEF items in an unusually negative manner relative to the clinical samples).

Authors based the development of this instrument on a combination of theory, clinical experience and literature review. They reported in the manual to be concerned about maintaining content and construct validity. First, authors reviewed research on executive function conceptualization, including development and assessment. They also consulted with experts in the field about executive function components. Nine components were included: initiate, sustain, inhibit, shift, organize, plan, self-monitor, working memory, and emotional control. Factor analysis yielded two major components or meta-domains (Metacognition and Behavioral Regulation).

In order to assess these components, the authors reviewed their clinical databases in the search for common symptoms and concerns among parents and teachers that were previously collected during clinical interviews. Based on their experience and expertise, in addition to a thorough review of current behavioral scales (including Reynolds and Kamphaus' BASC) they created a pool of 180 items aimed to cover and represent all nine- executive function components. After revision for reading and content issues, several items were eliminated and pools of 127 items for the teacher form and 129 items for the parent form were tested for readability (via RightWriter software version 4.0).

Before the ratings tryout, the items were tested for item-scale membership. Authors and a panel of experts rated each item's fit into one of the categories provided, and best fit agreement determined item membership. No items were deleted. Pilot data were collected from both a clinical population and a normative population. Analysis of reliability was performed with SPSS, generating a matrix correlation for each item with total score within each scale. Items with the lowest correlation were to be removed following a stepwise fashion to test internal consistency; however, no items were eliminated. A principal factor analysis with orthogonal rotation was used to test the structure of the scales and item membership. Furthermore, an inter-item correlation matrix was generated and studied, in addition to analysis of inter-rater agreement results.

For the final scale development, analyses of the tryout results were taken into account without eliminating any item. A pilot standardization procedure was performed as well as new statistical analyses. Four scales were correlating too high; therefore, the authors concluded that they were tapping the same constructs. Thus, items were combined into a larger pool for further analysis. A new factor analysis yielded two more internally consistent and stronger factors and a final set of 8 scales was established. Two more scales to assess validity were developed, the inconsistency scale and the negativity scale.

Given its relative novelty, fewer studies have included the BRIEF in their protocols for comparison with the WCST and other traditional measures. Although it is believed to add "ecological validity" while predicting the child's executive functioning in the everyday environment –versus a laboratory or clinical setting- (Gioia & Isquith, 2004; Gioia, Isquith, Kenworthy, & Barton, 2002), issues associated with the general shortcomings of ratings scales are also present in the literature (M. B. Denckla, 2002). Authors reported the utility of the BRIEF in identifying salient executive characteristics and dysfunctions on diverse developmental

disorders (ADHD, Traumatic Brain Injury, and Autistic Spectrum Disorders); they are cautious about the diagnostic limitations and the need to integrate this instrument into a broad assessment battery (Gioia, Isquith, Kenworthy et al., 2002). External studies have inconsistent results; for instance, one study documented its clinical utility in executive functions assessment for individuals with myelomeningocele and hydrocephalus (Kline, 2005; Mahone et al., 2002), whereas another discussed possible limitations of the instrument for measuring behavioral executive dysfunction in neurologically mildly affected cases (V. Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002).

Towards new alternatives in the assessment of executive functions

The *Behavior Assessment System for Children* –BASC (Reynolds & Kamphaus, 1992) is a multidimensional rating scale of externalizing, internalizing, and adaptive skills that includes questionnaires for parents and teachers of children and adolescents between the ages of 2 and 18. The validity of the BASC as an assessment tool for frontal lobe/executive function has been previously studied (Jarratt et al., 2005; Mahone et al., 2002; Riccio et al., 1994). Barringer and Reynolds developed a frontal lobe functioning scale from the BASC and presented it as a paper at the annual meeting of the National Academy of Neuropsychology in 1995 (Reynolds & Kamphaus, 2002). Barringer and Reynolds identified a subset of items from the BASC, which were hypothesized as being associated with frontal lobe/executive function using an expert approach by surveying editorial board members of three leading clinical journals in neuropsychology. A final set of 18 items produced a coefficient Alpha of .84, and comparison of various clinical groups showed high levels of discrimination (Reynolds & Kamphaus, 2002). Although Barringer and Reynolds' analysis included only the Parent Rating Scales, parallel forms of this set of items can be identified across different rating scales from the BASC,

including the Teacher Rating Scales for both age groups, children (6-11 years old) and adolescents (12-18 years old). Barringer and Reynolds's scale is useful for evaluating behaviors generally associated with frontal lobes functioning. Most of the items they selected from the BASC relate to executive *control* functions, specifically, inhibitory control, emotional control, and attentional control. Items related to other behavioral executive functions such as organization, planning, and decision making were not included.

Recently, the BASC-2 (Reynolds & Kamphaus, 2004) included the Frontal Lobe/Executive Control Scale (FLEC) as a supplemental scale derived from the early work of Reynolds and Barringer. Although promising, this scale has little known validity, and only one recent study by Sullivan and Riccio (2006) examined its convergent validity as a tool in the diagnosis of ADHD. Sullivan and Riccio utilized the scale score as a composite, and it demonstrated significant correlations ($p < .001$) with the BRIEF-parent-form's eight scales (Inhibit $r = .69$, Shift $r = .70$, Emotional Control $r = .66$, Initiate $r = .66$, Working Memory $r = .69$, Plan/Organize $r = .73$, Organization of Materials $r = .45$, Monitor $r = .71$), its two indexes (Behavioral Regulation $r = .80$, Metacognition $r = .76$), as well as its composite (Global Executive Composite $r = .83$). Significant correlations ($p < .001$) were also found with the Conners' Parents Rating Scale Short Form (Oppositional $r = .77$, Cognitive Problems/Inattention $r = .68$, Hyperactivity $r = .68$, and the ADHD Index $r = .63$) (Sullivan & Riccio, 2006). Finally, for the purposes of this study and given the large amount of behaviors that the BASC items measure, it is believed that a broader instrument (e.g., a screener) for the estimation of executive functions related to behavioral self-regulation in school-aged children can be developed.

Conclusions

Statement of the problem

Attempts to measure complex cognitive functions have historically generated controversies, and executive functions are not an exception. As it was presented here, multiple models have tapped different components of executive functions, given the great difficulty to operationally define them all together. Unlike Spearman's *g*, executive functions is a conceptual umbrella that groups elements that correlate poorly between each other, making psychometric analyses fail in their efforts to create a one *unitary* yet *diverse* model of the construct. However, this type of limitations should be encouraging more than disappointing, and researchers should combine efforts to approach this complex construct from newer perspectives, as long as needed, until a reasonable consensus is established.

In this direction, this research project aims to add a new angle, a new perspective to the bulk of research on executive functions and their measurement. Understanding how executive functions manifest themselves via regulation of everyday behavior is of great importance in fields such as pediatric neuropsychology and school psychology, but has been a research area poorly explored. The introduction of the BRIEF and the series of studies that followed it, pointed out many new directions to be explored in this line of research. Among them, both the analysis of the validity of such ratings and the creation of other measurement instruments appear imperative. Therefore, the ultimate goal of this project is the development of a new instrument for the screening of executive functions in school-aged children.

Purposes and significance of the study

Potential impact of the results of this research includes the following:

1. Development of a reliable and valid screener would provide the community of school psychologists, child psychologists, pediatricians, and other related professionals with a tool for the estimation of executive functions that can potentially be used for the early identification of children in need of academic and/or behavioral interventions.
2. Since the BASC has been standardized and validated in its Spanish form in the US, Spain, and Colombia, equivalent executive functions screeners can be extracted from these forms and tested for cross-cultural validation in Spanish-speakers, offering a bilingual instrument for the screening of children from Hispanic background.
3. Publication of this type of study may help establish foundations for the development of more screening and rating scales to complement traditional assessment methods of cognitive functioning.

Research Questions and Hypotheses

This project includes the development of a screener for the assessment of executive function, guided by two main questions:

1. *Is it possible to obtain a reliable screening measure of children's executive functions with behavioral rating scales?* Given previous results of the pilot analyses of Barringer and Reynolds (1995, cited by Reynolds & Kamphaus, 2002) it is hypothesized that a reliable measure to estimate executive functions related to initiation of behavior (problem solving, planning, decision making) and behavioral self-regulation (attentional control, behavioral control, and emotional control), can be obtained from a large behavioral rating scale such as the BASC.
2. *Is it possible to obtain a valid screening measure of children's executive functions with behavioral rating scales?* Given previous results of the empirical analyses of Sullivan and

Riccio (2006), it is hypothesized that this instrument's construct validity using a theoretically driven Confirmatory Factor Analysis, and Multiple Groups Analysis of Invariance, would yield a stable structure of factors that relate to executive self-regulating functions.

CHAPTER 3

METHODS

Original Instrument

The *Behavior Assessment System for Children* (BASC; Reynolds & Kamphaus, 1992) is a multidimensional rating scale of externalizing, internalizing, and adaptive skills that includes questionnaires for parents and teachers of children and adolescents between the ages of 2 and 18. The BASC rating scales are among the most widely used measures of child behavior in the United States, where it is ranked as either first or second in usage surveys (Pineda et al., 2005; Reynolds & Kamphaus, 2002; Sandoval & Echandia, 1994).

The *BASC Teacher Rating Scale-Child (6-11 years)*, or BASC- TRS-C (Reynolds & Kamphaus, 1992), is a 148-item, nationally standardized measure that yields nine problem behavior scales (Aggression, Conduct Problems, Hyperactivity, Anxiety, Depression, Somatization, Attention Problems, Learning Problems, and Atypicality) and four adaptive skills scales (Adaptability, Leadership, Social Skills, and Study Skills) as well as composite (or summary) scores. The 148 items are rated using a four-point Likert scale, anchored by 1 (never) and 4 (almost always) (Reynolds & Kamphaus, 1992).

Participants and Data Collection

Pearson Assessment gave the author permission to access and to analyze the original BASC standardization database for the TRS-C. This data was collected in two time periods, between the fall of 1988 and spring of 1991, and winter of 1997 through the spring of 1998. The standardization of the BASC included a vast sample representative of the population of U.S.

children, and it includes a range of diversity in terms of ethnicity, socioeconomic status, geographic region, and clinical problems (Reynolds & Kamphaus, 1992).

Data was collected in a total of 116 testing sites across the nation, using a school-based method, which allows for the inclusion of controlled representation of sub-samples (e.g., special education) in proportions that occurred in regular classrooms. For the purposes of this study, and to increase statistical power and generalizability of the results, both general-population norm and clinical samples were aggregated and analyzed as one sample. Table 3 reports the demographic characteristics of the sample, including age, gender, race, and geographic distribution. This sample distribution mirrored the population distribution for the general US population, with the exception of gender. There is a 20% discrepancy between male and female sub-samples, which was corrected in the original BASC standardization sample; for the purposes of this study, the distribution was held invariant since it is expected that a small sub-sample size discrepancy will not have a significant effect on the analysis. Table 4 includes the representation of the special education placement in the general-norm sample. This classification is based on the Special Education categories as defined by the Individuals with Disabilities Education Act (IDEA, 1990). The learning disability category represents the largest proportion of cases, with an overrepresentation of male versus female cases. This gender disproportion is consistent across different categories, and it follows the nation's special education placement statistics. Table 5 reports the demographic distribution of the clinical norm sample. Table 6 includes the representation of the clinical norm sample by primary diagnosis or classification. In this table, diagnoses of behavioral and emotional disorders were based on special education placement. Finally, Table 7 reports the representation of the integrated sample by developmental age. All the tables are presented in percentages (adapted from Reynolds & Kamphaus, 1992).

Table 3

Demographic representation of the general-norm sample (%)

Gender		Race				Geographic region			
Female	Male	African-American	Hispanic	White	Other	Northeast	North Central	South	West
40	60	24	7	67	2	8	35	36	21

Table 4

General norm sample by special education classification and gender (%)

Behavioral/Emotional Disturbance		Learning Disability		Mild Mental Retardation		Speech/Language Disorder	
Female	Male	Female	Male	Female	Male	Female	Male
0.3	0.9	2.6	5.9	0.8	1.2	1.6	1.9

Table 5

Demographic representation of the clinical-norm sample, N = 393 (%)

Gender		Race				Geographic region			
Female	Male	African-American	Hispanic	White	Other	Northeast	North Central	South	West
24	76	11	2	86	1	0	78	19	3

Table 6

Representation of the clinical norm sample by primary diagnosis or classification (%)

Behavior Disorder	Conduct Disorder	ADHD	Depression	Autism	Emotional Disturbance	Undifferentiated	Other
47	4	17	3	3	3	16	6

Table 7

Representation of the aggregated sample by developmental age (%)

6 years old	7 years old	8 years old	9 years old	10 years old	11 years old
11	16	17	23	17	16

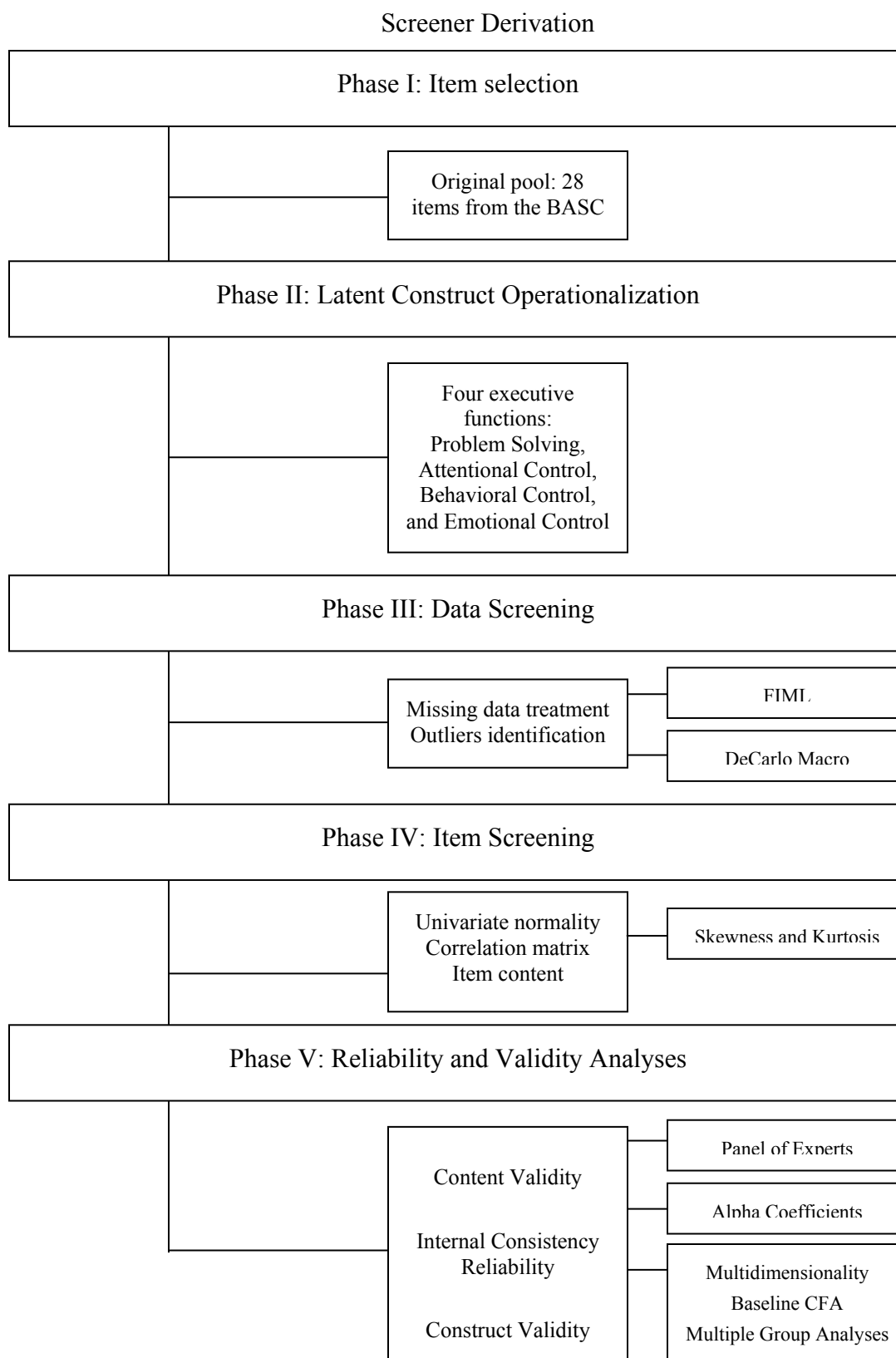


Figure 4, Screener derivation process model

Phase I: Item selection pool

For the development of this *executive functions* screener, all 148 items from the original BASC-TRS-C were carefully reviewed by the first author, with the aim of identification of items assessing behaviors that can be potentially classified as "executive." Thus, an initial set of 28 items believed to assess behaviors associated with executive functioning was extracted. Given the assumption that the executive function definition conveys both unitary and diverse elements, one of the goals for the screener development was to determine if it should be unidimensional or multidimensional. As presented earlier, Elliot (2003) asserted that "there is no intuitive lay concept that incorporates the essence of executive function" (p. 50), and yet, by consensus, it will be assumed here that such a self-regulatory yet complex executive system encompasses a set of coordinated cognitive abilities required in the planning, monitoring, and execution of behaviors oriented towards a goal (V. A. Anderson et al., 2002; Elliot, 2003; Lezak et al., 2004; Royall et al., 2002). Moreover, and following Baddeley's (1996) call for unifying efforts to better understand the processes involved in executive functions, the task of this screener derivation was to identify items that could serve for the estimation of one, two, or more of these processes.

