UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A COMPARISON OF CONTEMPORARY EXECUTIVE FUNCTION MODELS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MICHAEL AARON FANEROS
Norman, Oklahoma
2014

A COMPARISON OF CONTEMPORARY EXECUTIVE FUNCTION MODELS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY

Dr. Kirby Gilliland, Chair

Dr. Randa Shehab

Dr. Eugenia Cox-Fuenzalida

Dr. Scott Gronlund

Dr. Hairong Song

For Ila and Erik.

Acknowledgements

I give my thanks to my committee members who supported me and nurtured my research, Eugenia-Luz Cox-Fuenzilida, Ph.D., Scott Gronlund, Ph.D., Randa Shehab, Ph.D., and Hairong Song, Ph.D. I give special thanks to my mentor and committee chair, Kirby Gilliland, Ph.D., for this unending support and motivation. Andrea Vincent, Ph.D., has my gratitude for the use of lab equipment and statistical advice over the years. My heartfelt thanks go to my parents, who have always believed in me. Lastly, the words "thank you" are not enough to convey my appreciation for my wife Ila and son Erik in sharing this journey with me and for without whom life would be a little dimmer.

Table of Contents

List of Tables

List of Figures

Abstract

Research investigating executive function is difficult in two different ways. First, executive function has only recently evolved into its own field of study, and unlike the fields of personality or intelligence, it therefore does not have a wide body of prior research to draw upon. The second difficulty follows from this limited knowledge base. The many different approaches to the study of executive function have given a diverse set of ways to define and operationalize it, leading to numerous contradictions in the literature. This study addresses these contradictions by examining three different theories and four models related to those theories in an effort to understand how each fits the same data set. Models from Miyake and Friedman's (2012) and Faneros's (2011) factorial models, Peterson and Posner's (2012) biological model, and Barkley's (2012) clinical model are each described and analyzed using confirmatory factor analysis. Additionally, path analysis was used with each model to assess the model's stability and strength at predicting performance on a complex executive function task. The results indicate Peterson and Posner's theory provided a model that proved to have the best fit and was the most parsimonious of all the proposed models, χ^2 (10, N = 306) = 7.047, p =.72, CFI = 1.000, SRMR = 0.024, BIC = 4692. Each model's strengths and weakness were explored, and future directions considered.

Keywords: executive functioning, inhibition, attention

A Comparison of Contemporary Executive Function Models

Research investigating executive function is challenging to measure for two reasons. First, research exploring executive function is relatively recent. Though aspects attributed to the contemporary understanding of executive function date back to studies of the prefrontal cortex by Harlow in 1848, the term and concept of executive function does not appear in the literature until Pribram published his 1973 book chapter, "The Primate Frontal Cortex: Executive Of The Brain." And though the conception of executive function was initially defined in 1973, the area did not begin to generate a notable amount of research until the mid-1990s. Since then the amount of research published each year on executive function has continued to grow expanding our understanding of the concept, and diminishing the shortage of information in the field of study. This fast expansion in research has provided many insights, but also has led to contradictory evidence for the fundamental nature of executive function.

The second problem associated with executive function is how to define and thus operationalize executive function. There are many different definitions of the theoretical construct, many of which are contradictory or include concepts others do not address (a sample can be found in Barkley, 2012, p. 5-7). The lack of a common definition for executive function creates a situation where some definitions agree and even reference other conceptualizations of executive function, but also often contradict each other. Yet, the diversity of approaches provides strength in understanding executive functions as well. These different approaches address issues central to a particular paradigm that would be lost in others.

Executive function remains a vital topic to study despite these challenges. No matter what the definition, researchers agree that executive functions are those processes that contribute directly to tasks encountered in everyday life. Flying a plane, running a business, building a house, and grocery shopping are just a few tasks that require executive functioning. Neurological injury has been shown to directly impact executive functioning (Harlow, 1848), and executive functioning deficits cause a number of psychological problems (Barkley, 2012). Given the broad implications, understanding individual differences in executive function and how they relate to overall behavior are questions that drive research such as the study presented here. A study examining the predominant theories of executive function would help to highlight the weaknesses, strengths, and overlap in those theories, and provide evidence for which theory might best represent the domain of executive function.

This study intends to test the relationships and viability among four models drawn from three theories of executive function. Testing these models will highlight the weaknesses and strengths in each theory and reveal commonalities across theories. Since these theories had not been studied together before, a prediction of which model is stronger in prediction could not be made. Rather, the working assumption was each model will fit the data as predicted by the relevant theory and be able to predict complex executive function equally. Three dominant theories for conceptualizing and defining executive function are reviewed and then four models are developed for testing the viability of the measurement techniques of those theories are presented.

Theories of Executive Function

Reviewed here are three different theoretical approaches and four models of executive function and their corresponding definitions. While these theories approach the study of executive function quite differently and thus have differences in their resulting models, they also share some similarities. Both the similarities and differences will be discussed after a brief introduction to each. Broadly speaking, these theories can reflect a factorial or psychometric approach, an observational approach based on biological functioning, and a clinical approach, respectively.

Miyake & Friedman (2012): The Factorial Approach. Executive functions are typically defined as the interrelated cognitive processes specifically involved with data manipulation and attentional control. Friedman et al. (2007, p. 893) defined executive functions as “a family of cognitive control processes that operate on lower-level processes to regulate and shape behavior.” Updating, task shifting, and inhibition are three theoretical executive function processes that Miyake, Friedman, Emerson, Witzki, and Howerter (2000) focused on because these dimensions have a significant research history in the existing executive function literature. Updating describes an individual’s ability to monitor, code, and revise information for a given task. Shifting describes an individual’s ability to shift between tasks, operations, or mental sets. Inhibition describes an individual’s ability to deliberately inhibit prepotent responses when necessary. These dimensions are easily measured experimentally, and Miyake et al. have shown current measures of these functions adhere to a confirmatory factor structure. Additionally, updating, shifting, and inhibition have neurological correlates suggesting a biological basis for the theoretical processes (Miyake et al., 2000; Morris & Jones, 1990; Posner & Raichle, 1994).

In an update to their 2000 article, Miyake and Friedman (2012) reevaluated their model of executive function. In the new model, inhibition is part of a larger "common executive function" factor while updating and task switching account for specific types of executive functioning. This new model describes what Miyake and Freidman call the "unity and diversity" of executive functions; executive functions rely on some underlying general ability (unity) while at the same time are separable in measurement (diversity). This model loses some of the original model's explanatory power by merging inhibition with the rest of the variance of the model. The ability to measure the new general executive function factor is further complicated by merging inhibition with the variance from the three dimensions. The only way to measure this dimension now, according to the 2012 model, is to give a test battery that tests all sub-domains of executive function. Without the other task performance, there is no way to specify how the general executive function dimension contributes to performance as a whole, because part of it relies on the overall model variance. Additionally, inhibition in the new model is conflated with the model variance so that understanding how inhibition contributes to the model in general is lost.

The model presented by Miyake et al. (2000) has been used to identify individual differences in important psychological phenomena. For example, Altamirano, Miyake, & Whitmer (2010) found that people who tend to ruminate, an indicator for a predisposition for depression, also have higher errors in shifting tasks but fewer errors in updating tasks. Friedman et al. (2007) found teacher evaluations of attentional control of children ages 7 to 14 years were correlated with the three executive function processes in the Miyake et al. 2000 model. This result suggested attention problems were directly related to response inhibition. Friedman, Miyake, Robinson, and Hewitt (2011) followed up with a twin study investigating toddler's self-restraint and executive function 14 years later. They found that, in accordance with Miyake and Friedman's (2012) refinement of the model, less restrained children had a lower common executive function factor but were equivalent on the factors updating and shifting. Additionally, the twin analysis suggested a strong genetic component in the individual differences in self-restrain and executive function processes.

Miyake and Friedman (2012) utilized a battery of nine measures to research executive function: Letter Memory Task, Keep Track Task, Spatial 2-Back Task, Color-Shape Task, Number-Letter Task, Category-Switch Task, Antisaccade Task, Stroop, and Stop-Signal task (see Friedman et al., 2008, for specific descriptions of each task). These nine measures were then subjected to a confirmatory factor analysis (CFA) to support the conceptualized model of executive function. However, Faneros (2011) attempted to replicate Miyake et al. (2000) by using measures used in their original study as well as tasks from the Automated Neuropsychological Assessment Metrics (ANAM) battery to show correspondence to the three conceptual areas of executive function. In the Faneros study, the resulting CFA model failed to fit the Miyake et al. model. This result demonstrates a pitfall of the factorial approach in that the model and theory are only as good as the measures chosen to measure the constructs. A post-hoc exploratory factor analysis revealed a structure that grouped the Miyake et al. tasks into the updating factor and the ANAM tasks spread across three factors. Therefore there is still a strong possibility of a three-factor approach to executive function, but it may not be that of Miyake et al.

The results of the Faneros (2011) study suggest the factorial model of executive function might best be assessed using different measures than the tasks used by Miyake et al. (2000). The model derived using exploratory factor analysis found the selected tasks from Miyake et al. constrained to one factor while the other tasks chose by Faneros differentiated to three factors. The resulting model from Faneros did not match the model restructuring found in Miyake and Friedman (2012). Because factor analysis is dependent on the measures used and the performance of a particular sample, models often do not fit the same from experiment to experiment. The strength of the Faneros model, however, is that the same results came from both an orthogonal and oblique solution. The re-evaluation of the two models in this study will give more clarity to the theoretical structure of executive function.

Petersen & Posner (2012): The Biological Approach. Another executive function theory examines a biological approach. In 1990, Posner and Petersen presented a theory of attention based on the current knowledge of the biology of the brain. Twenty years later, they reviewed the research that led to an expansion and elaboration of the original framework. Petersen and Posner (2012) describe their theory of attention as three integrated networks: an alerting system, orienting mechanisms, and executive function. The alerting system is simply the physical mechanisms that allow for a stimulus to be recognized, such as the eye converting the light hitting the retina into electrical impulses that are relayed to the brain. This also includes the brain stem arousal system that is directly related to sustained attention tasks. Orienting mechanisms are the mental instructions and motor movements of the body to orient towards a particular stimulus (e.g., leaning forward and squinting at some far off object to better see it). Covert and overt shifts of attention are included in this process. There is some evidence that, while these systems are very integrated, there are distinct groups of cells responsible for each operation.

Executive function in the Petersen and Posner (2012) theory is described as focused attention, which is subject to an individual's limited attentional capacity. When a stimulus captures awareness in this theory, awareness for all other stimuli decreases slightly. This produces interference in the cognitive system, both on the initial stimulus captured (some attentional capacity will still be allocated to those items in the background) and the tracking of suppressed stimuli that would otherwise retain some cognitive load. This capturing of awareness relies on the idea of limited attentional capacity and focal attention (Jonides, 1983). Focal attention can be thought of as the conscious awareness of some object or idea at some cost to awareness to other items, much like a spotlight on a stage. One actor becomes illuminated while the other actors on stage recede.

From the physiological perspective, there is evidence for two separate systems governing the focusing of awareness and suppression of those stimuli not selected, defined as the executive control aspect of attention by Petersen and Posner (2012). This theory uses findings from Dosenbach, Fair, Cohen, Schaggar, and Petersen (2008) that show a dual network where there are two distinct physiological systems in controlling awareness. The cingulo-opercular control system provides stable maintenance of stimuli across trials. The fronto-parietal system initiates and adjusts control, and may be related to task switching and adjustments within trials. The two systems are independent according to functional magnetic resonance imaging (FMRI) studies, and yet still operate in tandem during a given task.

There is some research that disagrees with Petersen and Posner's (2012) attentional theory. Silk, Vance, Rinehart, Bradshaw, and Cunnington (2008) found there to be significantly less activation of the fronto-parietal network in children with attention deficit hyperactive disorder (ADHD) while engaged with the Raven's Progressive Matrices task, a task that places a high demand on the prefrontal cortex. Despite this difference, the ADHD and control groups did not differ significantly on performance during the task. Silk et al. did find overall more fronto-striatal activation in the normal controls as compared to the ADHD group. It is possible Silk et al. did not account for the cingulo-opercular control system because this network was identified by Dosenbach et al. (2008) just after publication of the Silk et al. study. An alternative reason for the difference in this finding is the idea of neuroplasticity: if a particular region of the brain is deficient for some reason, other regions may try to compensate. Consequently, the biological observations of Petersen and Posner may be fundamentally flawed in that the brain is able to accomplish the same processes with different regions. The theoretical processes may be the same, but the reliance on observations of activity in the brain still not fully understood can be misleading (Malabou & Rand, 2008).

This dual process theory of executive attention was used as an explanation for the development of self-regulation—that is, our ability to control our thoughts, our feelings, and our behavior. Rothbart, Sheese, and Posner (2007) describe the anatomical development of executive function and how the development of these processes reflects a child's developing ability for effortful control. While there are effortful control differences between children at 24 months, 36 months, and children between 4 years and 7 years old, results are surprisingly consistent from 7 years old to adulthood. Gardner, Dishion, and Posner (2006) found that higher ratings of inattention (lower self-regulation) in children were correlated with early tobacco use, and that the higher use of tobacco was associated with enhanced executive attention for those early tobacco users. The latter finding is probably due to withdrawal effects and may contribute to nicotine dependence to avoid any loss of attention during withdrawal.

This biological approach to understanding executive function is very attractive in that it is an approach to understanding human behavior through observation of physical events. Instead of inductively defining the components of executive function and choosing the tests to target those components, this approach observes processes in the brain during specific periods of time and attempts to identify the processes that regulate behavior and how those processes interact. This approach is limited however by the current level of technology. As measurement precision increases and new methods of observing mental events evolve, there will be changes to the understanding of those processes currently cited as executive functions. Additionally, as with the factorial approach, this behavioral approach removes any conscious element that may be going on but not fully understood.

Barkley (2011): The Clinical Approach. Barkley (1997) initially approached executive function from a clinical perspective by proposing a new theory of ADHD. As Barkley explored executive function in more detail, he began to uncover the many definitions and approaches this review explores. In the 2012 book, *Executive*

Functions, Barkley details four problems in current executive function research (p.1 – 35):

1. "Executive function" lacks a concise definition that distinguishes which functions are executive and which are not.
2. Current theories have moved away from capturing the difficulties of people with prefrontal cortex damage, especially problems with emotional, social, economic, and moral domains.
3. Current theories are "cold cognitive components" that lack the interaction of social networks and an identification of the self.
4. Executive functions have been studied without concern as to why executive functions exist.

Taking these problems of executive function research into account, Barkley (2012) states it is necessary to develop a concise theory of executive function apart from the psychometrics. Once a working theory of executive function is established, measurement and assessment can be naturally developed and explored. Barkley defines executive function as "*those self-directed actions needed to choose goals and create, enact, and sustain actions towards those goals*, or more simply as *self-regulation to achieve goals:* EF = SR" (emphasis in original text, p. 60). In his book, Barkley goes on to address each of the problems in current executive function research and how his theory addresses them; these particulars are not necessarily important for the research question presented in this review.

Like Miyake and Friedman (2012), Barkley (2012) suggested inhibition is an essential component to other executive functions, though Barkley holds a sense of self is required before inhibition can operate. His theory identifies six executive functions that are involved in self-regulation: self-awareness, self-restraint (inhibition), self-directed sensory-motor action (nonverbal working memory; imagination), self-speech (verbal working memory), self-directed appraisal (emotion-motivation), and self-play (innovation, problem solving). This theory is, according to Barkley, a hybrid approach based on the functions of the prefrontal cortex proposed by Bronowski, Fruster, Goldman-Rakic, and Damasio (as cited by Barkley, 2011, p. 555).

This hybrid theory puts self-awareness as the primary factor other executive functions rely upon to be able to function, because without self-awareness goals cannot be created. Inhibition follows closely after self-awareness in importance as it provides protection from interference to allow the other four executive function processes to facilitate goal-directed problem solving. This response inhibition refers to three overlapping processes: 1) inhibiting a prepotent, or dominant, response to an event to create a delay in responding, 2) interrupting an ineffective response, allowing for reevaluation of the situation, and 3) providing resistance to distraction. These processes are necessary for self-regulation and give an individual the ability to change their behavior to continue toward a desired goal.

The other four components of Barkley’s (2012) executive function theory are able to function because of inhibition. Self-directed sensory-motor action encapsulates nonverbal working memory, otherwise known as the visuospatial sketchpad (Baddeley, 1986). This process is both retrospective (reviewing or reliving past encounters) and prospective (imagining a future action). Self-speech is verbal working memory, defined by Baddeley (1986) as the phonological loop. This executive function process allows an individual the ability to self-assess, self-instruct, and problem solve, and apply self-control through language. Self-directed appraisal develops as somatic markers as a consequence of the preceding executive function processes (self-awareness, inhibition, self-sensing, and self-speech). Once a child can relive incidents, they can relive the emotional aspects of that event. This ability to have personal emotional events, separate from the event they have originated from, becomes a source of intrinsic motivation that is necessary for goal-directed behavior. The last process Barkley identifies in his executive function theory, self-play, involves the ability for an individual to engage in reconstitution of prior behaviors. Barkley describes this process happening in two stages: the analysis stage, where old behaviors are broken down into prime components, and the synthesis stage, where these prime components are assembled in ways to address a particular situation or obstacle. This executive function process is described in neuropsychology with the terms fluency, flexibility, and generativity (Barkley, 2011, p. 557).

These executive function processes are a set of tools used by individuals that facilitates adaptive functioning with inhibition being the foundation on which all other processes operate. Barkley is known as an expert on ADHD, and because he is approaching executive function from observations in clinical practice, it is not at all surprising his theory has been used in research and treatment of ADHD. This clinical approach to executive function has its limits. There is a potential when outlining a theory on general human behavior, especially after being immersed in a disordered population, to identify processes by being absent as opposed to being present. Gordon Allport (1955, p. 55) once said “A large share of our troubles lies in the fact that the elements we employ in our analyses are not true parts of the original whole." A skeptic of using the clinical approach to theory building, Allport was determined to understand the individual from a healthy perspective. From this point of view, the disorder does not dominate the theory or color it in such a way that may obfuscate the healthy processes that should be included in the theory. Barkley's background may interject some aspects that are symptomatic in an unhealthy population but do not pertain to healthy executive functioning. On the other hand, it is possible the deficits Barkley observed help explain executive function processes. This empirical study attempts to clarify this issue.

Model differentiation and commonality. These three approaches to the conceptualization of executive function are by their nature different. Each take significantly altering perspectives on how the components of executive function relate to one another. Miyake and Friedman (2012) posit theoretical processes of executive function that are separate and yet each jointly contributes to some ability on a task. However, there is concern that the Miyake and Friedman factor structure is not the best model to support the assessment of executive function (Faneros, 2011). The Petersen and Posner (2012) theory produced an interconnected model where "executive control" is a result of two separate functioning processes. The "much broader term executive function" (p. 83) is a result of the overall attentional system. Barkley's theory of executive attention is heavily integrated and all parts contribute to behavior. All of these theories address inhibition as an important part of the model. Most notably, the Miyake and Friedman and the Barkley theories emphasize the role of inhibition in their models as a foundation upon which other executive processes operate, even though

Miyake and Friedman include inhibition as “general executive function” in the model. Petersen and Posner also address inhibition and its relation to executive function, though they do not use the same language and point to the development of biological executive function processes as giving rise to inhibition.

Miyake and Friedman (2012) define inhibition as an individual’s ability to deliberately inhibit dominant, automatic responses when necessary, but operationally define inhibition as part of a greater executive function factor that influences the other executive functions, updating and shifting. While operationally defining their model in this way allows for a more parsimonious model with better psychometric model fit indices, Miyake and Friedman sacrifice predictive power by confounding the inhibition dimension with a significant portion of the model variance. This is one pitfall of using factor analysis to guide theory rather than theory guiding conceptualization of the variables needed to define a dimension. From a practical standpoint, moving inhibition into a causal role but including all other model variance sacrifices accurate prediction of its affect on the individual differences in overall performance.

Barkley’s (2011) theory assigns inhibition as the critical element of his model of executive function. Without inhibition, or at least some inhibition, the other executive functions Barkley includes in his model cannot function. As in Miyake and Friedman’s (2012) model, inhibition in Barkley’s model is not a dichotomous variable, and Barkley would likely agree everyone has some ability to inhibit actions. Unlike Miyake and Friedman’s theory, Barkley’s theory is not based on factorial evidence for inhibition as a separate executive function, nor does he cite model characteristics that suggest the necessity of inhibition for the other functions to work. Rather his theory of executive function was based on his own experience and was derived from his broader theory of ADHD. That said, Barkley's theory does not have much research support yet. Therefore the theory could be applied in a heuristic sense, while not yet necessarily having empirical evidence.

Petersen and Posner examine inhibition from a biologically-based approach. This theory contradicts the Miyake and Friedman and the Barkley theories by placing effortful control and self-regulation as products of executive function, rather than inhibition allowing other executive function processes to work. This perspective places goal-setting behavior and self-regulation as a result of mechanistic processes of attentional direction and management of information. This theory lacks an explanation for why exactly we attend to the stimuli on which we focus our attention, and therefore does not accurately account for goal-setting behaviors.

It is important to note that the Petersen and Posner (2012) biological theory of executive function and attention does not exactly match the biological explanations given by Miyake and Friedman (2012). Instead of the two attentional executive function networks, Miyake and Friedman adopt a prefrontal-cortex basal-ganglia working memory (PBWM) mechanism proposed by O'Reilly and Frank (2006). This proposal suggests that the prefrontal-cortex operates as the mechanism for stable maintenance of goals. The theory further proposes the basal-ganglia gates information to the prefrontal-cortex. This proposed model, and Miyake and Friedman's confirmatory analyses, have not yet been directly supported with imaging technology, but rather have been experimentally confirmed using computational modeling.

However, Faneros (2011) showed the factorial structure used by Miyake and Friedman may not be as robust and may even have a slightly different factorial structure.

Research Question

Detailed above are three theories of executive function with differing theoretical foundations and yet similar components. Could all three predict performance equally on a complex mental task involving executive function, or does one suggest a model that predicts performance better than another? The goal of this research investigation is to answer that question. The Faneros (2011) study attempted to understand the separate but unifying nature of executive function (Miyake et al., 2000) by testing the Miyake and Friedman model using the tests they cite, in addition to tasks from the ANAM battery. Using similar protocols as Faneros (2011), each model described above will be fit to a structural equation model (SEM) and the model fit for each will be evaluated and compared. The additional model developed by Faneros (2011) in the post-hoc exploratory factor analysis will be included in the comparative evaluation. There is a possibility that each model is robust at measuring one particular aspect of executive function. This will also be examined through model comparisons.

In order to test these theories, a model for each theory has to be constructed. Miyake and Friedman's (2012) theory has already provided a model, because the model construction was part of their theory building process (see Figure 1). Additionally, tasks that target the theoretical dimensions in the Miyake model are already established through the factor analysis process. The post-hoc model of executive function shown in Figure 2 and the corresponding tasks from Faneros (2011) are similarly already established.

Petersen and Posner (2012) and Barkley (2011) did not formalize a model, so a model has been constructed from their writings. Each presents a description of the necessary model components and how each component interacts with the others. A model for the Petersen and Posner (2012) and Barkley (2012) models was developed according to the model descriptions and can be seen in Figures 3 and 4, respectively. Note that self-awareness is not present in Barkley's model. Self-awareness, according to Barkley, is a dimension that is needed for other executive functions to exist but does not necessarily contribute to effectiveness of other processes. Self-awareness does not guarantee normal executive function (Barkley, personal communication, August 2, 2013). While Peterson and Posner have not personally endorsed the model developed from their 2012 article (Figure 3), Barkley has approved the proposed model (Figure 4) through a personal communication.

The model developed to test Petersen and Posner's theory comes directly from their 2012 model description. Here the first process to initiate attention is the alerting system, which is also the system that regulates sustained attention. Once an object arouses the system to pay attention to it, the orienting mechanisms focus the system onto the stimulus. These mechanisms are also responsible for the shifting of attention, both overt and covert, to follow the stimulus. Once attention has been shifted to the stimulus, executive control allows the mental system to monitor items within a task both across trials and within trials. Tasks that have traditionally been used to test for these dimensions will be used in this investigation. A visual vigilance task can be used for the alerting system, and Eriksen and Eriksen's (1974) Flanker task has been used for the orienting mechanisms in past research (Posner, 2012). Tasks that test the ability to monitor across trials, such as the 2-Back task, and tasks that test switching task sets, such as a letter-number task, can be used to test the respective executive control processes.