Four hypothetically and theoretically-driven clusters of items were derived from the original 28-item set. In this regard, the items function as observed variables that, when grouped, form factors representing constructs. The items themselves are only *indicators* of the construct, which is *latent* in nature. For the purpose of the screener development, it was hypothesized that the 28 BASC items could be grouped into four factors (or "*scales*" in measurement) that represented four essential executive functions: Problem Solving, Attentional Control, Behavioral Control, and Emotional Control. Specific abilities included in these four categories have been

documented to be very important components of executive functions. Some of these components have also been classified as “executive” given their “controlling” capacity over other functions such as emotion, attention, and inhibition of automatic behavior (e.g., Gioia et al., 2000; Royall et al., 2002).

Phase II: Latent construct operationalization and item distribution.

(1) The *Problem Solving* construct measures the ability to plan, problem solve, make decisions, and organize information towards the execution of a goal. It includes 10 items from i1 to i10, such as “*analyzes the nature of a problem before starting to solve it*” and “*makes decisions easily.*” It is hypothesized that these items can be identified as problem solving strategies, under the assumption that, first, they allow one to estimate skills required in goal attainment, and second, that achieving a “goal” is equivalent to problem solving. In other words, this construct assesses the question: *does this child demonstrate to know how to achieve a goal?*

(2) The *Attentional Control* construct includes seven items (i11 to i17) and measures the ability to focus (e.g., *makes careless mistakes*), sustain (e.g., *has trouble concentrating*), and shift attentional systems according to task demands (e.g., *has trouble shifting gears from one task to another*). One memory/working memory item was included (*forgets things*) to include an estimation of the effects of attentional problems in memory and therefore, in learning. These three elements of attention have been derived from Mirsky et al.’s (1991) model, but under the assumption that executive functions do have a regulatory role over these attentional systems, as was presented earlier by Norman and Shallice’s *Supervisory Activating System (SAS)*. Ratings within this factor attempt to answer the question: *can this child self-regulate attention?*

(3) The *Behavioral Control* construct is comprised of seven items (from i18 to i24), and was included in this screener under the assumption that executive functions play an important role in

the ability to self-regulate behavior (e.g., *acts without thinking*), including inhibition (e.g., *uses foul language*) and impulse control (e.g., *interrupts others when they are speaking*). This assumption is supported by the work of Barkley (1997) in ADHD. This construct answers the question: *can this child inhibit prepotent response and self-regulate behavior?*

(4) Items under the *Emotional Control* construct measure the ability to self-regulate emotional response to environmental and internal cues. This ability is of great importance when trying to examine the relationships between frontal/prefrontal areas and the limbic system, in terms of the regulatory function of executive systems over emotion. This factor is comprised of four items (i25-i28), representing a limited yet significant range of emotional dysregulation indicators (e.g., *changes moods quickly* and *throws tantrums*). This construct answers the question: *can this child self-regulate emotional expression?* Table 8 summarizes the items, their original distribution by factor, and the variable identification label.

Table 8

Original distribution of items per scale

Final BASC	Model	Factors / Item	
			<i>Problem Solving</i>
i37	I1	Analyzes the nature of a problem before starting to solve it	
i50	I2	Does extra credit	
i60	I3	Makes decisions easily	
i64	I4	Works hard, even in courses he or she does not like	
i83	I5	Gives good suggestions for solving problems	
i110	I6	Makes suggestions without offending others	
i111	I7	Asks to make up missed assignments	
i120	I8	Works well under pressure	
i124	I9	Has good study habits	
i148	I10	Is well organized	
			<i>Attentional Control</i>
i10	I11	Does not complete tests	
i18	I12	Is easily distracted from classroom	
i24	I13	Makes careless mistakes	
i55	I14	Has a short attention span	
i78	I15	Has trouble concentrating	
i115	I16	Forgets things	
i126	I17	Has trouble shifting gears from one task to another	

			<i>Behavioral Control</i>
i2	I18	Argues when denied own way	
i39	I19	Breaks other children's things	
i59	I20	Acts without thinking	
i82	I21	Interrupts others when they are speaking	
i94	I22	Uses foul language	
i127	I23	Hits other children	
i144	I24	Cannot wait to take turn	
			<i>Emotional Control</i>
i7	I25	Stays disappointed a long time if a favorite activity is cancelled	
i44	I26	Changes moods quickly	
i95	I27	Is easily upset	
i108	I28	Throws tantrums	

Phase III: Data Screening

Missing data treatment.

Typical missing data techniques include listwise and pairwise deletion, mean imputation, regression-based imputation, and hot-deck imputation (Enders & Bandalos, 2001). The validity of data imputation techniques has been questioned due to their tendency to produce biased parameter estimates. For instance, means imputation (i.e., missing values are replaced with the mean of the data from complete cases) not only produces underestimated variances and covariances but also biased parameters. Regression-based imputations (i.e., values are predicted based on values from other variables) are complex and require higher correlations among variables. Listwise and pairwise deletion techniques have been examined in the context of SEM and CFA and it has been observed that they are not only wasteful but also that they have a tendency to produce more biased estimates (Enders & Bandalos, 2001).

More recently, a Full Information Maximum Likelihood (FIML) approach has been presented. This method does not impute missing values but rather estimates parameters and standard errors by “borrowing information from the observed portion of data,” and by computing a “casewise likelihood function using only those variables that are observed for [the] case” with missing data (Enders & Bandalos, 2001; p. 434). FIML uses the following formula:

$$\log L_i = K_i - \frac{1}{2} \log |\Sigma_i| - \frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_i)' \Sigma_i^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_i) \quad (1)$$

in which K_i is a constant dependent of the complete data obtained for case i , and \mathbf{x}_i is the observed data. Mean vector and covariance matrix for the complete variables in case i are contained in $\boldsymbol{\mu}_i$ and Σ_i .

Enders and Bandalos (2001) performed a Monte Carlo simulation study in which they demonstrated the relative unbiased and efficient performance of FIML with respect to other missing data techniques. FIML was also superior in proportion of convergence, lower model rejection rates, and lower risk for Type 1-error. Thus, FIML was used to treat missing data in this study.

Outliers

Tests for multivariate normality and significant outliers were performed with a macro written by DeCarlo that can be run on SPSS (DeCarlo, 1997). The advantage of this macro is that it outputs critical values to test the significance of the Mahalanobis distances, providing a more objective tool to determine the outliers' deletion or inclusion in the analyses.

Phase IV: Item Screening.

This set of items is hypothetical, in other words, item analysis procedures are necessary in order to identify the best set of items for the screener purposes. Since the data has already been collected, the following steps are applicable to the item analysis procedure:

1. Analysis of frequency distributions: including analysis of means, standard deviation, skewness, and kurtosis. Standard deviations allow the identification of the variation in responses to items. Measures of skewness indicate whether the item's distribution deviates significantly from symmetry, with a value of 0 being a perfect symmetry. A critical value of $|2.0|$ was used to test skewness (Crocker & Algina, 1986). Kurtosis is a

measure of the degree to which the area in a distribution is primarily in the middle and the tails of a distribution. A value of 0 represents a perfect normal distribution. There is some debate about the adequate critical value to determine significant kurtosis. The BASC items use a 4-point Likert scale. It has been suggested that scales dealing with ordinal (non continuous) data such as Likert often present a nonnormal distribution, due to the nature of the items' content. For this study, values outside the range of ± 7.0 are problematic and represent considerable depart from normality, due to the risk of bias that univariate nonnormality presents (Cordon & Finney, 2006; Finney & DiStefano, 2006).

2. Generation of a correlation matrix to analyze the relationships among items and potential collinearity effects. If correlations are too high, collinearity is suspected (Kline, 2005).
3. However, since all the items in the scale are hypothetically expected to measure different aspects of the same construct, moderate correlation values are expected.
4. Examination of item content and wording. In this case, items that are suspicious of being too ambiguous are eliminated.

Phase V: Reliability and Validity Analysis

Content validity

A panel of 15 experts in neuropsychology and behavioral measurement were first informed about the purposes and goals of the screener, and subsequently, were asked to classify the 28 items into the four hypothesized executive components, in terms of their accuracy to measure problem solving, attentional control, behavioral control, and emotional control. The items were alphabetically arranged (by the first word) to keep the raters blind to the author's original classification. The item survey used is attached as an appendix (see Appendix A). Ten

(66.66%) of the surveys were returned and inter-rater reliability was calculated in terms of the percentage of agreement. A cut-off of 70% inter-rater agreement was established as a criterion.

Internal Consistency Reliability

Coefficient Alpha for each group of items (factors) is a way of estimating internal consistency reliability (Crocker & Algina, 1986). It is computed as:

$$k / (k-1) [1 - \sum \sigma^2_i / \sigma^2_c] \quad (2)$$

in which k is the number of items, σ^2_c is the variance of the whole test, and $\sum \sigma^2_i$ is the sum of the individual item variances. A homogenous set of items would produce higher coefficients.

The possible range goes from 0.00 to 1.00, with values closer to 1.00 being optimal, and 1.00 the maximum. According to Crocker and Algina (1986), when the purposes of the test are to measure a complex construct, lower values are expected.

Given the research question (*is it possible to obtain a reliable screening measure of children's executive functions with behavioral rating scales?*), it is hypothesized that a reliable measure to estimate executive functions such as problem solving, attentional control, behavioral control, and emotional control can be obtained from a large behavioral rating scale such as the BASC. SPSS 14.0 was utilized for the reliability analysis and to test the null hypothesis:

Ho: Cronbach's Alpha per Scale = 0.0 and Criterion Cronbach's Alpha for each scale > 0.8

Construct validity

Two well-supported methods of construct validation are included in this study: Confirmatory Factor Analysis and Measurement Invariance (or multiple group analysis). Both methods are performed in a SEM framework.

Confirmatory Factor Analysis (CFA). Construct validity refers to the degree to which a test can be considered to be an appropriate operational definition of a construct (Crocker &

Algina, 1986). For this study, a series of CFA was used to determine construct validity of the screener. Figure 5 presents the hypothetical CFA model to be tested.

This hypothetical model functions as a statistical model; therefore, directionality is assumed. In this case, the ellipses represent four latent variables believed to group as “self-regulatory” components of executive functions. The curved double-headed arrows represent relationships among all the latent constructs. Straight single-headed arrows represent factor loadings to be estimated. Although originally drawn for 28 indicators, the final model has 25 indicators [$25(26)/2 = 325$], 4 factor disturbances, and 52 free parameters to be estimated (21 factor loadings, 25 errors of measurement, and 6 correlations). It is hypothesized that the four factors will be correlated because they are components of the same construct. By allowing them to correlate, it is possible to specify the degree to which the four postulated executive functions are separable and share the same underlying ability or mechanism. CFA is the first choice over EFA because the aim is to observe the underlying factor model *a priori* (theoretically driven), instead of *a posteriori* (data driven). CFA models use the formula:

$$X = \lambda F + e \quad (3)$$

in which λF represents the weight of the variable into the factor, and e represents the error term (Kline, 2005).

These CFA procedures will allow us to answer the question: *Is it possible to obtain a valid measure of children's executive functions with behavioral rating scales?* It was hypothesized in this study that the instrument's construct validity using CFA (a priori) techniques would yield a stable structure of factors that are related to executive self-regulating functions. Given this hypothetical expectation, the next question to answer is if the CFA model is best fitted with one factor (*unidimensionality*) or more factors (*multidimensionality*). As it has

been presented earlier, it was hypothesized that this group of items would yield information about more than one component of executive functions; therefore, multidimensionality must be tested in contrast to a one-factor model. For this purpose, a series of models were run on Mplus 4.0 and tested for goodness of fit. These models are not nested, thus chi-square differences are only observed (L. K. Muthén & Muthén, 1998-2006).

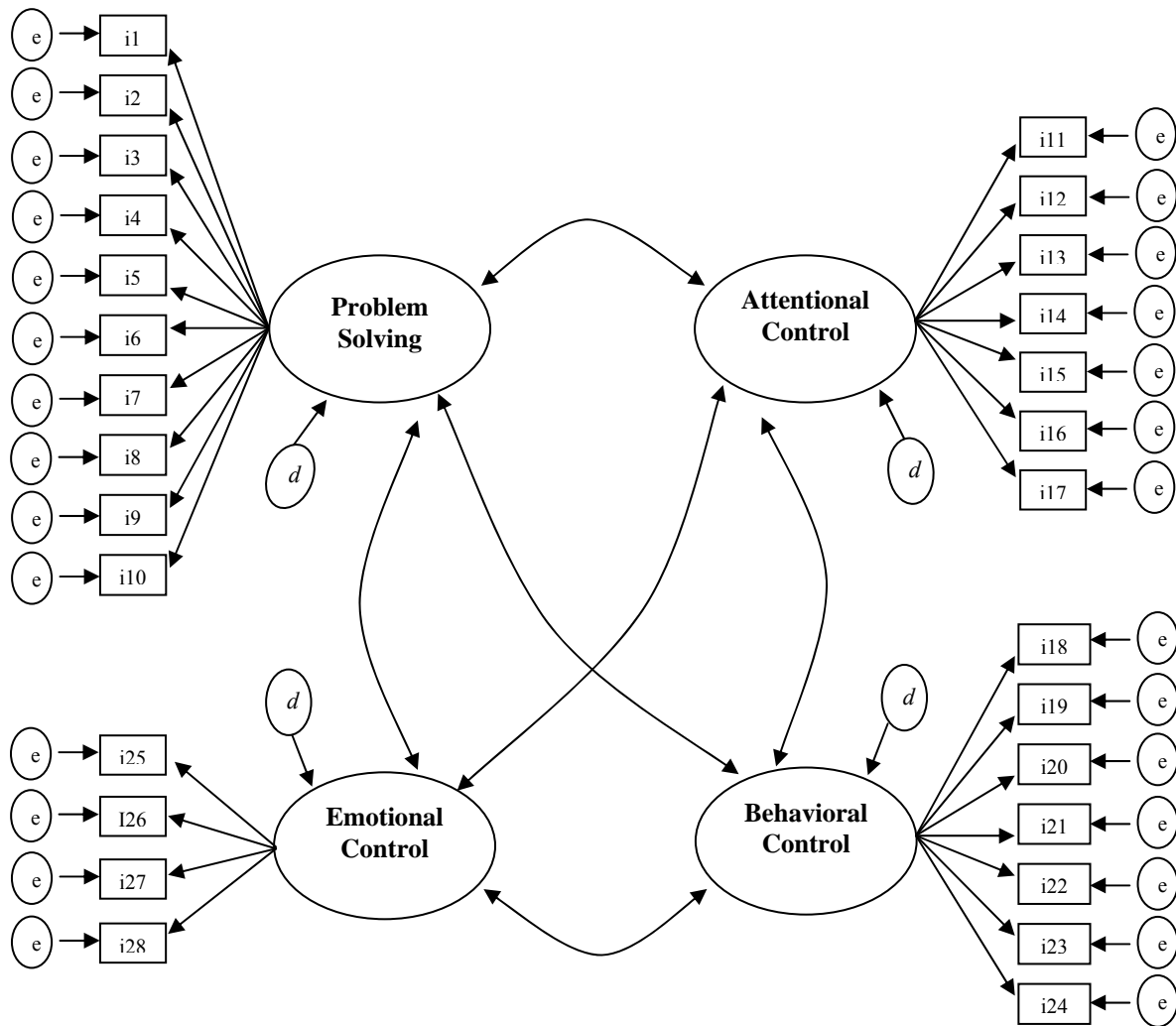


Figure 5, Hypothetical CFA model for the BASC Executive Functions Screener

Estimation Method

The Mplus 4.0 (L. K. Muthén & Muthén, 1998-2006) software was used to perform the CFA. SPSS 14.0 was utilized to produce a fixed ASCII format dataset (*.dat). Mplus utilizes a Maximum Likelihood (ML) method of estimation by default when data is continuous. This study utilizes items that are Likert-type (i.e., measuring latent constructs via observed responses that are distributed in a small number of ordinal categories from never (1) to almost always (4)). In these types of cases, it has been recommended to enhance validity by analyzing data as categorical (ordinal), rather than continuous (Flora & Curran, 2004; B. O. Muthén & Asparouhov, 2002).

Mplus default estimator for categorical data analyses is the Weighted Least Square with Means and Variance adjusted (WLSMV). According to Muthén and Muthén (1998-2006), “WLSMV [estimates parameters] using a diagonal weight matrix with standard errors and mean- and variance adjusted chi-square test statistic that use a full weight matrix.” Due to the vulnerability of other estimators for the common non-normality observed in categorical data, WLSMV adjustment of means and variances serves as a correction that is less computationally demanding, and produces estimates that are unbiased, consistent, and efficient. In contrast to other types of estimators (e.g., Weighted Least Squares -WLS), WLSMV uses the diagonal of the weight matrix in the estimation instead of the full weight matrix. Like WLS, WLSMV uses the full weight matrix to compute standard errors and chi-square (L. K. Muthén & Muthén, 1998-2006). There are two parameterizations provided by Mplus when using WLSMV as the parameter estimator: The Conditional Probability and the Latent Response Variable (LRV) formulations. LRV assumes that a continuous and latent response variable (y_i^*) underlies the observed categorical variable (y_i). Y_i is related to the latent response variable y_i^* via threshold

parameter (τ , or *intercepts*) between categories. For this purpose, polychoric correlations are calculated and used instead of Pearson's correlations (Flora & Curran, 2004). Figure 6 illustrates the model assumed by WLSMV (model adapted from Flora & Curran, 2004).

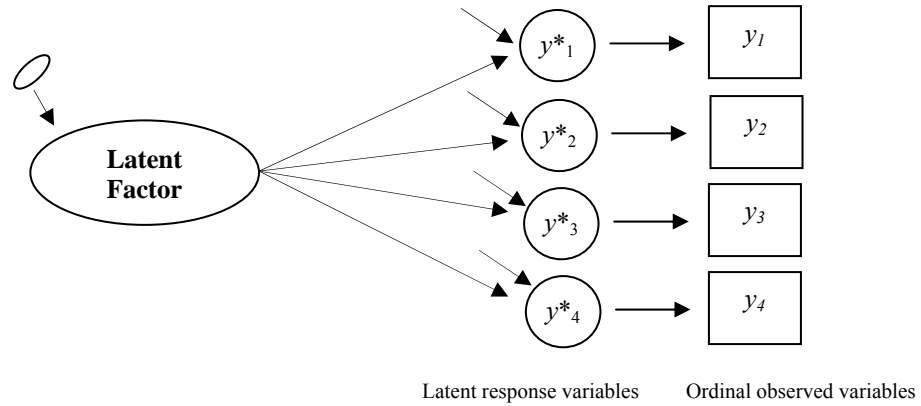


Figure 6, Model used by the WLSMV estimator under LVR formulation for ordinal data

Fit Indexes

There are two types of indicators of overall fit recommended by Hoyle and Panter (1995) and by Hu and Bentler (1999):

1. Absolute fit ("badness of fit") concerns the degree to which the covariances implied by the fixed and free parameters -and specified in the model- match the observed covariances from which free parameters in the model were estimated. Optimal fit is indicated by a value of zero. Hoyle and Panter (1995) recommended: χ^2 accompanied by df, sample size, and p -value. An adequate model is nonsignificant; therefore, a p value > 0.01 is used here as criterion.
2. Incremental fit ("goodness of fit") concerns the degree to which the model in question is superior to an alternative model, usually one that specifies no covariances among variables. Larger values indicate greater improvement of the model over an alternative.

Recommended: at least 2 Type-2 and Type-3 indexes (one from each). Tucker-Lewis Index (TLI) also known as Non-Normed Fit Index (NNFI) for Type-2, and Comparative Fit Index (CFI) for Type-3. A perfect, optimal value is 1.0, but is seldom obtained in practice (Cheung & Rensvold, 2002). A cutoff score of .95 for each is recommended by Hu and Bentler (1999), although some suggested that these are too high and that a cut-off of .90 is reasonable (Cordon & Finney, 2006). Due to these conflicting results, a range from .90 (lower bound) to .95 (optimal bound) is used as criterion in this study.

3. In addition, the Root Mean Square Error of Aproximation (RMSEA) will allow one to test the lack of fit of the sample data to the model. Values closer to zero are optimal. A cutoff score of .06 for RMSEA has been recommended by Hu and Bentler (1999), whereas others (e.g., Browne and Cudeck, 1993, as cited by Cordon and Finney, 2006) recommended values of .05 or less. For the purposes of this study, a range from .06 (lower bound) to .08 (optimal bound) is used as criterion for goodness of fit. Table 9 summarizes the fit indexes criteria used.

Table 9

Fit indexes and cut-off scores of goodness of fit examination

Fit Indexes		Criteria
χ^2		$p > 0.01$
H ₀ : $\sum = \sum(\theta)$		
Type 2	Optimal (Hu & Bentler, 1999)	Reasonable/lower bound (e.g., Cordon and Finney, 2006)
TLI	0.95	TLI ≥ 0.90
Type 3		
CFI	0.95	CFI ≥ 0.90
RMSEA	0.05	RMSEA ≤ 0.08

Abbreviations: χ^2 = Chi-square, H₀ Null Hypothesis, TLI = Tucker-Lewis Index, CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Aproximation.

Analyses of Invariance: Multiple Groups Approach (MGCFA)

The assessment of measurement invariance consists of a series of stepwise analyses to determine the extent to which items have equal meaning (or *stability*) across different groups of subjects (e.g., gender, ages, and races). This type of analysis enhances the examination of construct validity of the test because it provides evidence about construct-irrelevant invariance (Cordon & Finney, 2006; French & Finch, 2006). As French and Finch (2006) noted, new tests must meet current validity standards, including the analysis of invariance which limits the creation of biased measures. Invariance tests also facilitate better interpretation of true differences underlying latent constructs (e.g., Cordon & Finney, 2006).

There are three main steps of measurement invariance: Configural, Metric, and Scalar invariance tests. Configural invariance analysis allows identification of structural differences in the way that diverse groups interpret or conceptualize a construct. In other words, it tests the stability of the hypothesized factor structure across groups by evaluating if the same indicators (i.e., items) are associated with the same factors across groups (Rensvold & Cheung, 2001). Metric invariance analysis tests the hypothesis of equality of factor loadings:

$$\lambda_{ij}^{(1)} = \lambda_{ij}^{(2)} \quad (4)$$

It starts with the assessment of the whole model by adding constraints to the factor loadings and testing for the strength of the item-factor relationships across groups. If the model was found variant, examination of each factor (first) and then the items is recommended (Rensvold & Chung, 2001). Finally, the Scalar invariance assumes equivalence of intercepts across groups (Rensvold & Chung, 2001):

$$\tau_i^{(1)} = \tau_i^{(2)} \quad (5)$$

In other words, it sets constraints to the model in order to examine if respondents with equivalent values on the latent constructs have equivalent values on the items as well.

For the purposes of the screener, two multiple group analyses were performed. The first analysis includes the measurement invariance across sex (female vs. male). The second includes two groups, distributed by age: a “young” group, with children aged 6 to 8 years, and an “older” group, with children aged 9 to 11 years. It is hypothesized that the model will be held invariant across gender groups (i.e., there is not evidence supporting developmental sex-related differences in executive functions). It is also hypothesized that the model will be invariant across age (i.e., there may be developmental differences in level but not in the structure of the model of executive function itself).

Finally, a CFI difference (ΔCFI) ≤ -0.01 was the criterion used for the measurement invariance as recommended by Cheung and Rensvold (2002) and used in other studies (French & Finch, 2006; Kim, Cramond, & Bandalos, 2006).

CHAPTER 4

RESULTS

Phases I to IV: Data and Item-Level Screening

Prior to the identification of outliers, two cases were eliminated because they were missing 100% of the data. One more case was eliminated because the variable “sex” was unidentified. In addition, the DeCarlo macro (DeCarlo, 1997) identified five cases with significantly large Mahalanobis distances, which were eliminated as well. The overall effective sample size was 2,165 children. Further, the full dataset was assessed for univariate normality. These analyses included means, standard deviation, and values of skewness and kurtosis. Table 10 summarizes the results of these analyses. It is noted that most items’ skewness and kurtosis values are within criterion parameters for univariate normality ($[-2.0]$ and $[7.0]$, respectively), with the exception of item # 28, which has a value slightly over the cut-off score for skewness (2.2).

As can be observed in the table, all items are significantly correlated ($p < .01$), but correlations ranged from low to modest. Correlations ranged from a minimum of $-.10$ (between item # 3, “*makes decisions easily*” and item # 18, “*argues when denied own way*”) to a maximum of $.76$ (between item # 12, “*is easily distracted from classroom*” and item # 14, “*has a short attention span*”). Items were arranged in order from Factor 1 (Problem Solving) to Factor 4 (Emotional Control). Notice that the farther apart they are, the lower the correlations, and vice versa. Both items measure attentional control and specifically, both items assess the capacity to sustain attention on relevant information. Table 10 includes the data distribution and correlation matrix.

Table 10

Data distribution and correlation matrix

	i1	i2	i3	i4	i7	i9	i10	i11	i12	i13	i14	i15	i16	i17	i18	i19	i20	i21	i22	i23	i24	i25	i26	i27	i28
i1	1	.53	.54	.54	.48	.59	.52	-.29	-.48	-.39	-.50	-.45	-.33	-.37	-.19	-.19	-.36	-.23	-.21	-.27	-.17	-.20	-.19	-.21	-.21
i2	.63	1	.49	.53	.56	.59	.53	-.29	-.44	-.31	-.46	-.41	-.34	-.34	-.19	-.19	-.31	-.20	-.18	-.25	-.17	-.19	-.14	-.19	-.16
i3	.63	.58	1	.49	.42	.55	.49	-.33	-.43	-.33	-.47	-.43	-.35	-.38	-.10	-.16	-.29	-.15	-.11	-.16	-.12	-.19	-.16	-.24	-.18
i4	.62	.63	.56	1	.57	.73	.63	-.42	-.60	-.48	-.61	-.57	-.40	-.50	-.37	-.32	-.47	-.37	-.34	-.38	-.26	-.30	-.32	-.28	-.31
i7	.57	.66	.49	.65	1	.63	.54	-.27	-.44	-.32	-.47	-.42	-.32	-.33	-.28	-.24	-.35	-.29	-.26	-.33	-.17	-.23	-.19	-.22	-.24
i9	.68	.69	.61	.80	.72	1	.75	-.44	-.65	-.50	-.67	-.63	-.48	-.51	-.35	-.29	-.50	-.39	-.32	-.37	-.27	-.30	-.30	-.30	-.30
i10	.59	.61	.55	.70	.61	.81	1	-.38	-.65	-.50	-.63	-.60	-.51	-.50	-.33	-.32	-.52	-.41	-.27	-.36	-.29	-.30	-.29	-.29	-.29
i11	-.41	-.42	-.43	-.54	-.37	-.58	-.51	1	.45	.39	.45	.49	.38	.45	.24	.26	.33	.24	.25	.26	.25	.24	.27	.26	.26
i12	-.56	-.53	-.50	-.68	-.52	-.72	-.74	.56	1	.57	.76	.73	.49	.58	.41	.35	.60	.51	.33	.38	.38	.34	.35	.35	.34
i13	-.46	-.38	-.39	-.56	-.38	-.60	-.60	.48	.67	1	.53	.55	.49	.43	.31	.26	.49	.38	.23	.28	.27	.25	.27	.26	.21
i14	-.60	-.55	-.55	-.69	-.56	-.76	-.73	.57	.84	.63	1	.75	.49	.58	.36	.33	.60	.45	.31	.34	.34	.31	.32	.34	.33
i15	-.54	-.51	-.51	-.65	-.52	-.72	-.70	.61	.81	.64	.83	1	.50	.60	.35	.33	.58	.46	.32	.34	.32	.31	.32	.36	.35
i16	-.42	-.43	-.42	-.49	-.41	-.58	-.63	.47	.59	.57	.59	.60	1	.43	.22	.21	.43	.31	.16	.21	.23	.22	.21	.24	.18
i17	-.46	-.42	-.47	-.58	-.41	-.61	-.60	.57	.68	.52	.68	.70	.52	1	.38	.31	.50	.41	.28	.31	.32	.37	.33	.41	.37
i18	-.23	-.23	-.12	-.45	-.34	-.42	-.41	.33	.50	.40	.44	.43	.29	.47	1	.47	.52	.62	.53	.51	.47	.57	.54	.52	.57
i19	-.30	-.30	-.26	-.49	-.38	-.45	-.51	.41	.51	.40	.48	.49	.34	.47	.62	1	.46	.41	.49	.56	.35	.32	.39	.36	.45
i20	-.44	-.38	-.35	-.55	-.42	-.59	-.62	.44	.69	.59	.69	.66	.52	.60	.61	.65	1	.61	.42	.50	.50	.38	.43	.41	.41
i21	-.28	-.25	-.18	-.43	-.36	-.47	-.50	.33	.60	.45	.53	.54	.38	.49	.72	.59	.70	1	.46	.49	.57	.44	.45	.41	.45
i22	-.31	-.27	-.16	-.49	-.40	-.47	-.41	.39	.48	.35	.44	.46	.25	.42	.70	.68	.59	.63	1	.56	.38	.36	.39	.34	.47
i23	-.37	-.35	-.23	-.50	-.46	-.49	-.48	.35	.50	.37	.45	.46	.29	.43	.63	.74	.63	.62	.72	1	.42	.35	.40	.39	.46
i24	-.22	-.21	-.15	-.33	-.21	-.34	-.36	.35	.46	.33	.42	.40	.30	.41	.57	.51	.59	.67	.53	.54	1	.37	.38	.38	.36
i25	-.25	-.22	-.24	-.37	-.29	-.36	-.37	.33	.42	.33	.39	.39	.29	.46	.67	.46	.46	.52	.50	.47	.46	1	.47	.51	.46
i26	-.23	-.18	-.20	-.39	-.24	-.37	-.36	.37	.43	.34	.41	.41	.27	.42	.64	.57	.52	.54	.55	.52	.47	.58	1	.52	.53
i27	-.26	-.23	-.29	-.33	-.26	-.35	-.35	.34	.41	.31	.40	.44	.30	.49	.61	.50	.49	.48	.47	.49	.45	.60	.62	1	.53
i28	-.34	-.26	-.28	-.47	-.39	-.46	-.46	.39	.51	.32	.48	.51	.28	.54	.75	.62	.58	.62	.64	.63	.51	.63	.71	.71	1
N_i	2156	2142	2155	2152	2131	2136	2149	2153	2160	2159	2154	2146	2142	2145	2159	2146	2156	2152	2143	2146	2150	2161	2155	2149	2156
M	1.92	1.84	2.32	2.48	1.96	2.33	2.24	1.55	2.40	2.20	2.25	2.15	2.02	1.85	1.84	1.25	2.08	1.86	1.32	1.47	1.86	1.72	1.69	1.83	1.30
SD	.84	.85	.83	1.02	1.03	1.04	1.02	.76	.97	.76	1.00	.92	.69	.79	.92	.49	.84	.71	.58	.66	.81	.76	.73	.80	.62
SK	.68	.85	.30	.12	.77	.32	.41	1.4	.19	.42	.37	.48	.54	.70	.90	2.0	.53	.05	1.8	1.2	.75	.87	.87	.75	2.2
K	-.11	.14	-.42	-1.1	-.62	-1.0	-.94	1.5	-.93	.02	-.91	-.55	.62	.02	-.09	4.4	-.18	-.00	3.3	1.2	.08	.38	.38	.05	4.9

Pearson's correlations are placed to the right of the diagonal, and polychoric correlations to the left of the diagonal.

Abbreviations: N_i = item sample size, M = Mean, SD = Standard Deviation, SK = Skewness, K = Kurtosis

All Pearson's Correlations were significant at $p < .01$

Phase V: Reliability and Validity Analyses

Table 11

Distribution of items per scale after content validity analysis

Item	Inter-rater agreement %/ Factor	Factors / Item/ Cronbach's Alpha Coefficient	
			<i>Problem Solving</i>
i1	100/PS	Analyzes the nature of a problem before starting to solve it	
i2	90/PS	Does extra credit	
	10/EC		
i3	90/PS	Makes decisions easily	
	10/EC		
i4	70/PS	Works hard, even in courses he or she does not like	
	20/AC		
	10/BC		
i5	100/PS	Gives good suggestions for solving problems	
		<i>Note: this item was eliminated due to construct ambiguity</i>	
i6	60/PS	Makes suggestions without offending others	
	30/BC	<i>Note: This item was eliminated due to lack of inter-rater agreement</i>	
	10/EC		
i7	80/PS	Asks to make up missed assignments	
	10/AC		
	10/BC		
i8	50/PS	Works well under pressure	
	40/EC	<i>Note: This item was eliminated due to lack of inter-rater agreement</i>	
	10/AC		
i9	80/PS	Has good study habits	
	10/AC		
	10/BC		
i10	90/PS	Is well organized	
	10/AC		
			<i>Attentional Control</i>
i11	70/AC	Does not complete tests	
	20/PS		
	10/BC		
i12	100/AC	Is easily distracted from classroom	
i13	80/AC	Makes careless mistakes	
	10/PS		
	10/BC		
i14	100/AC	Has a short attention span	
i15	100/AC	Has trouble concentrating	
i16	80/AC	Forgets things	
	10/PS		
	10/EC		
i17	70/AC	Has trouble shifting gears from one task to another	
	20/PS		
	10/BC		
			<i>Behavioral Control</i>
i18	20/BC	Argues when denied own way	
	80/EC	<i>Note: item moved to the Emotional Control scale</i>	
i19	80/BC	Breaks other children's things	

	20/EC		
i20	100/BC	Acts without thinking	
i21	100/BC	Interrupts others when they are speaking	
i22	70/BC	Uses foul language	
	30/EC		
i23	80/BC	Hits other children	
	20/EC		
i24	100/BC	Cannot wait to take turn	
			<i>Emotional Control</i>
i25	100/EC	Stays disappointed a long time if a favorite activity is cancelled	
i26	100/EC	Changes moods quickly	
i27	100/EC	Is easily upset	
i28	70/EC	Throws tantrums	
	30/BC		

Abbreviations: PS = Problem Solving, AC = Attentional Control, BC = Behavioral Control, and EC = Emotional Control

Internal consistency reliability results

SPSS 14.0 was utilized for the reliability analysis. As observed in Table 12, internal consistency coefficients ranged from 0.805 to 0.890 for the constructs, which are considered indicators of moderate to high reliability. The “Cronbach’s Alpha if item deleted” feature was used in order to identify items that did not significantly contribute to the scale. None of the items demonstrated an increase in Alpha coefficient if deleted. A lower coefficient for the full screener (.640) was obtained as expected, and should be interpreted as a confirmatory indicator of the multidimensionality of this executive functions measure.