Developing a model for Barkley's (2012) theory is more complex and the defining dimensions are difficult to distill into a specific performance task. Indeed, Barkley admonishes any researcher trying to distil executive functions into a specific test battery (p. 194), even though he does admit test batteries "*may* have some role to play" in assessing inhibition, nonverbal and verbal working memory, planning, problem solving, and creativity/fluency/generativity. Barkley and Murphy (2010) assessed a group of ADHD adults and controls with a self-report measure (Deficits in Executive Function Scale, or DEFS, Barkley, 2011) and a test battery of executive function measures. Of the tests given, the Connors Continuous Performance Test (CPT) and the 5-Point Test of Design Fluency were the only ones that contributed to executive function impairment in an occupational setting. This may be due to the hierarchical nature of executive function (Barkley & Murphy, 2010, p. 169). Accordingly, CPT will be used to assess inhibition, the 5-Point Test (5PT; Tucha, Aschenbrenner, Koerts, & Lange, 2012) will be used to assess nonverbal memory, the automated version of Operation Span (Unsworth, Heitz, Schrock, & Engle, 2005) will be used to assess verbal memory, and Tower of Hanoi will be used to assess planning (self-play). Self-appraisal is a construct that has no performance-based method of testing like other executive functions. Barkley (2012) refers to self-appraisal as ultimately being the component which "provides the source of internal fuel – self-motivation" (p. 90). Given this, and understanding the dimension is perhaps much more encompassing, a measure of intrinsic motivation was given to measure self-appraisal. The Intrinsic Motivation Inventory (IMI, Ryan, Mims, & Koestner, 1983) is flexible enough to satisfy the requirements for this study.

Assessing executive function across these theories will require comparing performance on a single complex task that involves considerable executive functioning—in this case the Multi-Attribute Task Battery II (MATB-II). The MAT-II was developed by NASA to assess performance in a multitasking environment. The task battery presents a task environment and workload similar to that of piloting an aircraft, but simplified and easy to learn. This multi-task is very complex and yet training participants is fairly straightforward and quick. The model defined by each theory can be compared to performance on the MATB-II to determine which theory best predicts complex executive functioning behavior.

Method

Participants

A total of 322 participants were recruited from the University of Oklahoma campus through a freshman-level Psychology class; however, 16 participant's data were removed due to missing data. Of the remaining 306 participants, 207 (67.65%) were women and 99 were men with a mean age of 18.67 years and age range of 18 to 26 years. Thirty-one participants identified themselves as being left-handed, however for all computerized tasks they preferred to use the right-hand mouse-and-keyboard setup.

Materials

Assessment measures were drawn from prior experimentation from each of the theoretical paradigms. To test the Miyake and Friedman (2012) model, tasks Keep

Track, CPT 2-Back, Number-Letter, Category Switch, Stroop, and Stop-Signal were given. The Wisconsin Card Sort Task, CPT 2-Back, Automated Operation Span, Tower of Hanoi, Number-Letter, and Stop Signal tasks was used to test the Faneros (2011) model. A Visual Vigilance task, the Flanker task, CPT 2-Back, and Letter-Number was used to test the Petersen and Posner (2012) model. A Standard CPT task, Stop Signal, 5 Point Test of Design Fluency, Automated Operation Span, Intrinsic Motivation Inventory, and Tower of Hanoi was used to test the Barkley model. Lastly, the MATB-II was administered to assess each model's ability to predict complex executive function performance. A detailed description of each task follows. All tasks were given in a single two-hour session while seated at a computer station.

Keep Track. This task replicated the Keep Track task described in Friedman et al. (2008) and is a task that assesses the updating domain of executive function. Participants in this task were asked to keep track of the word that corresponded to one, two, or three target categories. At the beginning of the task participants were shown six possible categories (animals, colors, countries, distances, metals, and relatives) and the three exemplars for each category used in the task to ensure the participant would know how to categorize the exemplars. At the start of each trial the categories would appear at the bottom of the screen. Exemplars were then randomly presented on the screen for 1,500 ms each with the categories remaining on the bottom of the screen. After fifteen words were shown, the participant was asked to report the last word for each category by entering the answer using the computer's keyboard. One practice trial was given followed by ten recorded trials. Keep track was administered through OpenSesame (version 2.7.4; see Mathôt, Schreij, & Theeuwes, 2012). The dependent measure for Keep Track was percent correct.

CPT 2-Back (Two-Back). Two-Back is a task that assesses the updating domain of executive function. A participant was shown a series of numbers in the center of a computer screen. Each number remained on the screen for 500 ms, had a random interstimulus gap between 750 ms and 1500 ms, and the participant was given a maximum of 1000 ms to respond. The participant was instructed to keep track of the current and last two numbers shown. For each stimulus the participant was to respond by clicking the left button on the computer's mouse if the current stimulus was the same stimulus as the one before the prior stimulus (two back). Otherwise they were to click the right mouse button to indicate they were different. This task is very similar to the "spatial 2-back" task described in Friedman et al. (2008). Two-Back was administered through ANAM (version 4). The dependent measure for two-Back was percent correct.

Number-Letter. This task, adapted from Rogers & Monsell (1995), was designed to assess a participant's ability to switch between two task-sets needed to make judgments. Participants were presented with a number-letter pair and asked to determine if either the number was even or odd, or if the letter was a vowel or consonant, depending on a stimulus directly above the number-letter pair that directed the participant's attention to either the number or letter. Participants first saw 10 practice trials and then were given 96 trials. The dependent measure in this task was the switch cost, calculated by the average reaction time of trials that required a switch minus the average reaction time of the trials that did not require a switch. Number-

Letter was administered through ANAM (version 4) using a special stimulus file to adapt the ANAM task Logical Switch to Number-Letter.

Category Switch. This set-switching task is a replication of the Category Switch task detailed in Friedman et al. (2008). In each trial of this task participants were presented with a word that could be classified as (a) living or non-living or (b) bigger or smaller than a soccer ball. Sixteen words were presented randomly across 96 trials and were the same words used by Friedman et al. and Mayr and Kliegl (2005), from which this task was adapted: *table, bicycle, coat, cloud, pebble, knob, marble, snowflake, shark, lion, oak, alligator, mushroom, sparrow, goldfish,* and *lizard.* A symbol above each word was the cue to which categorization to use. A heart indicated the participant was to answer if the word indicated living or non-living. An arrow cross indicated larger or smaller. The dependent measure for Category Switch was switch cost, calculated by the average reaction time of trials that required a switch minus the average reaction time of the trials that that did not require a switch. This task was administered through OpenSesame (version 2.7.4).

Stroop. This task assesses a participant's processing speed and ability to inhibit prepotent responses. Participants were given three blocks of trials. In the first block the participant was asked to respond to color words on the screen by pressing an associated button (1 for RED, 2 for GREEN, and 3 for BLUE). In the second block, different colored stimuli were shown serially ("XXXX" in different colors) and the participant was instructed to press a button associated with the color. The last block instructed participants to read color words on the screen while pressing the button that is associated with the word's color (e.g. the word "BLUE" may appear on the screen in red). The dependent measure used in this task was interference cost, calculated by the differences in reaction time to block three trials and block two trials. This task was administered in ANAM (version 4).

Stop-Signal (Go-No-Go). Go-No-Go measures a participant's response inhibition. This task presented the participant with one of two stimuli: an upper case "X" or an upper case "O". The participant was asked to click the mouse button as fast as possible when seeing the "X", but not to respond to the "O". The participant was given 106 response trials interspersed with 24 non-response trials. The dependent measure for Go-No-Go was *d*-prime, an index of sensitivity based on signal detection theory (Macmillan & Creelman, 2005).

Berg's Card Sort Task (BCST). The BSCT is an implementation of the Wisconsin Card Sorting Test (Berg, 1948), a common clinical neuropsychological test sensitive to frontal lobe dysfunction. This task tested the participant's ability to update information in working memory and to shift working sets as new information is received. This task asked a participant to sort cards that have differing colors and symbols on them into piles according to a rule (e.g., according to color). The participant was not told the sorting rule but was required to figure out the rule through trial and error based on the feedback given when they sort each card. After each sort the participant was given feedback regarding if the card was sorted correctly according to the current rule. There was no time limit on sorting cards. After ten cards were sorted correctly, the rule was changed without informing the participant. This computerized task was administered with the Psychological Experiment Building Language (PEBL, Mueller, 2012). Fox, Mueller, Gray, Raber, and Piper (2013) demonstrated a high correlation between the typical 128-trial task and a short form 64-trial task. The participants in this study were given the short form. Percent of total perseverative responses was the dependent measure for the BCST.

Automated Operation Span (AOSpan). This task asked the participant to memorize a string of letters while at the same time answer a series of simple math problems (adding and subtracting). At the end of each memorization block the participant was asked to reproduce the letters in the order they were presented. This assesses the participant's ability to update working memory as well as test the participant's amount of available working memory. The version of Automated OSPAN in this study followed the same procedure used in Unsworth, Heitz, Schrock, and Engle (2005) and was administered in the Inquisit Lab (Version 4.0.3) software package.

Tower Puzzle (Tower). This task asked the participant to move blocks on three spindles to reproduce a target set. The participant was only able to move the blocks from spindle to spindle according to certain rules, and had a limited amount of time to place them. Additionally, the participant was asked to try and use as few moves as possible. This test assesses the participant's ability to inhibit moves from one spindle to another without some plan, and tests overall test planning. The task was administered through ANAM. One practice trial was given to insure the participant understood how to solve the task. Five trials were given and the mean score was used as the dependent measure.

Visual Vigilance. This task assesses a participant's ability to stay alert and respond to sparse-appearing stimuli. Participants were presented a blue field on a computer monitor and were asked to click the mouse button as soon as they detected when a small square appeared within the field (measuring .9 cm square). These stimuli appeared on the screen an average of once every 10 seconds with an inter-stimulus gap between 5 and 15 seconds. Participants were given 30 trials for a total of five minutes of testing. This was not a typical vigilance task that would take twenty to thirty minutes. However, this shorter task has been shown to be a viable alternative (Loh, Lamond, Dorrian, Roach & Dawson, 2004). Mean reaction time was used as this task's dependent measure. Visual Vigilance was administered through ANAM.

Flanker. This task was adapted from Eriksen and Eriksen's 1974 study to measure visual interference. Participants were presented an arrow on screen that pointed either to the left or the right and were asked to either press the left "Shift" key on the keyboard for arrows pointing left or the right "Shift" key for arrows pointing to the right. Participants were given twelve practice trials to ensure they were using the correct keys and then given 120 trials. A third of the trials (40) presented the stimulus on the screen by itself (defined as neutral trials). Forty of the trials presented two arrows on either side of the stimulus pointing the same direction (defined as coherent trials), while another 40 trials presented arrows in the same position pointing the opposite direction as the stimulus (defined as non-coherent trials). The three conditions were given randomly. The dependent measure for Flanker was the conflict reaction time, calculated by the difference of the mean reaction times of the non-coherence trials and the mean reaction times of the neutral trials. Flanker was administered through OpenSesame (version 2.4.0).

Standard CPT. The standard continuous performance task, or Standard CPT, assesses a participant's ability to attend to a task and inhibit prepotent responses. The participant was asked to memorize a character, in this case an "X." The task then displayed a sequence of characters on the screen, approximately one per second, with ten percent being the target stimulus. When the participant was directed to click the mouse as fast as possible when the target stimulus appeared, but let all other stimuli pass without responding. Similar to Go-No-Go, Standard CPT used the *d*-prime statistic as the dependent measure (Macmillan & Creelman, 2005) and was administered through ANAM.

5-Point Test of Design Fluency (5PT). The 5PT was originally developed by Regard, Strauss, and Knapp (1982) as a nonverbal working memory task. The task presents a participant with a sheet of paper with 40 five-dot matrices (Lee et al., 1997). Participants were asked to make as many designs as possible within a 3-minute time limit without repeating any designs. Should a participant complete 40 designs within the time limit they were given another sheet of figures to continue working. The dependent measure for this task was the number of original designs. This task was given via pencil and paper.

Intrinsic Motivation Inventory (IMI). The IMI is a multidimensional questionnaire intended to assess a participant's subjective experience related to an experiment. The four subscales on the standard 22-item IMI include interest/enjoyment, perceived competence, felt pressure and tension, and perceived choice. The interest/enjoyment subscale has been used in many studies as a self-report measure of intrinsic motivation (e.g., Deci, Eghrari, Patrick, & Leone, 1994; Plant & Ryan, 1985; Ryan, 1982; Ryan, Connell, & Plant, 1990; Ryan, Koestner & Deci, 1991;

Ryan, Mims & Koestner, 1983). Total score on the interest subscale was used as the dependent measure. This task was given via pencil and paper.