Table 12

Internal consistency reliability (Cronbach’s alpha coefficients by scale)

Scale	Cronbach’s Alpha Coefficient
Problem Solving	0.805
Attentional Control	0.890
Behavioral Control	0.842
Emotional Control	0.845
Full Screener	0.640

Construct validity:

Confirmatory Factor Analysis (CFA)

In order to answer the question about the appropriateness of considering the screener a unidimensional versus multidimensional instrument, a series of CFA models were analyzed in Mplus 4.0 (L. K. Muthén & Muthén, 1998-2006).

The first model (Model 1) corresponds to a one-factor/unidimensional model with 25 indicators. The latent construct was called “Executive Function” (EF), as it presumably measures executive functions as a whole (see Figure 7). The model rejection rate in Mplus was the highest (100%) and the model did not converge. Models with 2, 3, and 4 factors converged normally. Model 2 had two latent constructs, one called “Problem Solving” (PS) and a second construct referred to here as “Behavioral Self-Regulation”, (BSR). This second factor was comprised of items that originally corresponded to the factors Attentional Control, Behavioral Control, and Emotional Control (see Figure 8). A third model (Model 3) was comprised of three latent constructs: “Problem Solving” (PS), “Attentional Control” (AC), and “Behavioral/Emotional Control” (BEC) (see Figure 9). These last three models demonstrated increasingly better fit in terms of a lower (or less inflated) chi-square, more degrees of freedom, and higher fit indexes. Chi-square differences were observed but not tested because the models were not nested. CFI differences are included in Table 13. Although a small CFI difference between Model 3 (3 factors) and Model 4 (4 factors) was observed, RMSEA and TLI indexes are higher in Model 4. These results support the hypothesis that the group of items selected for this screener yields information about more than one component of executive functions; therefore, it is possible to reject the null hypothesis for unidimensionality. It also demonstrated that the theoretical model proposed (4 factors) is not only viable but testable. Figure 10 illustrates Model 4.

Table 13

Fit indexes for the unidimensional versus multidimensional model of executive functions

Model	WLSMV χ^2	df	CFI	Δ CFI	TLI	RMSEA
Model 1: 25 items, 1 Factor			No convergence			
Model 2: 25 items, 2 Factors	4856.138	84	0.827		0.944	0.162
Model 3: 25 items, 3 Factors	2245.639	106	0.923	-0.096	0.980	0.097
Model 4: 25 items, 4 Factors	1899.635	111	0.935	-0.012	0.984	0.086

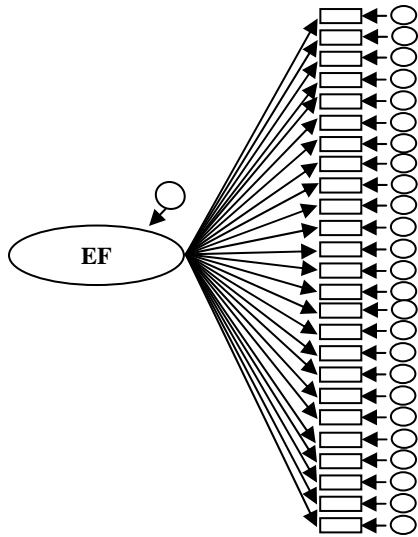


Figure 7, Unidimensional model 1

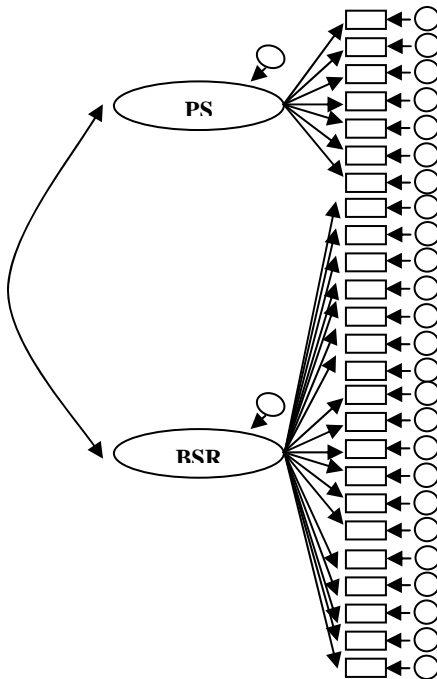


Figure 8, Multidimensional model 2

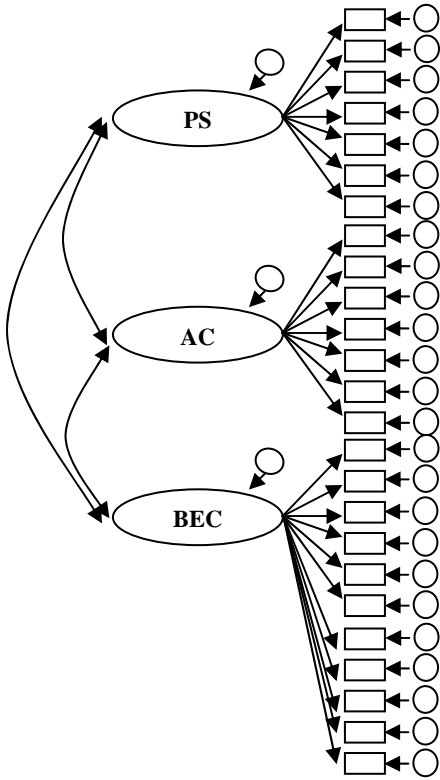


Figure 9, Multidimensional model 3

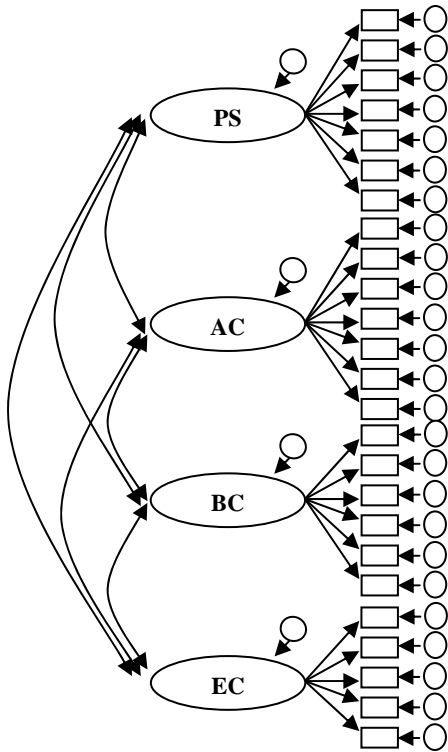


Figure 10, Multidimensional model 4

Once the multidimensionality of the model was confirmed, the next analysis consisted of examination of the model modification index recommended by the Mplus modeling software. This step commonly corresponds to an exploratory analysis (e.g., EFA) due to the fact that these modifications are data-driven, or *a posteriori*, rather than theoretically-driven. It has been observed that theory should prevail, as there is ongoing tendency to over-factorize latent constructs (Frazier & Youngstrom, 2007). A common practice is to negotiate between the two spheres, and make decisions on model changes based upon indexes but supported by theory, and guided by two principles: parsimony and replicability. Modification indexes for the 4 factor model, which will be called here *Baseline Model*, suggest allowing item #20 “Acts without thinking” to load in all factors. It was originally set to load only on the Behavioral Control factor. Table 14 summarizes the comparative fit indexes across models.

Table 14

Exploratory analysis following modification index recommendations

Model	WLSMV χ^2	Df	CFI	Δ CFI	TLI	RMSEA
Model 4 –Baseline	1899.635	111	0.935		0.984	0.086
Model 5: i20 loading on all	1469.593*	111	0.951	-0.016	0.988	0.075
Model 6: i20 loading on PS and BC	1550.552	112	0.948	0.003	0.987	0.077
Model 7: i20 eliminated	1403.811	105	0.952		0.988	0.076

As reported in Table 14, Model 4 (baseline) was compared against three more models. Model 5 included factor loadings from each factor to item#20; Model 6 included factor loadings from the Behavioral Control (BC) and Problem Solving (PS) factors. Finally, in Model 7, the factor loading of i20 was constrained to zero, eliminating the relationship between the indicator (i20) and the latent constructs. Figure 11 illustrates these models.

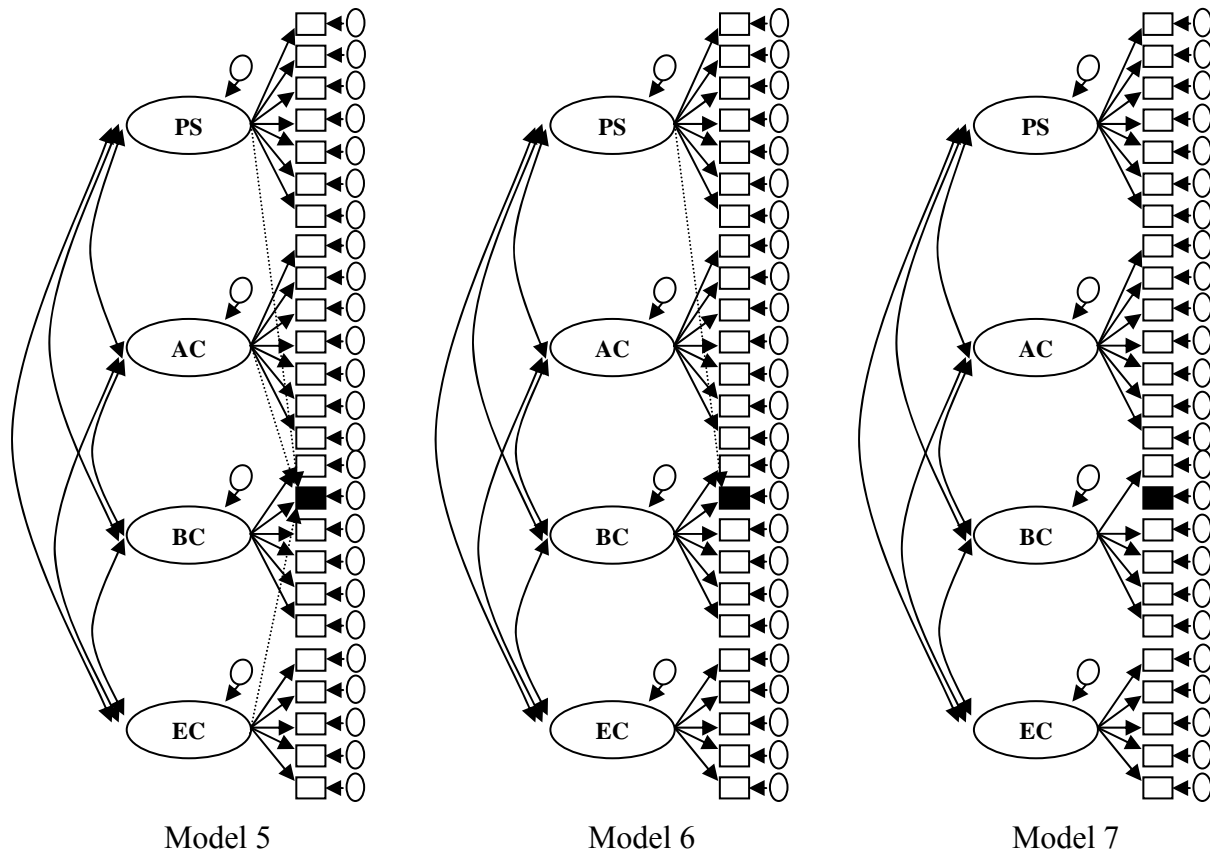


Figure 11, Test for model modifications

Parameters Estimates

By default, Mplus 4.0 constrains the loadings on the first item for each factor to 1.0. For this reason it is useful to report the standardized factor loadings as well. Table 15 summarizes the factor loadings for each item. Item #20 has a negative estimate due to two possible reasons; (1) it shares variance with more than one construct and needs to compensate for this; or (2) it is negatively worded whereas all the items on Problem Solving are positively worded. Factor loadings represent the strength of the relationship between the construct and the item in terms of shared variance. Standardized values closer to 1.0 are optimal. In case of cross loadings (e.g.,

i20), a lower estimate was expected. These estimates ranged from .618 to .936, which are significant and explain the overall goodness of fit of the model.

The following figure represents the final model configuration and includes the correlations among factors (see Figure 12). As can be observed in the figure, the correlations range from $r = -.502$ to $r = -.875$, which are considered moderate to high. Significant, moderate to high correlations were expected between this screener's four factors because there are strong relationships between the underlying latent construct they represent. It is possible that, as *executive functions*, these constructs are measuring components of the same underlying (and unifying) latent construct, which could be a second order construct in this model. Negative correlations were observed here and in the inter-item matrix. This negativity is explained by the wording orientation of the items. Problem Solving items included positively worded items (with the exception of item # 20 which is negatively worded), and the other three factors used only negatively worded items. However, raw scores need not be reversed for modeling purposes.

Table 15

Factor loadings

Variable & Factor	Unstandardized Factor Loading	Standardized Factor Loading	Standard Error	R ²
Problem Solving				
i1	1.000	.740	.000	.547
i2	.977	.723	.023	.523
i3	.901	.666	.022	.444
i4	1.164	.861	.022	.742
i7	.997	.738	.024	.544
i9	1.265	.936	.023	.877
i10	1.207	.893	.023	.797
i20	-.445	-.329	.027	.717
Attentional Control				
i11	1.000	.671	.000	.450
i12	1.357	.911	.031	.829
i13	1.083	.726	.028	.528
i14	1.364	.915	.033	.838
i15	1.327	.890	.031	.793

i16	1.015	.681	.029	.464
i17	1.169	.784	.029	.615
Behavioral Control				
i19	1.000	.823	.000	.677
i20	.751	.618	.025	.717
i21	1.042	.857	.022	.734
i22	.983	.809	.023	.654
i23	.997	.820	.022	.673
i24	.860	.707	.024	.500
Emotional Control				
i18	1.000	.884	.000	.782
i25	.842	.744	.018	.554
i26	.871	.770	.018	.593
i27	.853	.755	.018	.569
i28	1.011	.893	.017	.798

Analyses of Measurement Invariance: Multiple Groups Approach

Also recognized as multisample CFA (MCFA), measurement equivalence, and factorial invariance, the multiple groups approach allows for testing the structure of the screener across different groups (French & Finch, 2006). Measurement invariance is a stepwise procedure, which goes from testing the weakest form of invariance (Configural) to the strongest (Scalar). Two analyses were performed in this study. One included the factorial invariance examination across sex (female, male) for gender invariance; the second included the test of invariance across two age-groups (young = 6 - 8 years old; older = 9-11 years old) for developmental invariance.

Measurement Invariance across Gender. Prior to the invariance examination, the aggregated sample was divided into two groups, female and male, and the full CFA model (baseline model) was run independently for each sub-sample. The results of these comparisons served as earlier indicators of relative invariance. The model fit for the female group obtained a slightly better fit than the male group. Table 16 summarizes the model fit for each sub-sample. Table 17 contains all the factor loadings for the aggregated sample, and the independent female

and male sub-samples. The factor loadings in each group are quite close with respect to each other and to the aggregated sample.

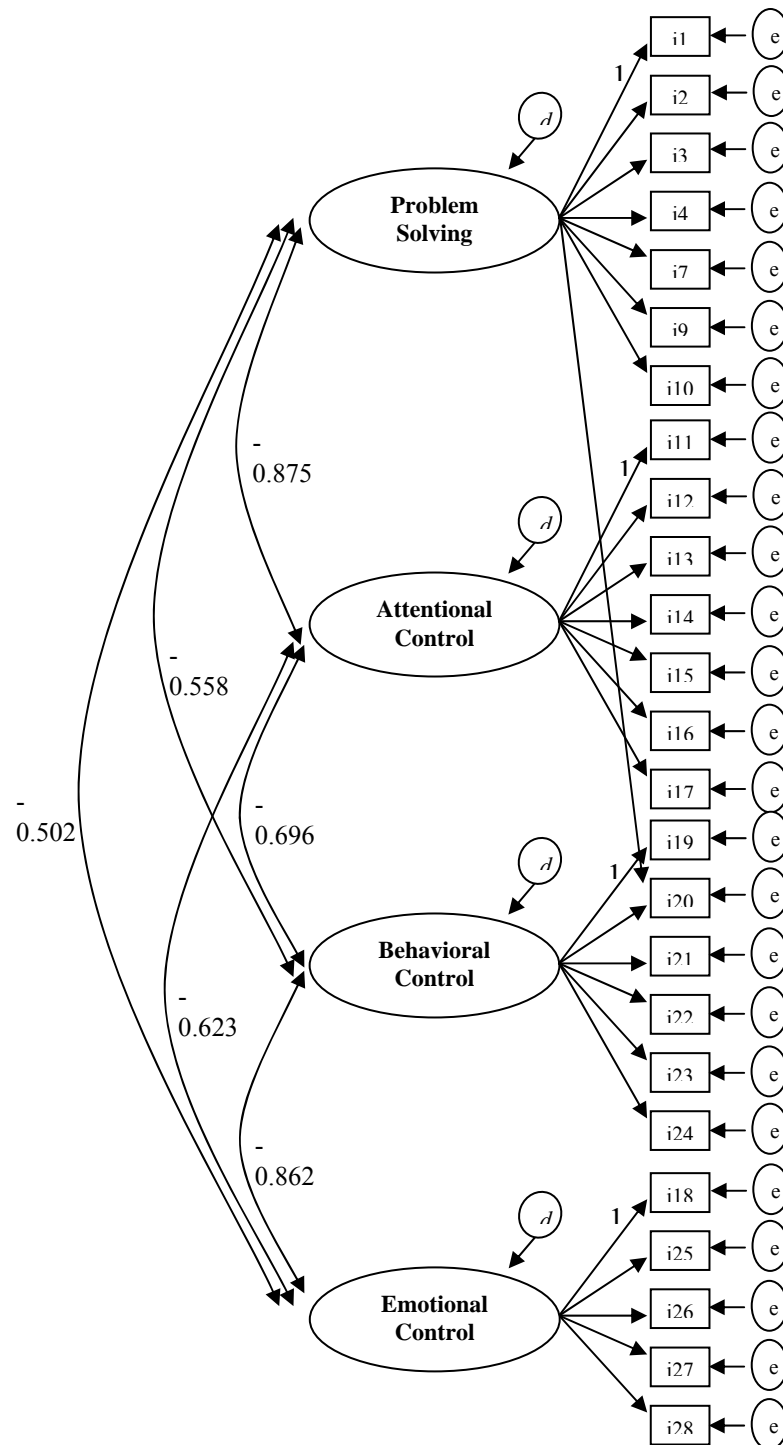


Figure 12, Final CFA model for the BASC Executive Functions Screener

Table 16

Model fit for each sub-sample (by sex)

Model	WLSMV χ^2	Df	CFI	TLI	RMSEA
Sex					
Female, $n = 860$	517.175	91	0.963	0.989	0.074
Male, $n = 1305$	990.686	105	0.937	0.983	0.080

Table 17

Standardized factor loadings for aggregated, only female, and only male sub-samples

Factors and Items	Aggregated	Female	Male
Problem Solving			
i1	.740	.773	.724
i2	.723	.737	.697
i3	.666	.738	.632
i4	.861	.855	.855
i7	.738	.732	.706
i9	.936	.952	.914
i10	.893	.931	.843
i20	-.329	-.346	-.318
Attentional Control			
i11	.671	.678	.664
i12	.911	.894	.911
i13	.726	.760	.683
i14	.915	.930	.895
i15	.890	.895	.878
i16	.681	.747	.633
i17	.784	.765	.780
Behavioral Control			
i19	.823	.820	.806
i20	.618	.601	.620
i21	.857	.848	.837
i22	.809	.801	.790
i23	.820	.834	.788
i24	.707	.667	.698
Emotional Control			
i18	.884	.845	.876
i25	.744	.764	.721
i26	.770	.748	.798
i27	.755	.754	.756
i28	.893	.919	.879

Once the initial CFA model has been tested in each subsample, a multiple group approach is appropriate. As explained earlier, examination of invariance includes three steps,

which are performed on a hierarchical stepwise method. The first model (step 1) included the analysis of configural invariance. In this first test, no equality constraints are imposed across models, producing a baseline model for the next test of invariance (Cordon & Finney, 2006). If achieved, configural invariance indicates that the same four components underlying the structure of the screener are being measured in the different groups. If configural invariance is not met, it could mean that one or more of the four components are conceptualized differently in each group; it can also mean that an extraneous uncontrolled variable is affecting the groups differently (Cheung & Rensvold, 2002; Rusticus & Hubley, 2006). As indicated in Table 4.9, the multidimensional four-factor model (baseline) has reasonably good goodness of fit indexes (CFI, TLI, and RMSEA). The chi-square is observed and included in the analysis, but it is not considered a criterion due to its sensitivity to sample size (e.g., easily inflated) in CFA and SEM modeling (Cheung & Rensvold, 2002; Kim et al., 2006; Rusticus & Hubley, 2006).

The next step is the examination of metric invariance. This level of assessment determines if the factor loadings of the four latent constructs are equal across groups. This test is performed by adding one more constraint to the configural model, which consists of setting all unstandardized factor coefficients equal across both groups. In other words, the equality of the strength of the relationship of the items and their factors across groups is tested (Cordon & Finney, 2006). If metric invariance is met, it is reasonable to assume that the item scores are scaled to the latent factor scores using the same unit of measurement across groups (Rusticus & Hubley, 2006). When metric invariance is not met, it could be indicative of poor operationalization of the latent variables (e.g., the four executive functions), or possibly that there are intrinsic differences across groups in how these four constructs are conceptualized.

Metric invariance examination in this model was demonstrated and supported by the goodness of fit indexes. Chi-square difference test was performed utilizing the DIFFTEST option of the Mplus 4.0 software, which utilizes a formula that corrects for the already adjusted chi-square produced by the WLSMV estimator. This test can be performed because the metric invariance model is nested within the configural model (Cordon & Finney, 2006). The CFI difference (ΔCFI) between the configural invariance model and the metric invariance model is $\Delta\text{CFI} = -.005$, which is within the cut-off score limits recommended by Cheung and Rensvold (2002) of $\Delta\text{CFI} \leq -.01$.

The next step in the analysis is the scalar invariance examination. This strongest level of examination imposes equality constraints on the intercepts. The intercepts allow prediction of values of the observed variables (i.e., the items) when the latent constructs (i.e., factors) are equal to zero (Cordon & Finney, 2006). In other words, if scalar invariance is met in this study, it would be expected that if two individuals have the same values on the latent constructs (i.e., the components of the screener), they would obtain the same values on the items regardless of their sex. When scalar invariance is not met, the measure should not be used for cross-group comparisons due to the existence of bias. This bias could be caused by different levels of response style (e.g., extreme or acquiescence) or by intrinsic features of the item that makes it more salient for one group than another (Rusticus & Hubley, 2006).

Scalar invariance was supported in the current analysis. As can be observed in Table 18, the CFI difference between the metric invariance model and the scalar invariance model is $\Delta\text{CFI} = -.011$, which is reasonably within the cutoff score of $-.01$ (Cheung & Rensvold, 2002). Other fit indexes such as TLI and RMSEA also demonstrated adequate goodness of fit.

Table 18

Measurement invariance for sex groups

Model	WLSMV χ^2	$\Delta \chi^2$	Df	CFI	Δ CFI	TLI	RMSEA
Step 1: Configural Invariance	1588.278		218	0.947		0.986	0.076
Step 2: Metric Invariance	1445.774	29.506 $p = .0301$	215	0.952	-0.005	0.988	0.073
Step 3: Scalar Invariance	1119.477	82.052 $p = .0000$	178	0.963	-0.011	0.988	0.070

Measurement invariance across developmental ages. Similar to the previous analysis, the aggregated full sample was divided into two groups, “young” comprised of children aged 6 to 8 years, and “older,” comprised of children aged 9 to 11 years. Measurement invariance was not expected across gender groups as much as it would be expected across ages (Becker et al., 1987; Davidson et al., 2006; Passler et al., 1985) making relevant a second measurement invariance examination. Table 19 summarizes the model fit for each sub-sample independently.

Table 19

Model fit for each sub-sample (by age)

Model	WLSMV χ^2	Df	CFI	TLI	RMSEA
<i>Age Group</i>					
Young (6-8), $n = 957$	815.774	96	0.940	0.981	0.089
Older (9-11), $n = 1208$	808.273	101	0.956	0.989	0.076

Table 20 presents the factor loadings for the aggregated sample and each age group independently. Notice that most of the items have similar factor loadings, with the exception of item #11, in which there is a larger difference. This item, “*Does not complete tests,*” has a stronger relationship to its factor (Attentional Control) in the “older” group than in the “young” group.

Table 20

Standardized factor loadings for aggregated (6-11), young (6-8), and older (9-11) samples

Factors and Items	Aggregated	Young (6-8)	Older (9-11)
Problem Solving			
i1	.740	.749	.733
i2	.723	.744	.704
i3	.666	.719	.634
i4	.861	.843	.873
i7	.738	.752	.727
i9	.936	.914	.953
i10	.893	.901	.887
i20	-.329	-.333	-.338
Attentional Control			
i11	.671	.591	.742
i12	.911	.895	.923
i13	.726	.712	.746
i14	.915	.909	.920
i15	.890	.877	.901
i16	.681	.630	.719
i17	.784	.774	.796
Behavioral Control			
i19	.823	.823	.828
i20	.618	.628	.603
i21	.857	.855	.862
i22	.809	.795	.829
i23	.820	.812	.832
i24	.707	.642	.761
Emotional Control			
i18	.884	.881	.898
i25	.744	.761	.732
i26	.770	.750	.784
i27	.755	.753	.757
i28	.893	.878	.900

As was the case for the analysis of invariance across gender, examination of configural, metric, and scalar invariance was performed in this second analysis. Configural invariance for the two age groups was met in this study. As reported in Table 21 goodness of fit indexes are within the cut-off scores. Specifically, the CFI (.946) is between the lower bound and the optimal bound recommended by Hu & Bentler (1999), the TLI is closer to the optimal level, and the RMSEA is at the cut-off point. These results demonstrate that the screener's same executive function components are being measured in the two age groups.

Metric invariance was also met in this study. The WLSMV adjusted chi-square is lower than in the adjacent model, yet and as expected in CFA models with large sample sizes, the $\Delta\chi^2$ test is significant ($p = .0018$), the CFI (the criterion) is .950, and the $\Delta\text{CFI} = -.004$, both values within the cutoff scores. TLI and RMSEA values improved as well. These results suggest that the two age groups are not only conceptualizing the four components of executive function similarly but also that the same unit of measurement is being used across the groups. Equality of factor loadings also demonstrates that the strength between the items and the latent constructs is similar across these two age groups.

The final step was examining the scalar invariance, which was also met by the screener's model. Goodness of fit indexes demonstrated that the model fit was adequate, and that it is possible to utilize the screener to perform mean comparisons across ages. Results are summarized in Table 21.

Table 21

Measurement invariance for two age groups

Model	WLSMV χ^2	$\Delta\chi^2$	<i>Df</i>	CFI	ΔCFI	TLI	RMSEA
Step 1: Configural Invariance	1742.818		218	0.946		0.987	0.080
Step 2: Metric Invariance	1635.108	40.425 $p=.0018$	220	0.950	-0.004	0.988	0.077
Step 3: Scalar Invariance	1255.042	93.759 $p=.0000$	181	0.962	-0.012	0.989	0.074

CHAPTER 5

DISCUSSION

Self-regulation refers to the many processes by which the human psyche exercises control over its functions, states, and inner processes. It is an important key to how the self is put together. Most broadly, it is essential for transforming the inner animal nature into a civilized human being. (Vogs & Baumeister, 2004, p. 1; as cited by Rueda, Posner, & Rothbart, 2005, p.1)

Data and Item Screening

The aim of the present study was the development of a screener for the assessment of executive functions using a set of selected items from the BASC – Teacher Rating Scale for children aged 6-11 (Reynolds, and Kamphaus, 1992). A baseline theoretically-driven model with four factors was created to scaffold the instrument creation. These four factors represented four important executive functions: a first factor that was called “*Problem Solving*,” which included behaviors related to planning and initiation of goal/task attack, and three more factors that are related to the self-regulatory mechanisms involved in the execution of the programs needed for the goal attainment, “*Attentional Control*,” “*Behavioral Control*,” and “*Emotional Control*.” These four components were treated as latent constructs in the CFA model, and their observed variables were 28 items from the BASC-TRS C that were believed to serve as indicators of the executive behaviors.

Prior to the examination of the model, data and item screenings were performed in several fashions. First, analysis of normality indicated that the data could be reasonably considered as normally distributed, as the levels of skewness and kurtosis remained within the criterion limits; however, it was decided that this data should be treated as categorical (ordinal

type) as it has been recommended in the most recent SEM literature (e.g., L. K. Muthén & Muthén, 1998-2006; Flora & Curran, 2004), due to the few Likert answer-points included in the original rating scale includes (never-1, sometimes-2, often-3, almost always-4). Second, few subjects were eliminated because 80% or more data was missing, or because they were identified as outliers by the DeCarlo macro on SPSS 14.0. Third, FIML, a sophisticated missing data treatment was chosen over the typical listwise deletion procedure to optimize data usage, due to the results of previous research demonstrating the stability and efficiency of this technique over other methods (Enders & Bandalos, 2001). Fourth, content validity analysis was performed via inter-rater agreement percentages. A survey was distributed among researchers in the fields of behavioral measurement and neuropsychology. Results demonstrated agreement on the majority of the items, yet two items were eliminated due to the lack of consensus (criterion $\geq 70\%$). And fifth, an additional item was eliminated because of unclear or ambiguous wording after informally asking elementary school teachers how they would interpret its content.

Latent Construct Analysis and Operationalization

Results in the content validity analysis indicated few problems on the operationalization of some of the constructs that were to be addressed. For instance, the “Problem Solving” factor received an average of 81% (out of 10 items) of agreement rate. Although this could mean that some of the items loading in this factor are measuring construct irrelevant content, factor loadings for this construct were significant and demonstrated that a large amount of the variance on the latent construct was explained by its indicators. However, three out of the original 10 items on this factor were eliminated to more accurately represent the target latent construct. These adjustments did not increase internal consistency coefficients (which are data-driven), but the modifications increased construct validity as it was observed on model goodness of fit.

It could be also argued that this first factor is attempting to measure too many components of the executive functions. Due to its intrinsic condition of screener imbedded on a larger behavior assessment rating scale (i.e., BASC), it was important to target a diversity of behaviors believed to be executive in nature. It is believe here that during behavioral regulation mediated by executive systems, some of the most important parts of the process are planning the goal execution, organizing information (input, output), making decisions about the goal/task approach, and programming the action. This is a complex step in the process, and it occurs in milliseconds prior to and during the motor execution. This function is highly dependent upon adequate prefrontal cortex functioning, and specifically, of the coordinated work of tertiary areas (Luria, 1973). In this context, the factor “Problem Solving” could also be recognized as a preparation for action-step towards goal execution.

From a neuroanatomical perspective, Fuster (1997) proposed that, due to its location within the frontal lobe, the prefrontal cortex has a very specific role which is the temporal organization of behavior. In this sense, there are two important functions involving prefrontal cortex, the initiation and the execution of action. The factor, Problem Solving, was included to attempt to measure this first component of Fuster’s model, the initiation of behavior towards goal attainment. Initiation of behavior relates to a very complex process that includes intention (motivation), drive, attention, effort, organization, decision making, planning, and programming of action. These elements are captured by some of the items loading in the Problem Solving factor. For instance, item # 2, “*Does extra credit*” and item # 4 “*Works hard, even in courses s/he does not like*” could be taken as motivation-related items, but they can also serve as behavioral indicators of drive and intention. It is necessary to note that a complex behavior, such as working hard or completing extra work implies functions beyond drive and intention, but it is

a limited yet efficient way to capture the concept. One could say that during problem solving, the primary goal is to achieve an action (the “problem”) and the prefrontal cortex, in coordination with other several systems, would be responsible for designing the “how to...” (the “solving”).

Furthermore, and as we know, the BRIEF is the only available scale on the market to behaviorally assess executive functions in children. Gioia et al (2000) reported in the BRIEF manual a similar situation regarding the Plan/Organize scale. They asserted that an empirical analysis of the BRIEF structure demonstrated that these two areas should be integrated into one factor, even though theoretically they could be separated into independent executive functions (p.19-20). Statistically, when high correlations among items are observed, the integration is recommended to avoid collinearity, which violates the assumption of interdependency among the indicators.

Another caution was raised by the content validity analysis and was related to difficulty in differentiating the placement of items between the Behavioral Control and the Emotional Control factors. Specifically, most of the experts agreed on the 4 items for the Emotional Control factor (average agreement rate = 92%), yet some disagreement was observed on the classification of the items under the Behavioral Control factor (average agreement rate = 78%). As a result, item #18, “*argues when denied own way,*” was moved from the original position (Behavioral Control) to the Emotional Control factor. This item was originally perceived as an indicator of inhibition and behavioral self-regulation; however, 80% agreement from the panel of experts demonstrated that it is a better operationalization of emotional self-regulation.