Multi-Attribute Task Battery (MATB-II). This task was developed to test multiple task management performance and requires participants to attend to four different tasks: system monitoring, tracking, communications, and resource management. See Figure 5 for an example of the MATB-II testing screen. The system monitoring task required a participant to "turn on" systems when they randomly "go off" and to keep system dial indicators within a shaded section of four bar graphs. The two-dimensional tracking task required a participant to keep a target in the center of a box by using a joystick. The communications task required a participant to monitor and respond to an audio prompt by changing a designated number to a number intermittently generated and presented by the software program. Finally, the resource management task required a participant to maintain a specific amount of "fuel" in four designated tanks within a complex set of tanks and pumps. During the test session, fuel level in the tanks became unbalanced due to pump failures and required the participant to re-balance fuel levels by turning on and off "pumps" that were connected to those tanks and a supply tank. MATB-II has 21 available measures for analyses, however many of these have been shown to be affected by range restriction (e.g., Caldwell & Ramspott, 1998; Singh, Tiwari, & Singh, 2010). Percent of correct responses was the dependent measure for the system monitoring and communications task. Mean tank level differences (actual level – target level) was the dependent variable for the resource management task. The root mean square of errors was the dependent measure for the tracking task. Additionally, a composite measure was defined by computing a z-score for each dependent measure and averaging them, thus giving equal weight for each task in a single composite measure. This composite measure was used in the predictive efficiency assessment of each model.

All analyses were carried out using R (version 3.0.3). The R package “psych” (version 1.4.3, Revelle, 2014) was used for all descriptive statistics. The R package “lavaan” (version 0.5.16) was used for all modeling analyses.

Procedure

All consent procedures and testing were conducted in a windowless room that provided control of lighting and other environmental factors. The room contained eight computer stations each with a folder with questionnaire materials, a pencil, a keyboard, a mouse with mouse pad, a joystick, headphones, and a computer monitor. Computer stations were separated by dividers. At time of testing, participants were invited into the room and asked to sit at a computer station. An ID number was placed on the station’s folder and was used for all tests to track participant data. The ID number was composed of the Julian day (e.g., November 1st = 307), testing session number for that day, computer station, and experimental condition. Participants were not assigned stations prior to testing. After all participants were seated, the participants were directed to the consent form in the folder and asked to follow along as the form was reviewed aloud by the research administrator. Once participants had any questions answered and consented to the study, they were directed to complete a demographics questionnaire. If the participant noted on their demographics questionnaire they were left-handed, they were given the option to use the mouse and keyboard in a left-handed fashion and tasks were appropriately modified. However, all left-handed participants indicated they preferred to use the mouse and keyboard in a right-handed fashion.

Once all participants completed the demographics questionnaire, participants were given a battery of tasks. Two versions of this task battery were given depending on the testing session condition. Condition one tasks were given in the following order: Go-No-Go, Tower, Two-Back, Standard CPT, Stroop, Visual Vigilance, Number-Letter, BCST, Flanker, Category Switch, Keep Track, and AOSpan. Condition two tasks were given in the following order: BCST, Flanker, Category Switch, Keep Track, AOSpan, Go-No-Go, Tower, Two-Back, Standard CPT, Stroop, Visual Vigilance, and Number-Letter. All participants were then given the opportunity to rest for up to five minutes, and then given the 5PT, the IMI, and MATB-II. After the MATB-II participants were thanked for participating, invited to ask questions, and offered the opportunity for their data not to be used for this study. No participant exercised this option. Testing took approximately two hours.

Results

An independent *t*-test for all dependent variables was used to assess test presentation effects. There were no significant differences between task presentation order for any of the dependent variables. ANOVA was used to test for time-of-day effects. There were no significant differences between testing times on any of the dependent variables. Therefore, presentation condition and time of day was not included in further analyses.

Table 1 details the fifteen dependent variables and their means, standard deviations, range, skew, and kurtosis. The distributions were close to a normal distribution with the exception of Visual Vigilance (skew = 1.99) and Keep Track (skew = -1.15). The distributions of Visual Vigilance (kurtosis = 5.58) and Keep Track (kurtosis = 1.27) were also leptokurtic. Table 1 lists the means, standard deviations, range, skew, and kurtosis for the MATB-II variables. Average performance on the MATB-II was close to a normal distribution though leptokurtic.

Figures 6 through 9 show correlations and distributions for each task included for each of the models. There were only four correlations that would be considered to have a low to moderate linear relationship. Two Back performance correlates with Tower, $r = .32$, $p < .01$, and FivePT, $r = .26$, $p < .01$. Go-No-Go correlates with Flanker, $r = .37$, $p < .01$, and Standard CPT ,$r = .29$, $p < .01$. Correlations of $r = .11$ and higher were significant to the $p = .05$ level, however this was likely due to the large sample size. The large sample enhanced statistical power, allowing small effects to be detected. Nevertheless, most correlations predicted by the models were significant. Overall performance on MATB-II was significantly correlated with Keep Track, $r = .13$, $p = .03$, but was not significantly correlated with any other measure.

Model fit was determined by confirmatory factor analysis using the lavaan package in R. The Miyake and Friedman (2012) model was defined with the following code:

```
model.MF2012 <- '
     #latent variable definitions
     F1 =~  KeepTrack + TwoBack
     F2 =~  NumLet + CatSwitch
     F3 =~  KeepTrack + TwoBack + NumLet + CatSwitch
            + Stroop + GoNoGo
     '
fit.MF2012 <- cfa(model.MF2012, data = selTests.MF,
     orthogonal=TRUE)
```

This model failed to converge after 10,000 iterations. Because of this, the model Miyake et al. (2000) presented, upon which the Miyake and Friedman (2012) model was based, was also fitted. The above code was used except for a redefinition of F3 (“`F3 =~ Stroop + GoNoGo`”) and the fit criteria did not specify orthogonal latent variables. The result was a model that had a fair fit, χ^2 (6, N = 306) = 12.643, p =.049, CFI = 0.924, SRMR = 0.051. Figure 10 reflects the resulting factor and variance loadings.

The following code was used to define the Faneros (2011) model:

```
model.F <- '
     #latent variable definitions
     F1 =~  CardSrt + TwoBack + AOSpan + Tower
     F2 =~  NumLet
     F3 =~  GoNoGo

     #variance definitions
     NumLet ~~ 0.22
     GoNoGo ~~ 0.24
     '
fit.F <- cfa(model.F, data = selTests.F)
```

Models that have a single indicator often have identification problems (Loehlin, 2004). Therefore, Number-Letter and Go-no-go were assigned fixed variances as 1 - reliability. Reliability for Number-Letter was computed using Cronbach’s α; reliability for Go-No-Go was computed by adjusting the split-half correlation with the Spearman-Brown prophecy formula. The model converged after 63 iterations, and the model had good fit characteristics, χ^2 (8, N = 306) = 6.802, p =.56, CFI = 1.000, SRMR = 0.030. Figure 11 reflects the resulting factor and variance loadings.

The following code was used to define the Peterson and Posner (2012) model:

```
model.PP <- '
     #regressions
```

```
		F3 ~ F2
		F4 ~ F2
		F2 ~ F1

		#latent variable definitions
		F1 =~ VisVig
		F2 =~ Flanker
		F3 =~ TwoBack
		F4 =~ LetNum

		#variance definitions
		Flanker ~~ 0.32
		'
	fit.PP <- cfa(model.PP, data = selTests.PP,
		std.lv=TRUE, orthogonal=TRUE)
```

The model converged after 43 iterations and had very good fit statistics, χ^2 (10, N = 306) = 7.047, p =.72, CFI = 1.000, SRMR = 0.024. Figure 12 reflects the resulting factor and variance loadings.

The following code was used to define the Barkley (2012) model:

```
	model.B <- '
		#regressions
		#F2 ~ F1
		#F3 ~ F1
		#F4 ~ F1 + F2 + F3
		#F5 ~ F1 + F2 + F3

		#latent variable definitions
		F1 =~ StdCPT + GoNoGo
		F2 =~ FivePT
		F3 =~ AOSpan
		F4 =~ IMI.Int
		F5 =~ Tower

	#variance definitions
		FivePT ~~ 0.37
		AOSpan ~~ 0.22
		IMI.Int ~~ 0.22
		Tower ~~ 0.23
		'
	fit.B <- cfa(model.B, data = selTests.B, std.lv=TRUE,
		orthogonal=TRUE)
```

Reliability for 5PT is the test-retest reliability as reported by Tucha et al. (2012). Reliability for AOSpan, IMI, and Tower were assessed using Cronbach's alpha. The model converged aft er 53 iterations, and had good fit statistics, χ^2 (9, N = 306) = 6.82, p = .656, CFI = 0.989, SMRM = 0.020. Figure 13 reflects the resulting factor and variance loadings.

To assess each model's ability to predict performance on a complex executive function task, a series of CFA models were tested to assess the stability of each model when introducing a new variable. Each model introduced the new variable according to the underlying theory. The factor loadings, interfactor correlations, and regression weights were allowed to vary instead of fixing them. This allowed comparisons from the base model upon which the new variable was applied. If the model's weights caused severe changes to the factor structure, then there is a possibility the underlying model had been misspecified. For the Miyake et al. (2000) and Faneros (2011) models, seven models each were tested with a varying number of paths from the latent variables to the complex executive function task observations. The Peterson and Posner (2012) and Barkley (2012) models each dictated two paths for observable behavior, and therefore three models for each were tested.

The Miyake et al. (2000) model and the Barkley (2012) failed to give any useful model fits. Each of the seven models for Miyake et al. and the three for Barkley either did not converge, yielded covariance matrices that were not positive definite (negative eigenvalues), or the model fit statistics were very poor. Three of the Faneros (2011) models also failed to give useful results, however four of the models fit (see Table 4). While these three models fit, none of the factor loadings were significant according to a

Wald test and the covariance between latent factors changed significantly. Other factor loadings remained relatively the same except for those unitary measures.

The three Peterson and Posner (2012) models had good fit statistics (see Table 5). However, the path from the circulo-opercular latent factor was not significant, and therefore the favored model was the single path from the fronto-parietal. Adding the MATB-II factor affected the overall loadings of other factors, the most significant being Standard CPT, which shifted from .96 to .46 with an equivalent increase for Visual Vigilance. The regressions also changed slightly, however the significance for each regression remained the same.

Discussion

This study was designed to investigate the theoretical foundation of executive function by comparing four contemporary models. This was accomplished by a four-step process. First, all models were tested for good model fit. Second, the parsimony for the four models as assessed by the Comparative Fit Index and the Akaike Information Criteria were compared. Third, a measure of complex executive function (MATB-II) was added to the model to assess model stability and the model's ability to predict performance. The fourth and last step in comparing these models was to compare the parsimony of those models that showed stability once MATB-II was added to the model.

The correlation matrices for each model showed weak to moderate correlations as typically found in studies of individual differences in executive function (Miyake et al., 2000). While many of the correlations were statically significant, there were four that would be considered to have a low to moderate linear relationship. Two Back performance correlates with Tower and FivePT, and Go-No-Go correlates with Flanker and Standard CPT. Most of the correlations predicted by the models were significant, even if the effects were very small. More importantly, analysis of the models of executive function yielded clear support for one model over the other three.

The Miyake and Friedman (2012) model failed to converge. The Miyake et al. (2000) model, run post-hoc, converged but the fit statistics had poor model fit indices. These results were similar to the results found in the 2011 study by Faneros. The Faneros model showed similar factor loadings as the exploratory factor analysis model determined in the 2011 study and has good model fit indices. However, Card Sort in this study was not significantly correlated with any other measure given, and its factor loading was very different from the model determined in the 2011 study. This could indicate a fourth latent factor.

The Peterson and Posner (2012) model had very good fit indices. All factor loadings were significant as determined by a Wald test; however none of the regression factors were significant. This indicates that the latent factor relationships were not as linear as Peterson and Posner suggest, but should be specified another way. It is possible the orienting mechanisms were not necessary for the cingulo-opercular and fronto-parietal systems to operate, but rather all three lower latent factors might depend on the alerting system. The significant correlations between Standard CPT and tasks that define the other latent factors reflect this potential.