Moreover, it could be argued that it is difficult to differentiate these two factors from each other, and this is demonstrated by the small CFI difference ($\Delta\text{CFI} = -0.012$) observed between the three factor model (Model 3, see Figure 9) and the four factor model (Model 4, see

Figure 10). This division was imposed in the CFA model given three main reasons, all of them related to the three major debates regarding executive functions discussed earlier in Chapter 2: the “*conceptualization problem*”, the “*neuroanatomical problem*”, and the “*measurement problem*.”

The first is related to the conceptualization of two important constructs, behavioral inhibition and behavioral self-control, and their *theoretical* relationship with the construct of emotional self-regulation. As Barkley (1997) asserted, behavioral inhibition refers to the ability to inhibit initial prepotent responses to external cues or event, stop an ongoing response producing a delay period, and protecting this delay from interferences (e.g., competing events). Self-control refers to the responses by the individual that “serve to alter the probability of their subsequent response to an event,” (Barkley, 1997; p. 51) influencing also the probabilities of occurrence of the consequences naturally to follow. Emotional self-regulation or control, is highly related to behavioral inhibition and self control, yet it is differentiated by the fact that the prepotent response that is being delayed is the expression of emotional reactions “that would have been elicited by the event and whose expression would have been a part of the expression of those prepotent responses” (Barkley, 1997; p.182).

Second, as it was presented in Chapter 2, there is supporting research demonstrating four differentiated brain pathways or circuits from the prefrontal cortex and other anterior, posterior, and subcortical brain structures, in relation to initiation of action, and self-regulation of behavior and emotion. The dorsolateral prefrontal circuit has been involved in planning, self-monitoring, and other higher cognitive functions (related to the latent construct “Problem Solving” in this screener study). This circuit includes connections between the dorsolateral caudate nucleus, the dorsolateral portion of the globus pallidus and the rostral substantia nigra reticulate, and the

parvocellular region of the medial dorsal and ventral anterior thalamic nuclei (Royall et al, 2002). The lateral orbitofrontal may be involved in the initiation of behaviors and the inhibition of prepotent behavioral responses, and it includes projections from cortical areas to the ventromedial caudate nucleus, the dorsomedial area of the internal globus pallidus, and the rostromedial portion of the substantia nigra; from there to the magnocellular region of the medial dorsal and ventral anterior thalamic nuclei, and back again to the lateral orbitofrontal areas (Royall et al., 2002). Barkley's (1997) studies supported this neuroanatomical correlation between right prefrontal region and ventromedial regions of the striatum in interference control and response inhibition. Rolls (2002) asserted that this regulatory function may be mediated by the involvement of orbitofrontal cortex in the decoding and representation of primary reinforcers, which in turn regulate reward-related and punishment-related behavior. Another circuit that appears to be involved in behavioral self-regulation and monitoring is the anterior cingulate circuit. This circuit includes pathways from the ventral striatum to ventral pallidum, rostromedial substantia nigra, and to the dorsomedial thalamic nuclei (Royall et al, 2002). Finally, there is one more important path related to emotional control. Fuster (1997) affirmed that the orbitomedial frontal cortex has a close connection to limbic structures (i.e., amygdala), through the "paralimbic" cortex. In this way, although there are common structures and networks involved in the self-regulation of behavior and emotion, it could be reasonable to separate them for measurement purposes.

Third, and back to the assumption that emotional control can be independent of behavioral control, the debate about what is the best way to estimate executive functions, here called "*the measurement problem*," has shed some light on alternative approaches that add ecological validity to the assessment of executive functions. In this regard, the success of the

BRIEF serves as an example of new tendencies, and a model to follow. Gioia et al's (2000)

BRIEF includes two separate scales, "Inhibit" and "Emotional Control", which are integrated on a composite index score denominated "Behavioral Regulation Index." It could be argued that the two factors "Emotional Control" and "Behavioral Control" should be also integrated into one factor for the screener purposes. From a measurement perspective, there are some arguments in favor of and against this idea.

On one hand, and in favor of the separation, there is little available research in the BRIEF's Emotional Control and Inhibition scales, and in relation to diagnosis of disorders characterized by executive dysfunction (e.g., Jarrat, Riccio, & Siekierski, 2005), and most of the research has been looking at the BRIEF index scores (e.g., Mahone et al, 2002). Therefore, little is known about the performance of these two separate constructs in relation to typical/average behavior and psychopathology. One study by Gioia et al (2002) tested 4 different CFA models using Maximum Likelihood estimation, in a clinical sample of children with mixed diagnoses, and assessed with the BRIEF. Authors reported that for this clinical sample, a 3-factor model obtained the best fit indexes. Interestingly, the BRIEF scale Emotional Control loaded in an independent factor from Inhibit and Self-monitoring. The three factors were called "Metacognition," "Emotional Regulation," and "Behavioral Regulation." Emotional regulation included "Emotional Control" and "Shift" (Gioia et al, 2002). "Shift" refers to cognitive flexibility and the ability to make transitions (Gioia et al., 2000). Similar to these results, and in contrast to the BRIEF original model, items included under the "Emotional Control" factor in the screener presented in this study are indicators of behaviors related not only to emotional self-regulation but also to emotional and cognitive flexibility (e.g., *"stays disappointed a long time if a favorite activity is cancelled"*).

On the other hand, and against the separation, it could be argued that the high and significant correlation between the two factors ($r = -.862$), the small CFI difference between the 3-factor model and the 4-factor model ($\Delta\text{CFI} = 0.012$), in addition to the lower agreement rates by the panel make possible to consider further analysis of the current factorization. In this regard, Frazier & Youngstrom (2006) recalled the importance of parsimony when developing tests, in terms of the number of factors they attempt to address. They asserted that “recent commercial tests of cognitive ability are not adequately measuring the number of factors they are purported to measure by test developers” (p. 180). They also reminded test developers of the recent increase of overfactorization. It would be advisable to further examine the 4-factor versus a 3-factor model, and across different samples.

To this point, the pertinence of allowing three factors on the model (i.e., Problem Solving, Behavioral Control, and Emotional Control) has been revised, in terms of the red alerts identified from the content validity analysis. Yet there is one more factor that should be addressed, the “Attentional Control.” Starting at an item level, inter-rater agreement rates for the items in this factor were higher, and a reasonable consensus was achieved (average rate = 85%). It was noticed during the item screening analysis that some of the items within this factor had the highest correlations across the matrix. In other words, it could be argued that some of the items are redundant in terms of behavioral indicators on a screening instrument. A look into the statistical analysis for this construct may serve as a counterargument for this assumption. Factor loadings were generally significant, ranging from .671 to .915, and stable across gender and developmental ages. The R^2 values which indicate the contributions of each item into the shared latent factor variance (in other words, the amount of factor variance explained by each item),

ranged from .450 to .838, which are within significant ranges. Moreover, Cronbach's Alpha coefficients did not significantly increase if any of the seven items were to be deleted.

Furthermore, and at the construct level, it could be argued that *attentional control* should be considered a second-order or higher level self-regulatory ability, responsible for the modulation of behavior. This assumption comes from the extensive works of Michael I. Posner and his colleagues. In brief, and as was described in Chapter 2, Posner and colleagues asserted that voluntary, controlled attention is associated with dopaminergic-modulated networks such as the anterior cingulate, basal ganglia, and lateral prefrontal cortex, which are responsible for the "Executive Attention" or executive control system (Rueda, Posner, & Rothbart, 2005). This "executive" aspect of attention is not considered by these authors as an executive function, but rather as part of a larger and extensive network of cortical and subcortical areas. Posner called this system the attentional "organ." In its complexity, this organ appears to be a reasonable theoretical framework for the study of executive functions, if one assumes that either executive functions are regulated by this organ, or that executive functions are not more than attentional functions in nature (the "*conceptualization problem*").

The "Attentional Control" factor includes several items that are related to the functions of Posner's attentional mechanism, but as it is presented by Posner and his colleagues, it would be difficult to operationalize it as a latent construct using behavioral indicators on a screening instrument (e.g., ratings), whereas it has been successfully assessed as a "systemic organ" with laboratory and performance-based measures (the "*measurement problem*"). Therefore, the aim of this factor is to estimate the *executive* ability to regulate attentional functions, such as focusing, sustaining, and shifting. This theoretical assumption is supported by the early work of Norman and Shallice on the Supervisory Attentional System (e.g., Shallice, 1982), the work of

Baddeley on the “Central Executive” (Baddeley, 1986, 1996, 1998), Fuster’s research on prefrontal functioning (e.g., Fuster, 1997), and the Mirsky’s model of attention (e.g., Mirsky et al, 1991).

Model Modifications and Item-Analysis

At an item level, there is one more issue regarding the factorial structure of the screener that worth discussing here. Modification indexes produced by the modeling software recommended the inclusion of a factor loading from all the constructs to item # 20, “*Acts without thinking.*” This item was originally conceived as a Behavioral Control item, and the panel of experts agreed 100% about its contribution to the operationalization of such factor. The question here is, what is it about this item that may, or may not, make it contribute to all constructs? A look to this item reveals that it captures the essence of executive functions, in the sense that a behavior such as acting without thinking not only represents the lack of self-control but also the lack of planning and organization of the action prior to its execution, or in the best scenario, the inefficient preparation for the initiation of action. An excellent example of this type of behavior is observed in children with Attention Deficit Hyperactivity Disorder (Barkley, 1997).

Similar arguments could be made for the inclusion of factor loadings from the attentional control and emotional control constructs to the item #20. In a strict sense, attentional control is a factor included to examine behaviors related to the ability to focus, sustain, and shift attentional systems over relevant information. One could present the argument that attentional systems are regulated by or depend upon the intentionality and the drive to attain a goal (Fuster, 1997); or that acting without thinking may relate to an automatic-driven (posterior) more than voluntary-driven (anterior) focusing of attentional systems (Rueda, Posner, Rothbart, 2005) on information that is irrelevant or relevant depending upon the context (e.g., survival). However, in the context

in which this item is presented, acting without thinking suggests a poor, problematic, and inefficient behavior, related to overall lack of control.

Furthermore, and regarding the relationship between this item and the Emotional Control factor, it is important to note that Emotional Control is proposed as a measure of the ability to self-regulate emotional response to environmental and internal cues. From this definition, it can be asserted that “acting” without thinking could include the measurement of emotional responses that have not been mediated by any cognitive constraints, like in an “emotional response” without thinking, type of behavior (e.g., *Throw tantrums*, item # 28). However, in order to keep the instrument’s parsimony, and in addition to the support for differentiating between behavioral self-control and emotional control presented earlier, it was decided to differentiate between behaviors driven by other sources (e.g., aggression in item # 23, *“Hits other children”*), and emotions expressed through behavior (e.g., sadness, disappointment in item #27, “is easily upset”). This last source of behavioral expression is common to the items loading into the Emotional Control factor.

Unidimensionality versus Multidimensionality

One more issue worth discussing here refers to single-factor model versus multiple-factor model comparisons. Results clearly pointed out the need for expansion on the factorization, from a non-converging model (unidimensional) to a three-factor and four-factor model that not only converged but also obtained reasonably adequate goodness of fit indexes. It was discussed earlier the appropriateness of using a four versus a three-factor model. Yet, the question is what these results mean in terms of the “unity versus diversity” dilemma in the conceptualization of executive functions. And even more relevant for the purposes of this study, how do these results impact the current ways of assessing executive functions?

There are different angles to address these questions. From a statistical standpoint, it appeared appropriate to factorize the set of 25 items into three or four components. The decision of utilizing a four-factor model was made based on theory, both *a priori* –as the four constructs were originally defined before examining the fit of the model, and *a posteriori* –as some statistically-driven (or data-driven) modifications were applied to the model. From a theoretical standpoint, the construct validity analyses demonstrated the pertinence of including four components of executive functions in the screener. As it was stated earlier, it is difficult to confine the definition of executive function into one statement (Elliot, 2003), as much as it is difficult to avoid numerating its components while defining the construct (Miyake, Emerson, and Friedman, 2000) and to list all of the possible components under the executive functions umbrella (Baddeley, 1996). It was decided here to utilize the term “executive functions” as a manifest of an understanding of this latent construct in terms of its diversity, yet believing that, as it is the case for “g,” executive function could be approached as a second-order latent construct that unifies the diversity of its components. The invariant sustainability of the model across different groups may serve as evidence in favor of this statement. In other words, it seems appropriate to approach the assessment of executive function from its components, via indicators, and from there estimate the latent construct as a composite.

Having said that, the question that may come to mind would be *is the composite more useful than the analysis of its components?* This is a difficult question to address, and it has been largely discussed in the intelligence testing literature. It is known from analysis of overall composite scores versus part scores that the full composite is more stable (Canivez & Watkins, 2001; Neisser et al., 1996; Raguette, Campbell, Berry, Schmitt, & Smith, 1996), has the best predictive validity (Hunter & Hunter, 1984), and has a large amount of scientific and theoretical

support. In contrast, we also recognize that part scores (e.g., indexes, scales, factors) allow clinicians to perform profile analysis, in which patterns of subtest scores to assess individuals' strengths and weaknesses are interpreted. Having access to this part scores facilitates diagnosing and the selection of appropriate interventions (Glutting, McDermott, Konold, Snelbaker, & Watkins, 1998).

According to Kamphaus (2001), profile analyses can be of two ilk, normative and ipsative. Normative subtest profiles compare an individual's subtest scores with those of a norm-referenced group, whereas ipsative analyses look at intraindividual differences. Interestingly, and despite their popularity, profile analysis is not supported by empirical evidence, yet it would be recommended in clinical practice and when working with special populations (e.g., bilingual assessment) and for screening purposes.

For the purposes of this screener, it appears appropriate to create a composite score, which would be multidimensional in nature while it also underlines the "unity" of the construct. Access to the four scales (factors) would help clinicians to make further intervention decisions in areas such as planning and organizing (initiation of behavior) and in behavioral and emotional self-regulation. The Attentional Control scale would serve for identification of attention problems, commonly observed in ADHD and other pediatric psychopathologies.

Measurement Invariance and Construct Validity

Two independent multiple group CFA analyses were performed in this study, in order to identify measurement equivalence across groups. The first analysis included a stepwise examination of configural, metric, and scalar invariance across female and male sub-samples. The second analysis included the same three examinations across two age groups, children 6 to 8 years old (Young Group) and children 9 to 11 years old (Older Group). Configural invariance

was supported by the model in both cases, which demonstrates that the same four-factor structure underlying the screener is being measured across groups, in that “the same indicators are associated with the same factors for both groups” (Resvold & Cheung, 2001, p. 29). Metric invariance was also met. According to Rusticus and Hubley, (2006, p. 828) if metric invariance is met, “the measure may be used to examine structural relationships or correlations between the construct of interest and other constructs across groups.” Due to the overall fit of the model no further metric invariance testing of each factor or items was necessary. Finally, the strongest type of examination, scalar invariance, was also supported by the model. This means that the measure is free of bias across groups, and can be used for both genders and across developmental ages 6 to 11. These finding are of great importance for future research involving the BASC executive functions screener.

Although gender differences were not expected in this model, it could be argued that differences across developmental ages should be identified by the model, due to results obtained in experimental research in the development of executive functions (e.g., Passler, Isaac, & Hynd 1985; Becker, Isaac, & Hynd, 1987; Davidson, Amso, Anderson, & Diamond, 2006). These differences could be identified in two ways, shape and level. Configural, metric and scalar invariance examinations serve as tests for evaluating differences in shape (factorial and construct structure). As it was reported, measurement invariance was supported. Differences in level indicate whether one group “has more amount” of latent construct than the other; in this case, it would be necessary to perform a latent means difference test in order to identify level differences.

CHAPTER 6

CONCLUSIONS

Summary of Findings

Executive function is a complex construct, difficult to conceptualize, not easy to localize, and specially, challenging to operationalize. What was called earlier “*the measurement problem*” could be actually presented as a synthesis of the other two major debates surrounding this construct: diversity versus unity or “*conceptualization problem*,” and unit versus network or “*neuroanatomical problem*.” The main purpose of this study was to create a measurement instrument that is reliable and valid, and that can provide clinicians and researchers with a short yet efficient tool, for the ecological assessment of executive functions in elementary school-aged children.

In one of her most recent publications, Martha Denckla (2007) stated that if complex perceptual systems such as visual and auditory have been divided into components or sub-systems, and have been colloquially named the “what” and “where” functions, thus executive functions should be recognized as the “how” and “when” functions. Although it is not explicit, she may have realized the connection between this bold statement and well-studied theories of prefrontal functioning; for instance, Fuster’s theory. In brief, Fuster (1997) proposed that prefrontal cortex functions could be summarized into the initiation and execution of action. In this way, Fuster’s *initiation* would be Denckla’s *how* functions, in terms of how to perform an action. This function would involve planning, organizing the information, decision making, and

programming the execution. Moreover, Fuster's *execution of action* corresponds to Denckla's *when* functions, in terms of monitoring, controlling, stopping, delaying, and executing behaviors.

Furthermore, it is possible to assert that the term *executive function* represents a range of diverse, organize, and complex network of neuronal activities, involving dorsolateral and orbitofrontal prefrontal cortex, and their extensive connectivity with subcortical structures such as the anterior cingulate, thalamus, caudate, globus pallidus, substantia nigra, paralimbic association cortex and limbic system, among others. Executive input and output appears to flow across organized circuits. The dorsolateral circuit is implicated in planning, decision making, problem solving, set shifting, and working memory, among other functions; the lateral orbitofrontal circuit is involved in self-regulation of social behavior, and inhibition of prepotent responses; and the anterior cingulate circuit appears to be involved in attentional control, self-monitoring, and error correction. Finally, pathways from the orbitomedial prefrontal cortex and limbic systems are implicated in emotional self-regulation.

In its simplicity and parsimony, this conceptualization of executive functions served as an excellent framework for the operationalization of the screener for the behavioral assessment of executive functions in children that is presented here. Consequently, a CFA model comprised of 25 items loading into four latent construct or factors that represent four executive functions was developed, and its statistical properties were examined. These four functions were named "Problem Solving," "Attentional Control," "Behavioral Control," and "Emotional Control." Reliability analysis demonstrated moderate to high internal consistency in each factor. Content validity helped to eliminated construct irrelevant indicators. And finally, the analysis of construct validity served for the examination of the screener's supporting baseline model. Tests of the fit of nested and no-nested variations of the baseline model showed that the four-factor model was

stable and had the best fit indexes. Multiple Group CFA examination across gender and age also supported the hypothesis that a reliable and valid executive functions measure could be obtained from a behavioral rating scale, such as the BASC. If implemented in assessment practice, this screener will provide clinicians and researchers with a reliable and valid tool for the estimation of four executive functions in children aged 6 to 11 years. Further analysis is needed in order to recommend this instrument for diagnosis purposes, yet an early identification of executive dysfunctions can prompt further assessment and intervention implementations.

Although far from trying to answer to the “unity versus diversity” dilemma about the definition of executive functioning (the “*conceptualization problem*”), and even farther from answering Teuber’s (1971) dilemma about the “unity or diversity” of frontal lobe functions (the “*neuroanatomical problem*”), the results obtained in the screener’s CFA using a single-factor model versus multiple-factor models may serve as evidence of the need for examining executive functions separately into a range of constructs. In other words, these results provide us with new perspectives into the aforementioned “*measurement problem*” in the assessment of executive functions. Furthermore, and by proxy, these results may provide guidelines for further research, in terms of how we understand and conceptualize executive functions, and where to look at in terms of prefrontal networks associated with these components.

Clinical Implications, Limitations, and Future Research

As presented, this study adds to our current knowledge on the behavioral assessment of executive functions, and provides with a screening tool for the estimation of four executive functions in children. This line of research was initiated with the early work of Reynolds and Barringer (1995) on BASC items believed to assess frontal lobe functioning, and strongly introduced in the assessment market by Gioia and colleagues (2000) with the launching of the

Behavioral Rating Inventory Executive Function, BRIEF. However, until today several adult rating scales for the behavioral and ecological assessment of executive functions have been created, yet only the BRIEF has served as the unique instrument available for children in the last decade.

This study demonstrated the utility of deriving screening tools from a large pool of items, such as the pool provided by the BASC. Yet, there are some limitations that should be addressed before recommending this instrument for clinical use. First, analysis of convergent validity would be necessary to establish if this instrument is comparable and competitive with the BRIEF. Second, it is necessary to create T-scores for each one of the four constructs imbedded in the model, so norm comparisons are possible. Third, introduction of the composite score will be optimal for clinical applications and cross-instruments comparisons. Fourth, further analysis of the clinical validity of the instrument would enhance its utility as an assessment instrument. It would be recommendable to utilize a large clinical database to examine the instrument's sensitivity and sensibility and therefore, discriminant accuracy rate. Fifth, due to its nature, this screener was conceived within the limited frame of available items from the BASC. Therefore, there are important executive functions (e.g., working memory) that could not be included in the model, and are of great importance on the assessment of this construct. It is recommended that if this instrument were to become available for clinical use, it should be utilized as a complementary tool in the screening of executive functions. Finally, the best mechanism for the implementation of this instrument would be its introduction into the most recent BASC-2 scoring software making it accessible, cost-efficient, and viable for further external research.

Future research should include cross-cultural validation of the model, convergent and concurrent validity analyses in comparison to the BRIEF and to performance-based gold

standard measures of executive functions, and finally, correlation between behavioral indicators and morphological and functional neuroimaging of the prefrontal cortex in pediatric population.

REFERENCES

- Anderson, S. W., Damasio, H., Tranel, D., & Damasio, A. R. (2000). Long-Term Sequelae of Prefrontal Cortex Damage Acquired in Early Childhood. *Developmental Neuropsychology*, 18(3), 281-296.
- Anderson, V. (2001). Assessing Executive Functions in Children: Biological, Psychological, and Developmental Considerations. *Pediatric Rehabilitation*, 4(3), 119-136.
- Anderson, V., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between Cognitive and Behavioral Measures of Executive Function in Children with Brain Disease. *Child Neuropsychology*, 8(4), 231-240.
- Anzai, Y., & Simon, H. A. (1979). The Theory of Learning by Doing. *Psychology Review*, 86(2), 124-140.
- Ardila, A., Galeano, L. M., & Rosselli, M. (1998). Toward a Model of Neuropsychological Activity. *Neuropsychology Review*, 8(4), 171-190.
- Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between Intelligence Test Scores and Executive Function Measures. *Archives of Clinical Neuropsychology*, 15(1), 31-36.
- Aylward, G. P. (2002). Cognitive and Neuropsychological Outcomes: More than IQ Scores. *Mental Retardation and Developmental Disabilities Research Review*, 8(4), 234-240.
- Baddeley, A. (1983). Working Memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 302(1110), 311-324.
- Baddeley, A. (1986). *Working Memory*. Oxford Oxfordshire/New York: Clarendon Press ; Oxford University Press.

- Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal Of Experimental Psychology*, 49A(1), 5-28.
- Baddeley, A. (1998). The Central Executive: A Concept and Some Misconceptions. *Journal of the International Neuropsychological Society*, 4(5), 523-526.
- Baddeley, A., & Hitch, G. (1974). Working Memory. In I. G. A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47-90). New York: Academic Press.
- Barkley, R. A. (1997). *ADHD and the Nature of Self-Control*. New York: Guilford Press.
- Barkley, R. A., Murphy, K. R., Dupaul, G. I., & Bush, T. (2002). Driving In Young Adults with Attention Deficit Hyperactivity Disorder: Knowledge, Performance, Adverse Outcomes, and the Role of Executive Functioning. *Journal of the International Neuropsychological Society*, 8(5), 655-672.
- Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-Making Deficits, Linked to a Dysfunctional Ventromedial Prefrontal Cortex, Revealed in Alcohol and Stimulant Abusers. *Neuropsychologia*, 39(4), 376-389.
- Bechara, A., & Martin, E. M. (2004). Impaired Decision Making Related to Working Memory Deficits in Individuals with Substance Addictions. *Neuropsychology*, 18(1), 152-162.
- Becker, M. G., Isaac, W., & Hynd, G. W. (1987). Neuropsychological Development of Nonverbal Behaviors Attributed to "Frontal Lobe" Functioning. *Developmental Neuropsychology*, 3(3 & 4), 275-298.
- Bialystok, E., & Martin, M. M. (2003). Notation To Symbol: Development In Children's Understanding Of Print. *Journal of Experimental Child Psychology*, 86(3), 223-243.
- Bolla, K. I., Cadet, J. L., & London, E. D. (1998). The Neuropsychiatry Of Chronic Cocaine Abuse. *Journal of Neuropsychiatry and Clinical Neurosciences*, 10(3), 280-289.

- Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., Et Al. (1998).
The Reliability And Internal Validity Of The Wisconsin Card Sorting Test.
Neuropsychological Rehabilitation, 8(3), 243-254.
- Brand, M., Kalbe, E., Labudda, K., Fujiwara, E., Kessler, J., & Markowitsch, H. J. (2005).
Decision-Making Impairments In Patients With Pathological Gambling. *Psychiatry
Research*, 133(1), 91-99.
- Brand, M., Labudda, K., Kalbe, E., Hilker, R., Emmans, D., Fuchs, G., Et Al. (2004). Decision-
Making Impairments In Patients With Parkinson's Disease. *Behavioral Neurology*, 15(3-
4), 77-85.
- Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The Ecological
Validity Of Tests Of Executive Function. *Journal of the International
Neuropsychological Society*, 4(6), 547-558.
- Canivez, G. L., & Watkins, M. W. (2001). Long-Term Stability Of The Wechsler Intelligence
Scale For Children-Third Edition Among Students With Disabilities. *School Psychology
Review*, 30(2), 438-453.
- Chamberlaine, E. (2003). Behavioural Assessment Of The Dysexecutive Syndrome (Bads) - Test
Review. *Journal of Occupational Psychology, Employment and Disability*, 5(2), 33-37.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-Of-Fit Indexes For Testing
Measurement Invariance. *Structural Equation Modeling*, 9(2), 233-255.
- Clark, C., Prior, M., & Kinsella, G. (2002). The Relationship Between Executive Function
Abilities, Adaptive Behaviour, And Academic Achievement In Children With
Externalising Behaviour Problems. *Journal of Child Psychology and Psychiatry*, 43(6),
785-796.

- Clark, L., & Manes, F. (2004). Social And Emotional Decision-Making Following Frontal Lobe Injury. *Neurocase*, 10(5), 398-403.
- Cohen, M. J., Riccio, C. A., Kibby, M. Y., & Edmonds, J. E. (2000). Developmental Progression Of Clock Face Drawing In Children. *Child Neuropsychology*, 6(1), 64-76.
- Conners, C. K. (1995). *Conners Continuous Performance Test*. Toronto: Multi-Health Systems.
- Cordon, S. L., & Finney, S. J. (2006). *Measurement Invariance of the Mindful Attention Awareness Scale Across Adult Attachment Style*. Unpublished Manuscript.
- Crocker, L. M., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: Holt Rinehart And Winston.
- Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal Executive And Cognitive Functions In Rodents: Neural And Neurochemical Substrates. *Neurosciences and Biobehavioral Review*, 28(7), 771-784.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development Of Cognitive Control And Executive Functions From 4 To 13 Years: Evidence From Manipulations Of Memory, Inhibition, And Task Switching. *Neuropsychologia*, 44(11), 2037-2078.
- Decarlo, L. T. (1997). On The Meaning And Use Of Kurtosis. *Psychological Methods*, 2(3), 292-307.
- Demakis, G. J. (2003). A Meta-Analytic Review Of The Sensitivity Of The Wisconsin Card Sorting Test To Frontal And Lateralized Frontal Brain Damage. *Neuropsychology*, 17(2), 255-264.
- Denckla, M. B. (1996a). Research On Executive Function In A Neurodevelopmental Context: Application Of Clinical Measures. *Developmental Neuropsychology*, 12(1), 5-15.

- Denckla, M. B. (1996b). A Theory And Model Of Executive Function: A Neuropsychological Perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, Memory, And Executive Function* (Pp. 263-278). Baltimore, Md: P.H. Brookes.
- Denckla, M. B. (2002). The Behavior Rating Inventory Of Executive Function: Commentary. *Child Neuropsychology*, 8(4), 304-306.
- Denckla, M. B. (2007). Executive Function: Binding Together The Definitions Of Attention-Deficit/Hyperactivity Disorder And Learning Disabilities. In L. Melltzer (Ed.), *Executive Function In Education* (Pp. 5-18). New York: Guilford.
- Duff, S. C. (2000). What's Working In Working Memory: A Role For The Central Executive. *Scandinavian Journal of Psychology*, 41(1), 9-16.
- Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid Intelligence after Frontal Lobe Lesions. *Neuropsychologia*, 33(3), 261-268.
- Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the Frontal Lobe: The Organization Of Goal-Directed Behavior. *Cognitive Psychology*, 30(3), 257-303.
- Duncan, J., Johnson, R., Swales, M., & Freer, C. (1997). Frontal Lobe Deficits After Head Injury: Unity And Diversity Of Function. *Cognitive Neuropsychology*, 14(5), 713-741.
- Elliott, R. (2003). Executive Functions And Their Disorders. *British Medical Bulletin*, 65, 49-59.
- Enders, C. K., & Bandalos, D. L. (2001). The Relative Performance Of Full Informacion Maximum Likelihood Estimation For Missing Data In Structural Equation Models. *Structural Equation Modeling*, 8(3), 430-457.

- Ernst, M., Kimes, A. S., London, E. D., Matochik, J. A., Eldreth, D., Tata, S., Et Al. (2003). Neural Substrates Of Decision Making In Adults With Attention Deficit Hyperactivity Disorder. *American Journal of Psychiatry*, 160(6), 1061-1070.
- Espy, K. A. (2004). Using Developmental, Cognitive, And Neuroscience Approaches To Understand Executive Control In Young Children. *Developmental Neuropsychology*, 26(1), 379-384.
- Espy, K. A., Kaufmann, P. M., & Glisky, M. L. (2001). New Procedures To Assess Executive Functions In Preschool Children. The *Clinical Neuropsychologist*, 15(1), 46-58.
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The Contribution Of Executive Functions To Emergent Mathematic Skills In Preschool Children. *Developmental Neuropsychology*, 26(1), 465-486.
- Evans, J. J., Chua, S. E., McKenna, P. J., & Wilson, B. A. (1997). Assessment Of The Dysexecutive Syndrome In Schizophrenia. *Psychological Medicine*, 27(3), 635-646.
- Ewing-Cobbs, L., Prasad, M. R., Landry, S. H., Kramer, L., & DeLeon, R. (2004). Executive Functions Following Traumatic Brain Injury In Young Children: A Preliminary Analysis. *Developmental Neuropsychology*, 26(1), 487-512.
- Finney, S. J., & Distefano, C. (2006). Nonnormal And Categorical Data In Structural Equation Models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling. A Second Course* (Pp. 269-314). Greenwich, Ct: Information Age.
- Fletcher, P. C., Shallice, T., & Dolan, R. J. (1998). The Functional Roles Of Prefrontal Cortex In Episodic Memory. I. Encoding. *Brain*, 121 (Pt 7), 1239-1248.

- Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1998). The Functional Roles Of Prefrontal Cortex In Episodic Memory. Ii. Retrieval. *Brain*, *121* (Pt 7), 1249-1256.
- Flora, D. B., & Curran, P. J. (2004). An Empirical Evaluation Of Alternative Methods Of Estimation Fro Confirmatory Factor Analysis With Ordinal Data. *Psychological Methods*, *9*(4), 466-491.
- Frazier, T. W., & Youngstrom, E. A. (2007). Historical Increase In The Number Of Factors Measured By Commercial Tests Of Cognitive Ability: Are We Overfactoring? *Intelligence*, *35*, 169-182.
- French, B. F., & Finch, W. H. (2006). Confirmatory Factor Analytic Procedures For The Determination Of Measurement Invariance. *Structural Equation Modeling*, *13*(3), 378-402.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not All Executive Functions Are Related To Intelligence. *Psychological Science*, *17*(2), 172-179.
- Fuster, J. M. (1995). Memory And Planning. Two Temporal Perspectives Of Frontal Lobe Function. *Advances in Neurology*, *66*, 9-19; Discussion 19-20.
- Fuster, J. M. (1997). *The Prefrontal Cortex - Anatomy, Physiology, And Neuropsychology Of The Frontal Lobe*. Philadelphia: Lippincott-Raven.
- Fuster, J. M. (2002). Frontal Lobe And Cognitive Development. *Journal of Neurocytology*, *31*(3-5), 373-385.

- Gilotty, L., Kenworthy, L., Sirian, L., Black, D. O., & Wagner, A. E. (2002). Adaptive Skills And Executive Function In Autism Spectrum Disorders. *Child Neuropsychology*, 8(4), 241-248.
- Gioia, G. A., & Isquith, P. K. (2004). Ecological Assessment Of Executive Function In Traumatic Brain Injury. *Developmental Neuropsychology*, 25(1-2), 135-158.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Brief: Behavior Rating Inventory Of Executive Functions*. Odessa, Fl: Psychological Assessment Resources.
- Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles Of Everyday Executive Function In Acquired And Developmental Disorders. *Child Neuropsychology*, 8(2), 121-137.
- Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Espy, K. A. (2002). Confirmatory Factor Analysis Of The Behavior Rating Inventory Of Executive Function (Brief) In A Clinical Sample. *Child Neuropsychology*, 8(4), 249-257.
- Glutting, J. J., Mcdermott, P. A., Konold, T. R., Snelbaker, A. J., & Watkins, M. W. (1998). More Ups And Downs Of Subtest Analysis: Criterion Validity Of The DAS With An Unselected Cohort. *School Psychology Review*, 27, 599-612.
- Godefroy, O., Cabaret, M., Petit-Chenal, V., Pruvo, J. P., & Rousseaux, M. (1999). Control Functions Of The Frontal Lobes. Modularity Of The Central-Supervisory System? *Cortex*, 35(1), 1-20.
- Goldman-Rakic, P. S., Bates, J. F., & Chafee, M. V. (1992). The Prefrontal Cortex And Internally Generated Motor Acts. *Current Opinion in Neurobiology*, 2(6), 830-835.
- Gordon, M. (1983). *The Gordon Diagnostic System*. Dewitt, Ny: Gordon Systems.