The Barkley (2012) model showed very good fit indices. All factor loadings were significant; however this could be because of the single measures used to define four of the latent variables. Two of eight regressions were significant. This too could be because of the unitary nature of the four latent variables. Adding additional measures to capture each variable could lead to a better understanding between the latent variables.

To compare these models, a comparison of the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC) was made between models. These related fit indices can be used to compare non-nested models given the same data set (Busemeyer & Diederich, 2010) by taking the lowest AIC and BIC calculated value. Table 2 lists the AIC and BIC values associated with each model. Of the four fitted models, the Miyake et al. (2000) model appears to be the most parsimonious followed closely by the Peterson and Posner (2012) model. Given the poor fit of the Miyake et al. model as indicated by the χ^2 statistic, the better model trading goodness of fit for parsimony was the Peterson and Posner model. The Barkley (2012) model has the highest AIC and BIC values due to the complexity of the model.

Adding MATB-II to the models cause all the models some systemic problems. The Miyake et al. (2000) model and the complex Barkley (2012) model failed to maintain a stable structure and failed to converge or had covariance matrices that were not positive definite. The Faneros (2011) model had good fit; however, all the paths analyzed were not significant according to a Wald test. The Peterson and Posner (2012) model proved to have good fit and significant factor loadings for the frontal-parietal latent variable path, but adding the variable causes instability in the rest of the model. These results indicate that, as predicted by the model comparisons above, the Peterson and Posner model was more stable and a better fit than the other models. Nevertheless, this model does have some weaknesses in being able to integrate all parts effectively. To conclude, all the models had some instability in predicting complex executive functioning, although the Peterson and Posner model was the best predictor of complex executive function performance.

These models's ability to predict performance relies on the relationships between dependent variables and the latent variables. According to the correlations in the Miyake et al. (2000) model, Stroop was more related to Keep Track and two Back than Go No Go. Adding additional inhibition tasks, such as Standard CPT and Flanker, may prove a better fit and more robust model. The Faneros (2011) model can be strengthened by similar additional measures of the latent constructs. The primary weakness in the two models, however, is the lack of explanatory relationship between the latent variables. Instead, the two models allow the sub-domains of executive function to work concurrently, perhaps related by the fact the processes inhabit the same brain. The explanation of how each sub-domain of executive function works together is something the Peterson and Posner (2012) and Barkley (2012) models do well.

The Barkley (2012) model has very good model fit due to its complexity, however the model did not emerge as the favored model because the other models fit while being more parsimonious. Additionally only two of the eight regressions were significant. This would suggest that these latent variables could be collapsed. Self-sensing and self-speech could be collapsed into a self-orienting latent factor and self-appraisal and self-play into a self-performance management latent factor. This suggestion is not from an exploratory factor analysis but from the theoretical background given by Barkley in his 2012 book *Executive Functions*. Barkley's basis for the separation of these factors was drawn from his own clinical work with people with ADHD and his understanding of how each factor can impact an individual's ability to plan and work towards goals. Given the results in this study, it is not necessary to separate these factors in a non-ADHD population.

The Peterson and Posner (2012) and Barkley (2012) theories have biological and behavioral evidence of each process and how the processes interact. This evidence informs the respective model, even if the evidence comes from distinctly different methodologies. The Peterson and Posner theory suggests executive functioning behavior arises from the attentional system, and this supposition is supported given the Peterson and Posner model was the only one that could account for performance on the complex executive functioning task. However, the model does not predict the resulting path correctly. The results of the path analysis show that the fronto-parietal sub-system was crucial in determining complex executive functioning, leaving the cingo-opercular system's impact on overall functioning unclear. The fact that the regressions in the model were not significant could be an indicator of why complex performance was not accurately predicted. Reassessing the theory and understanding how these regressions were connected could potentially yield a much better understanding of executive functioning.

Across all of the theories and their respective models, it was clear that the idea of inhibition is a central concept and probably plays a large role in executive function. However, based on the results of this study, it is still unclear as to the number of latent factors of executive function beyond inhibition. Based on the best fitting model of both the base model and the extended model predicting complex behavior, there was good evidence for the four-factor model of executive functioning. The Peterson and Posner (2012) model showed the centrality of inhibition by placing the latent factor at the top of the model and theorize all other factors were influenced by inhibition. Likewise, Barkley (2012) had inhibition at the top of his model and Miyake and Friedman (2012), though the model didn't fit the data, theorized inhibition to be central to executive functioning.

The goal of this study was to bring greater clarity to the complex area of research that is executive function. The study assessed three theoretical perspectives on executive function by comparing the ability for the respective models to fit a sample of drawn from a normal population. Overall, the Peterson and Posner (2012) model was the most robust model for fitting the data and predicting performance on a complex executive function task. The failure of the Miyake and Friedman (2012) model to fit the data reveals that executive function should be thought of as a process model such as Peterson and Posner or Barkley (2012) have proposed. This study also addressed the definition problem of executive function by guiding future definitions to include this important fact of thinking.

Philosophically speaking, human beings are individuals that exist within the confines of time. Ray Cummings (1973) once wrote "Time is what keeps everything from happening at once." Humans think from word to word or image to image. Time is necessary to understand the history of human civilization and human evolution. That knowledge cannot be obtained all at once, despite how much some people wish differently. The Peterson and Posner (2012) theory has described our cognitive processes happening in a linear fashion. Cognitive data is processed and then decisions are made. Future theories of executive function should reflect this.

References

Allport, G. W. (1955). *Becoming: basic considerations for a psychology of personality*. New Haven, CT US: Yale University Press.

Altamirano, L. J., Miyake, A., & Whitmer, A. J. (2010). When mental inflexibility facilitates executive control: Beneficial side effects of ruminative tendencies on goal maintenance. *Psychological Science*, *21*(10), 1377-1382. doi:10.1177/0956797610381505

Automated Neurological Assessment Metrics (Version 4) [Computer software]. (2013). Seattle, WA: Millisecond Software LLC.

Baddeley, A. (1986). *Working memory*. New York, NY US: Clarendon Press/Oxford University Press.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD.*Psychological Bulletin*, *121*(1), 65-94. doi:10.1037/0033-2909.121.1.65

Barkley, R. A. (2011). Attention-deficit/hyperactivity disorder, self-regulation, and executive functioning. In K. D. Vohs, R. F. Baumeister (Eds.) , *Handbook of self-regulation: Research, theory, and applications (2nd ed.)* (pp. 551-563). New York, NY US: Guilford Press.

Barkley, R. A. (2011). *Barkley Deficits in Executive Functioning Scale (BDEFS)*. New York, NY US: Guilford Press.

Barkley, R. A. (2012). *Executive functions: What they are, how they work, and why they evolved*. New York, NY: Guilford Press.

Barkley, R. A., & Murphy, K. R. (2010). Impairment in occupational functioning and adult ADHD: The predictive utility of executive function (EF) ratings versus EF tests. *Archives Of Clinical Neuropsychology*, *25*(3), 157-173. doi:10.1093/arclin/acq014

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology, 39:1*, 15-22. doi: 10.1080/00221309.1948.9918159

Busemeyer, J. R, & Diederich, A. (2010). *Cognitive Modeling.* Thousand Oaks, CA: Sage Publications, INC.

Caldwell, J. A., & Ramspott, S. (1998). Effects of task duration on sensitivity to sleep deprivation using the multi-attribute task battery. *Behavior Research Methods, Instruments & Computers*, *30*(4), 651-660. doi:10.3758/BF03209483

Cummings, Ray. (2007). *Girl in the Golden Atom.* Retrieved from http://www.gutenberg.org/ebooks/21094 (Original work published 1922)

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, *62*, 119-142.

Dosenbach, N. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control.*Trends In Cognitive Sciences*, *12*(3), 99-105. doi:10.1016/j.tics.2008.01.001

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149. doi:10.3758/BF03203267

Faneros, M. A. (2011). *Composite Components of Cognition: Development of the Executive Function Index.* Thesis (M.S.): University of Oklahoma.

Fox CJ, Mueller ST, Gray HM, Raber J, Piper BJ (2013) Evaluation of a Short-Form of the Berg Card Sorting Test. *PLoS ONE 8*(5): e63885. doi:10.1371/journal.pone.0063885

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., Corley, R. P., & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science*, *18*(10), 893-900. doi:10.1111/j.1467-9280.2007.01997.

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology, 47*(5), 1410-1430. doi:10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal Of Experimental Psychology: General, 137*(2), 201-225. doi:10.1037/0096-3445.137.2.201

Gardner, T. W., Dishion, T. J., & Posner, M. I. (2006). Attention and adolescent tobacco use: A potential self-regulatory dynamic underlying nicotine addiction. *Addictive Behaviors*, *31*(3), 531-536. doi:10.1016/j.addbeh.2005.05.018

Harlow, J. M. (1848). Passage of an iron rod through the head. *Boston Medical and Surgical Journal, 39*, 389-393.

Inquisit (Version 4.0.3) [Computer software]. (2013). Seattle, WA: Millisecond Software LLC.

Jonides, J. (1983). Further toward a model of the mind's eye's movement. *Bulletin Of The Psychonomic Society*, *21*(4), 247-250.

Lee, G. P., Strauss, E., Loring, D. W., McCloskey, L., Hawirth, J. M., & Lehman, R. A. W. (1997). Sensitivity of figural fluency on the Five-Points Test to focal neurological dysfunction. *The Clinical Neuropsychologist, 11,* 59:68.

Loehlin, J. C. (2004). *Latent Variable Models* (4th Ed.). Mahwah, NJ: Lawrence Erlbaum Associates, INC., Publishers.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments & Computers,36*(2), 339-346. doi:10.3758/BF03195580

Macmillan, N. A., & Creelman, C. (2005). Detection theory: A user's guide (2nd Ed.). New York, NY US: Cambridge University Press.

Malabou, Catherine; Rand, Sebastian (Trans). (2008). *What should we do with our brain?* New York: Fordham University Press.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

Mayr, U., & Kliegl, R. (2000). Complex semantic processing in old age: Does it stay or does it go?. *Psychology And Aging, 15*(1), 29-43. doi:10.1037/0882-7974.15.1.29

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (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. doi:10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions.*Current Directions In Psychological Science*, *21*(1), 8-14. doi:10.1177/0963721411429458

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal Of Psychology,81*(2), 111-121. doi:10.1111/j.2044-8295.1990.tb02349.

Mueller, S. T. (2012). PEBL: The psychology experiment building language (Version 0.12) [Computer experiment programming language]. Retrieved Nov. 2012 from http://pebl.sourceforge.net.

O'Reilly, R. C., & Frank, M. J. (2006). Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. *Neural Computation*, *18*(2), 283-328. doi:10.1162/089976606775093909

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review Of Neuroscience, 35*73-89. doi:10.1146/annurev-neuro-062111-150525

Plant, R. W., & Ryan, R. M. (1985). Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally-controlling styles. *Journal of Personality, 53*, 435-449.

Posner, M. I. (2012). Imaging attention networks. *Neuroimage*, *61*(2), 450-456. doi:10.1016/j.neuroimage.2011.12.040

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review Of Neuroscience, 13*25-42. doi:10.1146/annurev.ne.13.030190.000325

Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York, NY US: Scientific American Library/Scientific American Books.

Pribram, K. H. (1973). The primate frontal cortex: Executive of the brain. In , *Psychophysiology of the frontal lobes* Oxford England: Academic Press.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills, 123,* 17-22.

Revelle, W. (2014) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, http://CRAN.R-project.org/package=psych Version = 1.4.3.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal Of Experimental Psychology: General, 124*(2), 207-231. doi:10.1037/0096-3445.124.2.207

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/.

Rothbart, M. K., Sheese, B. E., & Posner, M. I. (2007). Executive attention and effortful control: Linking temperament, brain networks, and genes. *Child Development Perspectives, 1*(1), 2-7. doi:10.1111/j.1750-8606.2007.00002.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43,* 450-461.

Ryan, R. M., Connell, J. P., & Plant, R. W. (1990). Emotions in non-directed text learning. *Learning and Individual Differences, 2,* 1-17.

Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Varied forms of persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion, 15,* 185-205.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal Of Personality And Social Psychology*, *45*(4), 736-750. doi:10.1037/0022-3514.45.4.736

Silk, T. J., Vance, A., Rinehart, N., Bradshaw, J. L., & Cunnington, R. (2008). Dysfunction in the fronto-parietal network in attention deficit hyperactivity disorder (ADHD): An fMRI study. *Brain Imaging And Behavior*, *2*(2), 123-131. doi:10.1007/s11682-008-9021-8

Singh, A. L., Tiwari, T., & Singh, I. L. (2010). Performance feedback, mental workload and monitoring efficiency. *Journal Of The Indian Academy Of Applied Psychology*, *36*(1), 151-158.

Tucha, L., Aschenbrenner, S., Koerts, J., & Lange, K.W. (2012). The five-point test: Reliability, validity and normative data for children and adults. *PLoS ONE 7*(9): e46080. doi:10.1371/journal.pone.0046080

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498-505. doi:10.3758/BF03192720

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the Executive Function Theory of Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *Biological Psychiatry*, *57*(11), 1336-1346. doi:10.1016/j.biopsych.2005.02.006

Table 1

Descriptive Statistics of dependent variables

Task	*M*	*SD*	Range	Skew	Kurtosis
BDEFS	33.59	7.96	40	0.7	0.25
IMI - Interest	16.67	9.56	40.14	0.33	-0.69
Two Back	64.37	15.37	68.41	-0.25	-0.52
Keep Track	0.77	0.15	0.8	-1.15	1.27
Number-Letter	236.35	249.68	1355.95	0.32	-0.08
Category Switch	237.55	198.83	1658.15	0.68	2.7
AOSpan	26.75	12.76	55	-0.01	-0.16
Card Sort	31.44	6.11	39.11	0.48	-0.59
Tower	1839.16	321.98	1516.44	-0.09	0.51
Visual Vigilance	327.60	54.76	333.96	1.99	5.58
FivePT	34.02	10.18	60	-0.81	29.85
Stroop	17.65	7.72	37.36	-0.22	1.29
Flanker	95.27	48.95	273.7	0.12	-0.55
Go-no-go	3.59	1.4	6.44	-0.07	0.01
Standard CPT	5.39	1.93	6.49	-0.87	-1.2
MATB-II SysMon	0.41	0.200	0.88	0.59	0.25
MATB-II Comm	0.87	0.17	0.70	-1.41	1.55
MATB-II Tracking	55.37	31.06	16283	2.09	4.14
MATB-II Resource	-569.63	772.80	3585.7	-0.56	-0.39
MATB-II Performance	0.00	0.50	3.35	0.05	1.02

Note. *N* = 306. All values represent raw, nonstandardized scores.

Table 2

Pearson Correlations Coefficients for Target Dependent Variables

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. IMI - Interest	**---**														
2. Two Back	**0.18**	**---**													
3. Keep Track	0.11	**0.35**	**---**												
4. Number-Letter	**0.12**	0.03	**0.11**	**---**											
5. Category Switch	-0.02	0.1	**0.13**	**0.29**	**---**										
6. AOSpan	0.04	**0.24**	**0.23**	0.09	0.09	**---**									
7. Card Sort	-0.05	-0.07	-0.08	-0.02	-0.01	-0.1	**---**								
8. Tower	**0.19**	**0.32**	**0.15**	0.03	0.02	0.16	0	**---**							
9. Visual Vigilance	**-0.2**	**-0.19**	**-0.18**	-0.06	-0.06	-0.05	0.01	-0.09	**---**						
10. FivePT	0.09	**0.26**	0.08	0.06	0.08	**0.24**	-0.05	**0.12**	0.01	**---**					
11. Stroop	-0.04	**0.14**	**0.12**	-0.09	**-0.13**	0.07	0	0.1	-0.04	0.06	**---**				
12. Flanker	0.02	0.08	0.03	0.06	0.08	0.07	0.03	0.04	-0.03	0.02	0.07	**---**			
13. Go-no-go	0.03	**0.12**	0.11	**0.13**	0.1	0.15	-0.05	0.05	**-0.13**	0.04	0.04	**0.37**	**---**		
14. Standard CPT	**0.12**	**0.12**	**0.21**	0.07	**0.11**	0.05	-0.05	0.08	**-0.17**	0.04	0.03	**0.16**	**0.29**	**---**	
15. MATB-II Overall	0.1	0.09	0.03	**0.13**	0.1	-0.04	0.02	-0.04	0.03	0.05	-0.04	0.03	-0.03	0.01	**---**

Note. $N = 306$. $p < .05$ are in boldface.

Table 3

Model Fit Indices

Model	*df*	χ^2 (*p*)	CFI	SRMR	AIC	BIC
Miyake et al. (2000)*	6	12.643(.05)	0.924	0.051	4612	4668
Faneros (2011)	8	6.802(.56)	1.000	0.030	4709	4757
Peterson and Posner (2012)	10	7.047(.72)	1.000	0.024	4625	4692
Barkley (2012)	3	3.725 (.29)	0.989	0.020	4952	5019

Note. Good model fits typically will not have a significant X^2, a Bentler Comparative Fit Index (CFI) equal to or greater than 0.95, and a Standardized Root Mean Square Residual (SRMR) of less than 0.08. A lower Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are desired when comparing two models.
*Model run post-hoc.

Table 4

Path Analysis for MATB-II on to Faneros (2011) model

					Coefficients for paths		
Model	*df*	χ^2	CFI	BIC	Shifting	Updating	Inhibition
1. 2 Paths (shifting, updating)	11	16.42	0.925	5235	.34	.59	---
2. 2 Paths (updating, inhibition)	11	16.42	0.925	5235	---	.17	.39
3. 1 Path (shifting)	**12**	**16.60**	**0.936**	**5229**	**.32**	**---**	**---**
4. 1 Path (inhibition)	12	16.60	0.936	5229	---	---	.25

Note. Endorsed model is in bold.

Table 5

Path Analysis for MATB-II on to Peterson and Posner (2012) model

					Coefficients for paths	
Model	*df*	χ^2	CFI	BIC	Circulo-opercular	Fronto-parietal
1. 2 Paths	15	12.73	1.0	5166	-.05	.22*
2. 1 Path (Circulo-opercular)	16	18.12	.98	5166	.05	---
3. 1 Path (Fronto-parietal)	**16**	**13.05**	**1.0**	**5161**	**---**	**.20***

Note. Endorsed model is in bold.

*p < .01.

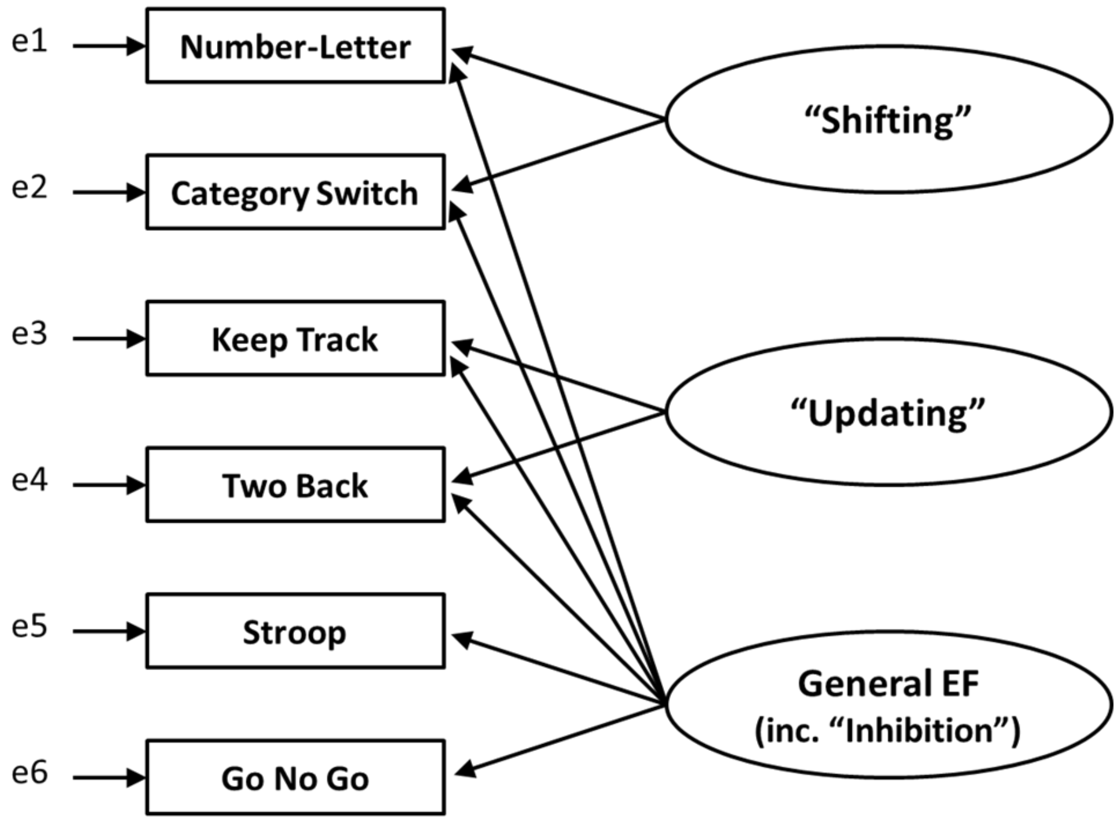

Figure 1. Model of executive function as presented by Miyake and Friedman (2012).

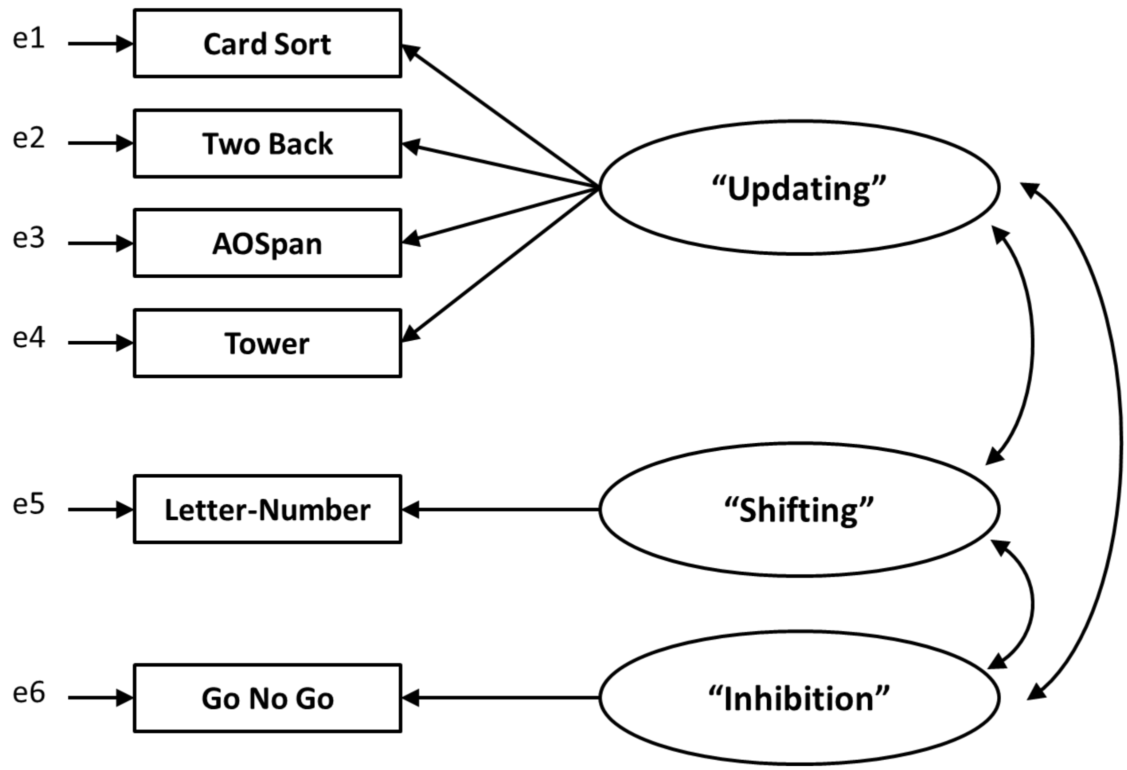

Figure 2. Model of executive function resulting from exploratory factor analysis (Faneros, 2011).