- Gordon, M. (1991). *Instruction Manual For The Gordon Diagnostic System, Model Iii-R*. Dewitt, Ny: Gordon Systems.
- Grace, J., & Malloy, P. (2001). *Frontal Systems Behavior Scale (Frsbe): Professional Manual*. Lutz, Fl: Psychological Assessment Resources.
- Grafman, J., & Litvan, I. (1999). Importance Of Deficits In Executive Functions. *Lancet*, 354(9194), 1921-1923.
- Greve, K. W., Brooks, J., Crouch, J. A., Williams, M. C., & Rice, W. J. (1997). Factorial Structure Of The Wisconsin Card Sorting Test. *British Journal of Clinical Psychology*, 36 (Pt 2), 283-285.
- Greve, K. W., Stickley, T. R., Love, J. M., Bianchini, K. J., & Stanford, M. S. (2005). Latent Structure Of The Wisconsin Card Sorting Test: A Confirmatory Factor Analytic Study. *Archives of Clinical Neuropsychology*, 20(3), 355-364.
- Harlow, J. M. (1999). Passage Of An Iron Rod Through The Head. 1848. *The Journal Of Neuropsychiatry And Clinical Neurosciences*, 11(2), 281-283.
- Heaton, R. K. (1981). *Wisconsin Card Sorting Test (Wcst)*. Odessa, Fl: Psychological Assessment Resources.
- Hoyle, R. H., & Panter, A. T. (1995). Writing About Structural Equation Models. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, And Applications* (Pp. 158-176). Thousand Oaks: Sage.
- Hu, L., & Bentler, P. M. (1999). Cutoff Criteria For Fit Indexes In Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling*, 6(1), 1-55.

- Hughes, C., & Graham, A. (2002). Measuring Executive Functions In Childhood: Problems And Solutions? *Child And Adolescent Mental Health*, 7(3), 131-142.
- Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence For Executive Dysfunction In Autism. *Neuropsychologia*, 32(4), 477-492.
- Hunter, J. E., & Hunter, R. V. (1984). Validity And Utility Of Alternate Predictors Of Job Performance. *Psychological Bulletin*, 96(72-98).
- Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive Function In Preschool Children: Examination Through Everyday Behavior. *Developmental Neuropsychology*, 26(1), 403-422.
- Jarratt, K. P., Riccio, C. A., & Siekierski, B. M. (2005). Assessment Of Attention Deficit Hyperactivity Disorder (Adhd) Using The Basc And Brief. *Applied Neuropsychology*, 12(2), 83-93.
- Jöreskog, K. G., & Söborm, S. (1993). Lisrel 8: Structural Equation Modeling With The Simplis Command Language. (Version 8.54). Hillsdale, N.J.: Scientific Software International.
- Kafer, K. L., & Hunter, M. (1997). On Testing The Face Validity Of Planning/Problem-Solving Tasks In A Normal Population. *Journal of the International Neuropsychological Society*, 3(2), 108-119.
- Kamphaus, R. W. (2001). *Clinical Assessment Of Child And Adolescent Intelligence* (2nd Ed.). Boston: Allyn & Bacon.
- Kerr, A., & Zelazo, P. D. (2004). Development Of "Hot" Executive Function: The Children's Gambling Task. *Brain and Cognition*, 55(1), 148-157.
- Kertesz, A., Davidson, W., & Fox, H. (1997). Frontal Behavioral Inventory: Diagnostic Criteria For Frontal Lobe Dementia. *Canadian Journal Neurological Sciences*, 24(1), 29-36.

- Kibby, M. Y., Cohen, M. J., & Hynd, G. W. (1999). Clock Drawing Performance In Children With Attention Deficit Hyperactivity Disorder. *Archives Of Clinical Neuropsychology*, 14(8), 690-691.
- Kibby, M. Y., Cohen, M. J., & Hynd, G. W. (2002). Clock Face Drawing In Children With Attention-Deficit/Hyperactivity Disorder. *Archives of Clinical Neuropsychology*, 17(6), 531-546.
- Kim, K. H., Cramond, B., & Bandalos, D. L. (2006). The Latent Structure And Measurement Invariance Of Scores On The Torrance Test Of Creative Thinking-Figural. *Educational And Psychological Measurement*, 66(3), 459-477.
- Kline, R. B. (2005). *Principles And Practice Of Structural Equation Modeling* (2nd Ed.). New York: Guilford.
- Korkman, M., Kirk, U., & Kemp, S. (1998). *Nepsy: A Developmental Neuropsychological Assessment, Manual*. San Antonio, Tx: Psychological Corporation.
- Kray, J., Eber, J., & Lindenberger, U. (2004). Age Differences In Executive Functioning Across The Lifespan: The Role Of Verbalization In Task Preparation. *Acta Psychologica (Amsterdam)*, 115(2-3), 143-165.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions Of Executive Functioning: Evidence From Children. *British Journal Of Developmental Psychology*, 21, 59-80.
- Lezak, M. D. (1982). The Problems Of Assessing Executive Functions. *International Journal Of Psychology*, 17, 281-297.
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment*. New York: Oxford University Press.

- Logan, G. D. (1994). On The Ability To Inhibit Thought And Action: A User's Guide To The Stop Signal Paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory Processes In Attention, Memory, And Language*. San Diego, Ca: Academy Press.
- London, E. D., Ernst, M., Grant, S., Bonson, K., & Weinstein, A. (2000). Orbitofrontal Cortex And Human Drug Abuse: Functional Imaging. *Cerebral Cortex*, 10(3), 334-342.
- Luria, A. R. (1973). *The Working Brain: An Introduction To Neuropsychology*. New York: Basic Books.
- Mahone, E. M., Zabel, T. A., Levey, E., Verda, M., & Kinsman, S. (2002). Parent And Self-Report Ratings Of Executive Function In Adolescents With Myelomeningocele And Hydrocephalus. *Child Neuropsychology*, 8(4), 258-270.
- Malloy, P., & Grace, J. (2005). A Review Of Rating Scales For Measuring Behavior Change Due To Frontal Systems Damage. *Cognitive and Behavioral Neurology*, 18(1), 18-27.
- Manchester, D., Priestley, N., & Jackson, H. (2004). The Assessment Of Executive Functions: Coming Out Of The Office. *Brain Injury*, 18(11), 1067-1081.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., Et Al. (2002). Decision-Making Processes Following Damage To The Prefrontal Cortex. *Brain*, 125(Pt 3), 624-639.
- Mani, T. M., Bedwell, J. S., & Miller, L. S. (2005). Age-Related Decrements In Performance On A Brief Continuous Performance Test. *Archives of Clinical Neuropsychology*, 20(5), 575-586.
- Milner, B. (1982). Some Cognitive Effects Of Frontal-Lobe Lesions In Man. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 298(1089), 211-226.

- Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis Of The Elements Of Attention: A Neuropsychological Approach. *Neuropsychology Review*, 2(2), 109-145.
- Miyake, A., Emerson, M. J., & Friedman, N. P. (2000). Assessment Of Executive Functions In Clinical Settings: Problems And Recommendations. *Seminars in Speech and Language*, 21(2), 169-183.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity And Diversity Of Executive Functions And Their Contributions To Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41(1), 49-100.
- Moritz, S., Birkner, C., Kloss, M., Jahn, H., Hand, I., Haasen, C., Et Al. (2002). Executive Functioning In Obsessive-Compulsive Disorder, Unipolar Depression, And Schizophrenia. *Archives of Clinical Neuropsychology*, 17(5), 477-483.
- Muthén, B. O., & Asparouhov, T. (2002). Latent Variable Analysis With Categorical Outcomes: Multiple-Group And Growth Modeling In Mplus [Electronic Version]. *Mplus Web Notes*, 22 From <http://www.statmodel.com/examples/webnote.shtml>.
- Muthén, L. K., & Muthén, B. O. (1998-2006). *Mplus User's Guide. Fourth Edition*. Los Angeles, Ca: Muthén & Muthén.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Ceci, S. J., & Al, E. (1996). Intelligence: Knowns And Unknowns. *American Psychologist*, 51(2), 77-101.
- Neylan, T. C. (1999). Frontal Lobe Function: Mr. Phineas Gage's Famous Injury. *The Journal Of Neuropsychiatry And Clinical Neurosciences*, 11(2), 280-281.

- Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological Development Of Behavior Attributed To Frontal Lobe Functioning In Children. *Developmental Neuropsychology*, 1(4), 349-370.
- Pineda, D. A., Aguirre, D. C., Garcia, M. A., Lopera, F. J., Palacio, L. G., & Kamphaus, R. W. (2005). Validation Of Two Rating Scales For Attention-Deficit Hyperactivity Disorder Diagnosis In Colombian Children. *Pediatric Neurology*, 33(1), 15-25.
- Pineda, D. A., Henao, G. C., Puerta, I. C., Mejia, S. E., Gomez, L. F., Miranda, M. L., Et Al. (1999). [The Use Of Brief Questionnaire In The Diagnosis Of Attention Deficit. Study Group Of The Manizales University Foundation]. *Revista de Neurologia*, 28(4), 365-372.
- Pineda, D. A., Kamphaus, R. W., Mora, O., Puerta, I. C., Palacio, L. G., Jimenez, I., Et Al. (1999). [Use Of Multidimensional Scale For Parents Of Children Aged 6 To 11 For The Diagnosis Of Attention Deficit With Hyperactivity]. *Revista de Neurologia*, 28(10), 952-959.
- Pineda, D. A., Kamphaus, R. W., Mora, O., Restrepo, M. A., Puerta, I. C., Palacio, L. G., Et Al. (1999). [A System Of Multidimensional Behavior Assessment. A Scale For Parents Of Children From 6 To 11 Years Of Age. Colombian Version]. *Revista de Neurologia*, 28(7), 672-681.
- Pineda, D. A., & Merchan, V. (2003). Executive Function In Young Colombian Adults. *International Journal of Neurosciences*, 113(3), 397-410.
- Posner, M. I. (In Press). *Attention As An Organ*. Paper Presented At The Neurobiology Of Perception And Communication: From Synapse To Society.

Posner, M. I., & Rothbart, M. K. (1998). Attention, Self-Regulation And Consciousness.

Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 353(1377), 1915-1927.

Rabbitt, P. (1997). *Methodology Of Frontal And Executive Function*. Hove, East Sussex: Psychology Press.

Raguet, M. L., Campbell, D. A., Berry, D. T. R., Schmitt, F. A., & Smith, G. T. (1996). Stability Of Intelligence And Intellectual Predictors In Older Persons. *Psychological Assessment*, 8(2), 154-160.

Rahman, S., Sahakian, B. J., Hodges, J. R., Rogers, R. D., & Robbins, T. W. (1999). Specific Cognitive Deficits In Mild Frontal Variant Frontotemporal Dementia. *Brain*, 122 (Pt 8), 1469-1493.

Rensvold, R. B., & Cheung, G. (2001). Testing For Metric Invariance Using Structural Equation Models: Solving The Standardization Problem. In C. A. Schriesheim & L. L. Neider (Eds.), *Equivalence In Measurement* (Vol. 1, Pp. 25-50). Greenwich, Ct: Information Age.

Rey, A., & Osterrieth, P. A. (1993). Translations Of Excerpts From Andre Rey's Psychological Examination Of Traumatic Encephalopathy, And P.A. Osterrieth's The Complex Figure Copy Test. *The Clinical Neuropsychologist*, 7, 4-21.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System For Children (Basc)*. Circle Pines, Mn: American Guidance Services (Ags).

Reynolds, C. R., & Kamphaus, R. W. (2002). *The Clinician's Guide To The Behavior Assessment System For Children*. New York: The Guilford Press.

- Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System For Children -2nd Edition*. Bloomington, Mn: Pearson Assessment.
- Riccio, C. A., Cohen, M. J., Hynd, G. W., & Keith, R. W. (1996). Validity Of The Auditory Continuous Performance Test In Differentiating Central Processing Auditory Disorders With And Without Adhd. *Journal Of Learning Disabilities*, 29(5), 561-566.
- Riccio, C. A., Hall, A. J., Morgan, A., Hynd, G. W., Gonzalez, J. J., & Marshall, R. M. (1994). Executive Function And Wisconsin Card Sorting Test: Relationship With Behavioral Rating And Cognitive Ability. *Developmental Neuropsychology*, 10, 215-229.
- Riccio, C. A., & Reynolds, C. R. (2001). Continuous Performance Tests Are Sensitive To Adhd In Adults But Lack Specificity. A Review And Critique For Differential Diagnosis. *Annals Of The New York Academy Of Sciences*, 931, 113-139.
- Riccio, C. A., Reynolds, C. R., Lowe, P., & Moore, J. J. (2002). The Continuous Performance Test: A Window On The Neural Substrates For Attention? *Archives of Clinical Neuropsychology*, 17(3), 235-272.
- Riccio, C. A., Reynolds, C. R., & Lowe, P. A. (2001). *Clinical Applications Of Continuous Performance Tests. Measuring Attention And Impulsive Responding In Children And Adults*. New York: John Wiley & Sons, Inc.
- Robbins, T. W. (1996). Dissociating Executive Functions Of The Prefrontal Cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 351(1346), 1463-1470; Discussion 1470-1461.
- Rolls, E. T. (2002). The Functions Of The Orbitofrontal Cortex. In D. T. Stuss & R. T. Knight (Eds.), *Principles Of Frontal Lobe Function* (Pp. 354-375). New York: Oxford.

- Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., & Beck, L. H. (1956). A Continuous Performance Test Of Brain Damage. *Journal of Consulting Psychology*, 20, 343-350.
- Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., Et Al. (2002). Executive Control Function: A Review Of Its Promise And Challenges For Clinical Research. A Report From The Committee On Research Of The American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neurosciences*, 14(4), 377-405.
- Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The Development Of Executive Attention: Contributions To The Emergence Of Self-Regulation. *Developmental Neuropsychology*, 28(2), 573-594.
- Rusticus, S. A., & Hubley, A. M. (2006). Measurement Invariance Of The Multidimensional Body-Self Relations Questionnaire: Can We Compare Across Age And Gender? *Sex Roles*, 55, 827-842.
- Sandoval, J., & Echandia, A. (1994). Behavior Assessment System For Children. *Journal Of School Psychology*, 32(4), 419-425.
- Scott, S. K., Holmes, A., Friston, K. J., & Wise, R. J. (2000). A Thalamo-Prefrontal System For Representation In Executive Response Choice. *Neuroreport*, 11(7), 1523-1527.
- Senn, T. E., Espy, K. A., & Kaufmann, P. M. (2004). Using Path Analysis To Understand Executive Function Organization In Preschool Children. *Developmental Neuropsychology*, 26(1), 445-464.
- Shallice, T. (1982). Specific Impairments Of Planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 298(1089), 199-209.

- Shallice, T., & Burgess, P. (1996). The Domain Of Supervisory Processes And Temporal Organization Of Behaviour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 351(1346), 1405-1411; Discussion 1411-1402.
- Shallice, T., & Burgess, P. W. (1991). Deficits In Strategy Application Following Frontal Lobe Damage In Man. *Brain*, 114 (Pt 2), 727-741.
- Strayhorn, J. M. (1993). Statistical Case Puzzle. The Case Of The Despondent Student. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32(2), 466-467.
- Stuss, D. T., & Benson, D. F. (1986). *The Frontal Lobes*. New York: Raven Press.
- Suchy, Y., Leahy, B., Sweet, J. J., & Lam, C. S. (2003). Behavioral Dyscontrol Scale Deficits Among Traumatic Brain Injury Patients, Part Ii: Comparison To Other Measures Of Executive Functioning. *The Clinical Neuropsychologist*, 17(4), 492-506.
- Sullivan, J. R., & Riccio, C. A. (2006). An Empirical Analysis Of The BASC Frontal Lobe/Executive Control Scale With A Clinical Sample. *Archives Of Clinical Neuropsychology*, 21(5), 495-501.
- Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation Of Brain Function Associated With Response Inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(10), 1231-1238.
- Teuber, H. L. (1972). Unity And Diversity Of Frontal Lobe Functions. *Acta Neurobiologiae Experimentalis (Wars)*, 32(2), 615-656.
- Tirapu-Ustarroz, J., Munoz-Cespedes, J. M., & Pelegrin-Valero, C. (2002). [Executive Functions: The Need For The Integration Of Concepts]. *Revista de Neurologia*, 34(7), 673-685.

- Watanabe, K., Ogino, T., Nakano, K., Hattori, J., Kado, Y., Sanada, S., Et Al. (2005). The Rey-Osterrieth Complex Figure As A Measure Of Executive Function In Childhood. *Brain Dev*, 27(8), 564-569.
- Watkins, L. H., Rogers, R. D., Lawrence, A. D., Sahakian, B. J., Rosser, A. E., & Robbins, T. W. (2000). Impaired Planning But Intact Decision Making In Early Huntington's Disease: Implications For Specific Fronto-Striatal Pathology. *Neuropsychologia*, 38(8), 1112-1125.
- Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A Normative-Developmental Study Of Executive Function: A Window On Prefrontal Function In Children. *Developmental Neuropsychology*, 7, 131-149.
- Wilson, B. A., Evans, J. J., Emslie, H., Alderman, N., & Burgess, P. (1998). The Development Of An Ecologically Valid Test For Assessing Patients With A Dysexecutive Syndrome. *Neuropsychological Rehabilitation*, 8(3), 213-228.