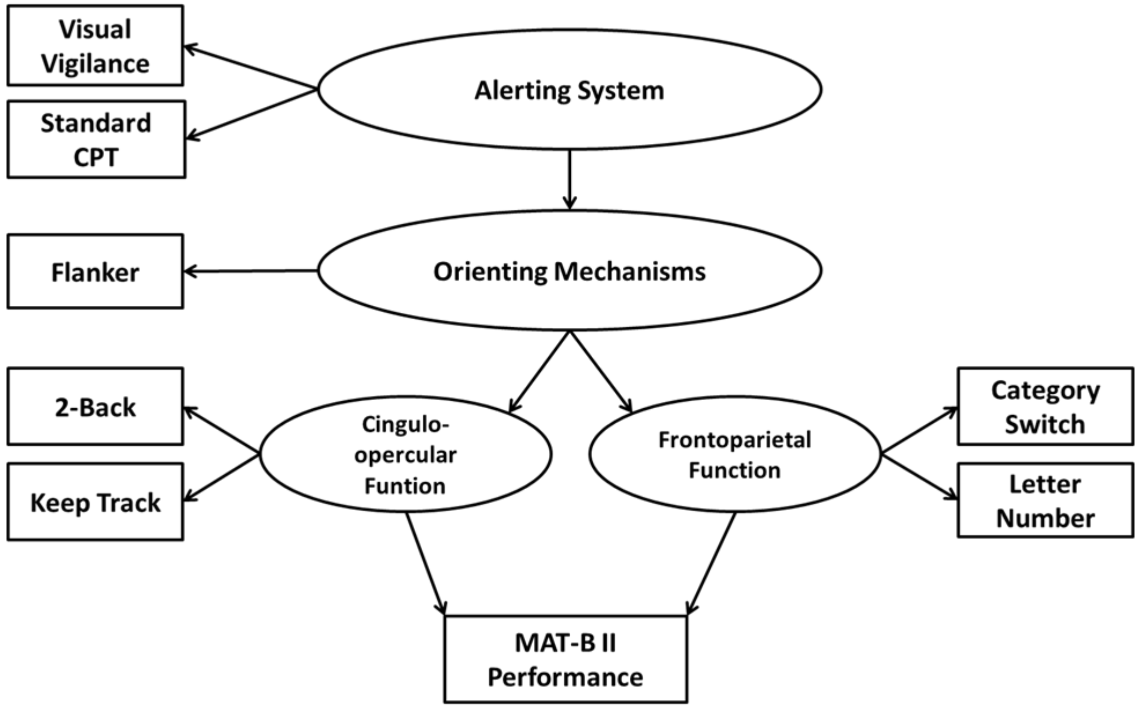

Figure 3. Model of attention based on Petersen and Posner (2012).

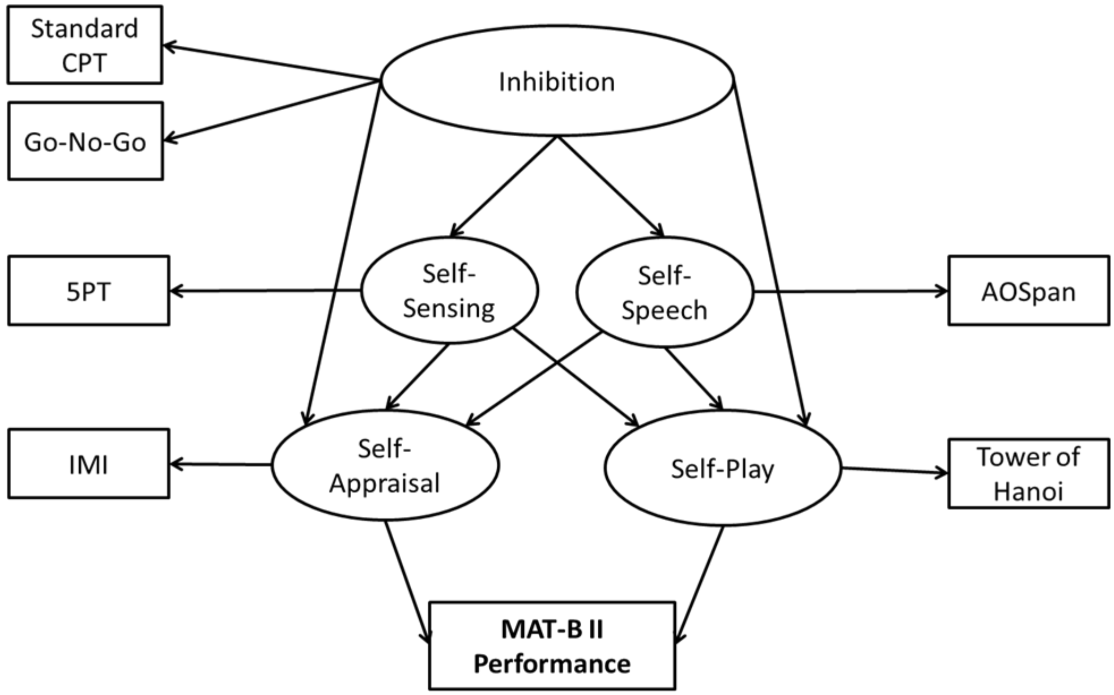

Figure 4. Model of executive function based on Barkley (2012). Note the process Self-Awareness is not contained in this model.

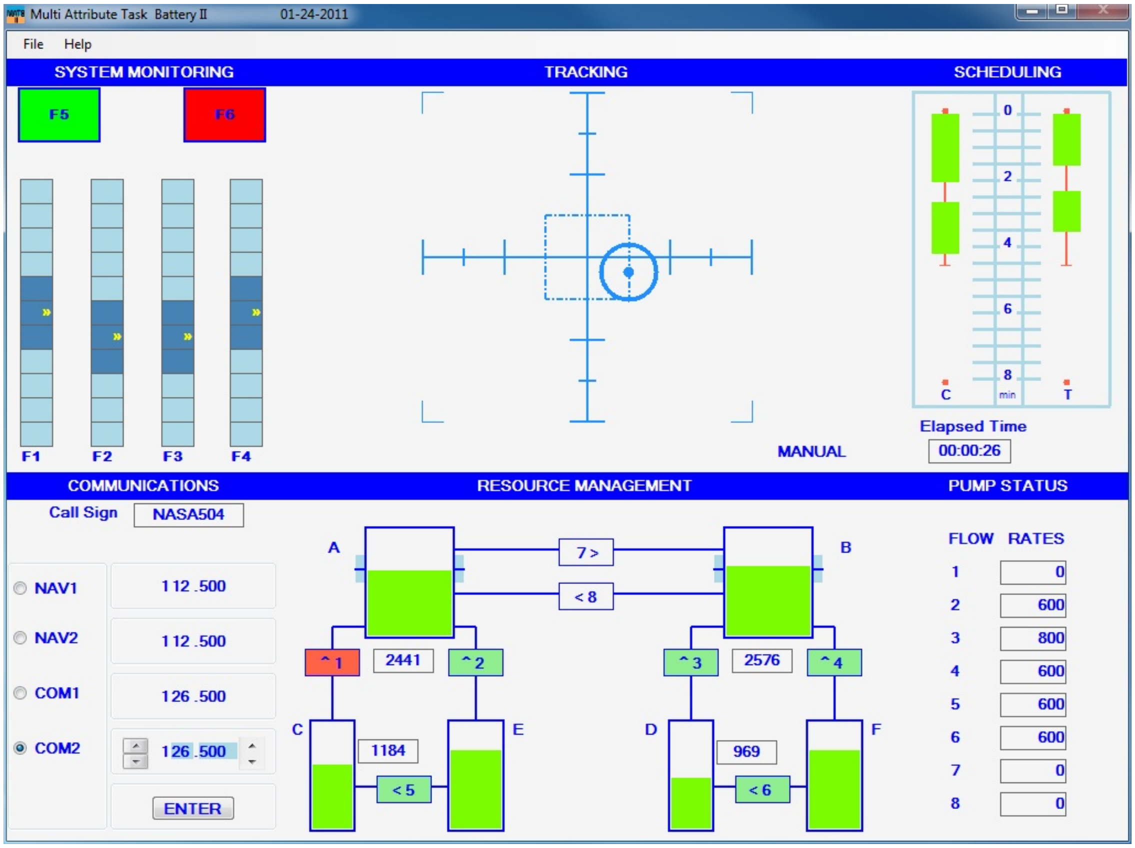

Figure 5. Example of a typical Multi-Attribute Task Battery II testing screen.

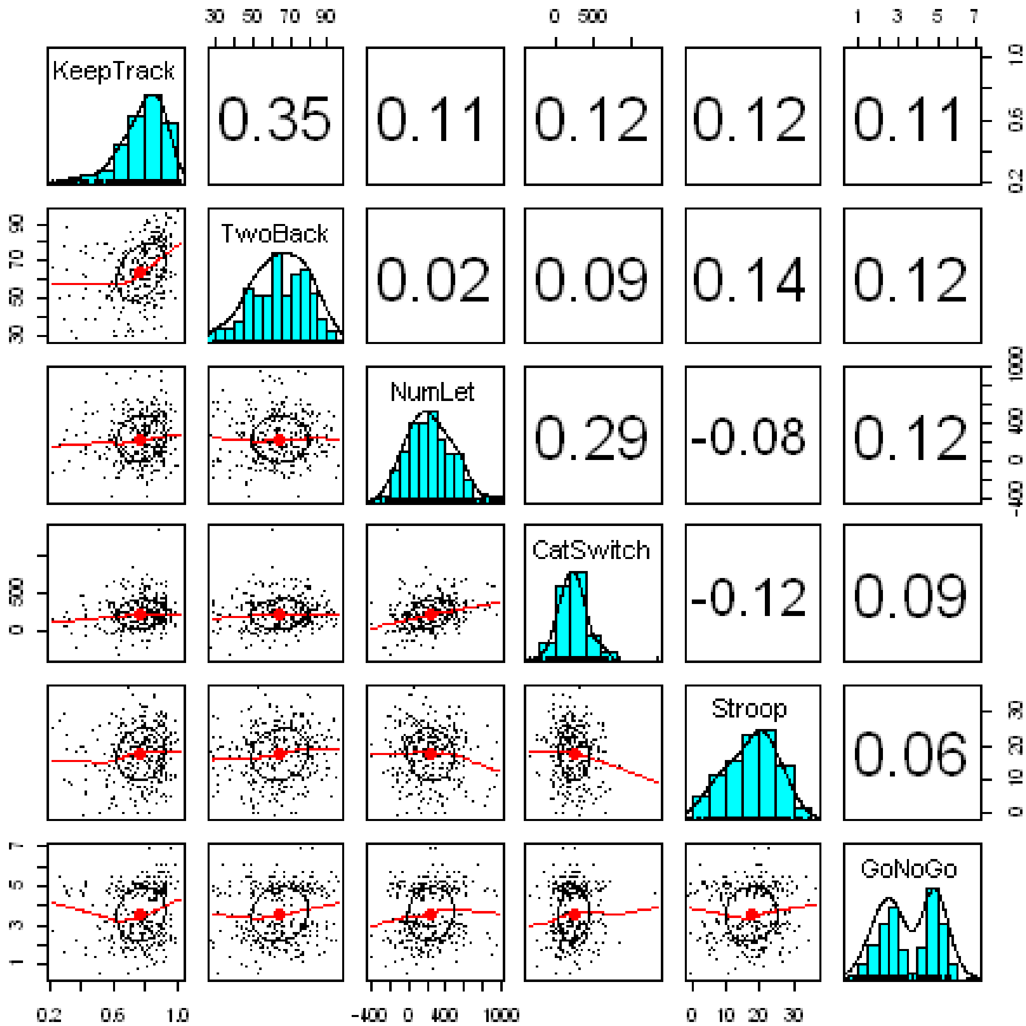

Figure 6. Correlations, distributions, and dot plots of the tasks used in the Miyake and Friedman (2012) model.

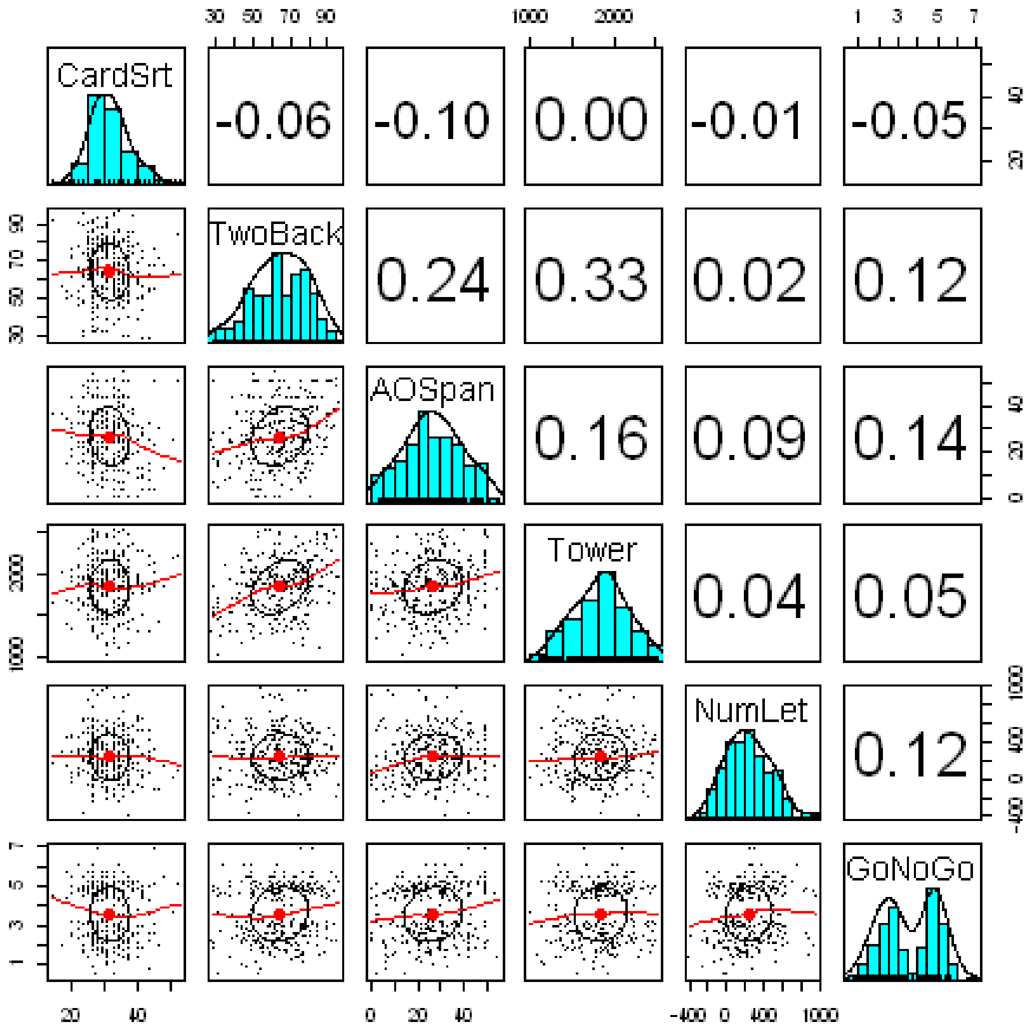

Figure 7. Correlations, distributions, and dot plots of the tasks used in the Faneros (2011) model.

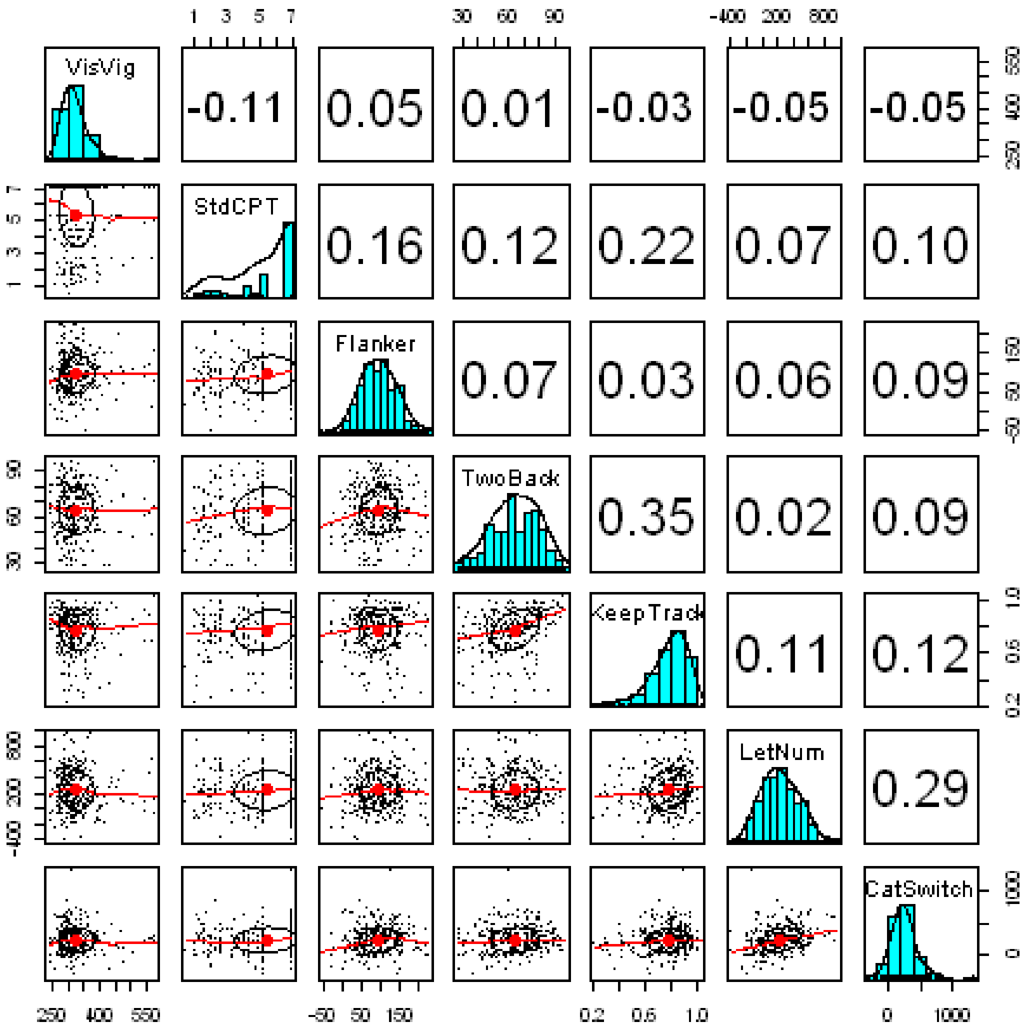

Figure 8. Correlations, distributions, and dot plots of the tasks used in the Peterson and Posner (2012) model.

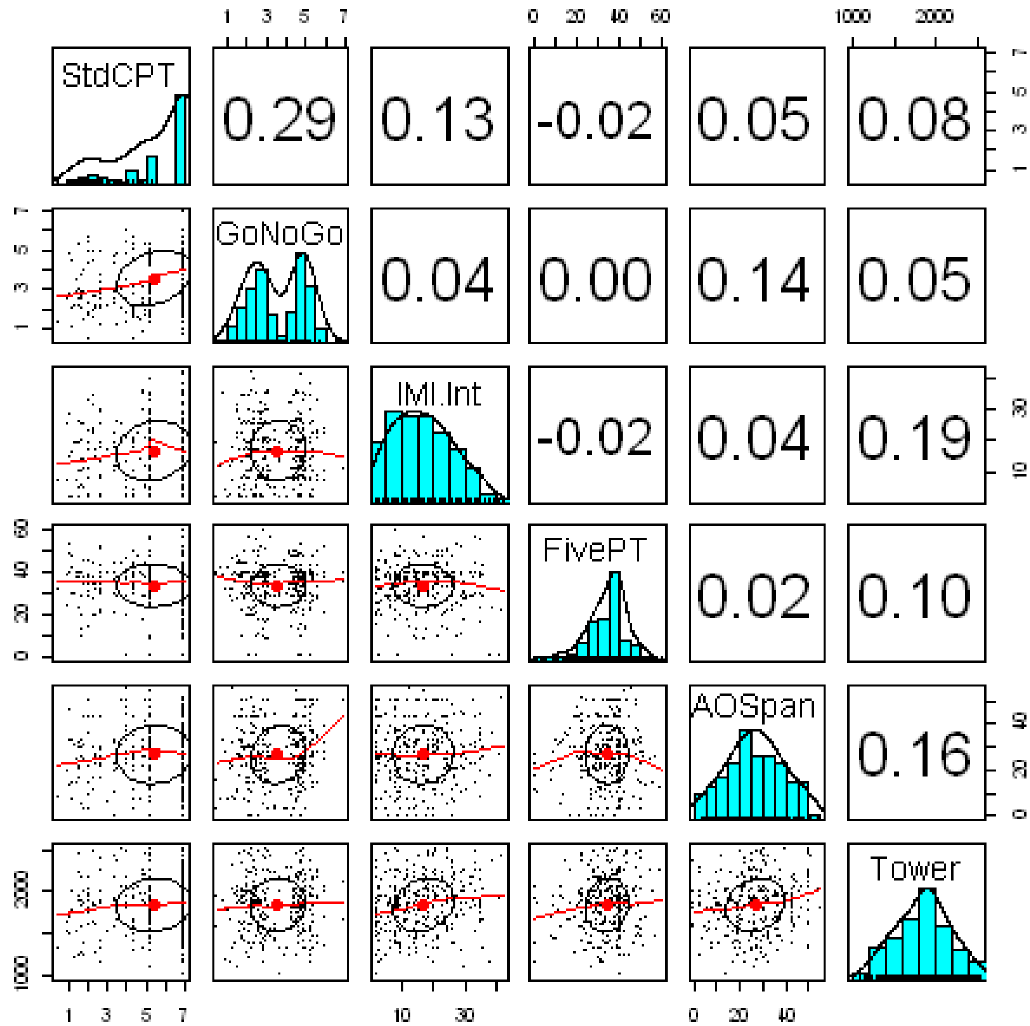

Figure 9. Correlations, distributions, and dot plots of the tasks used in the Barkley (2012) model.

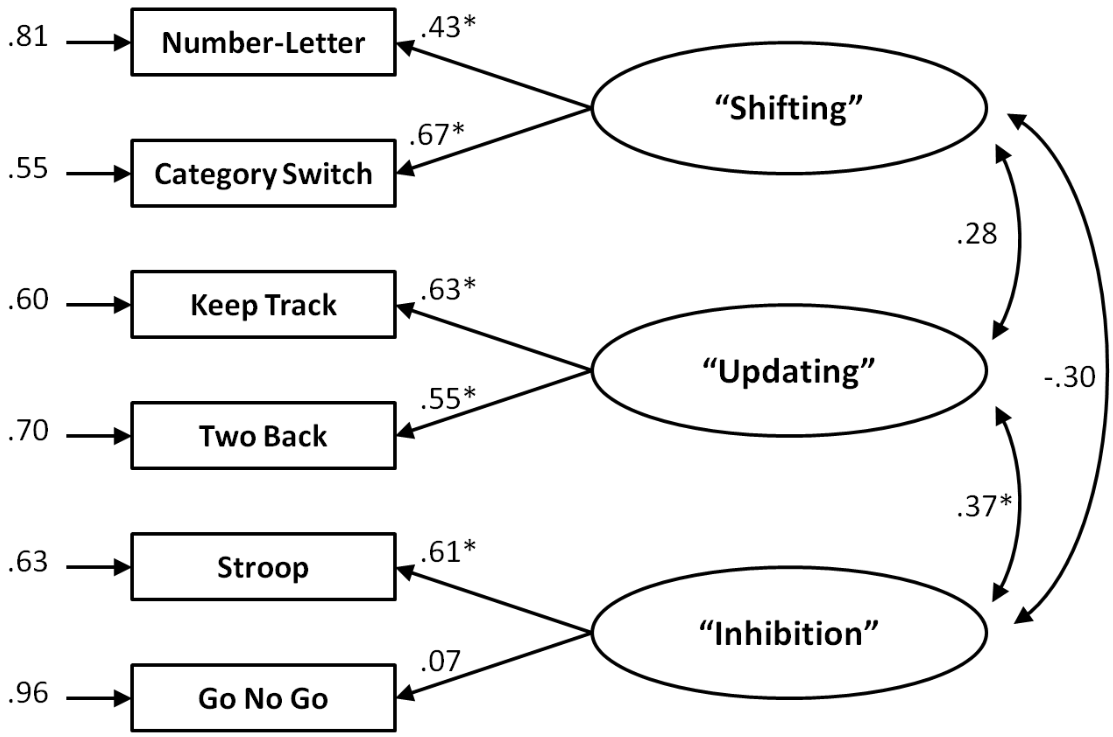

Figure 10. Factor loadings for the Miyake et al. (2000) model. Factor loading significance determined by a Wald test.
$^*p < .05$.

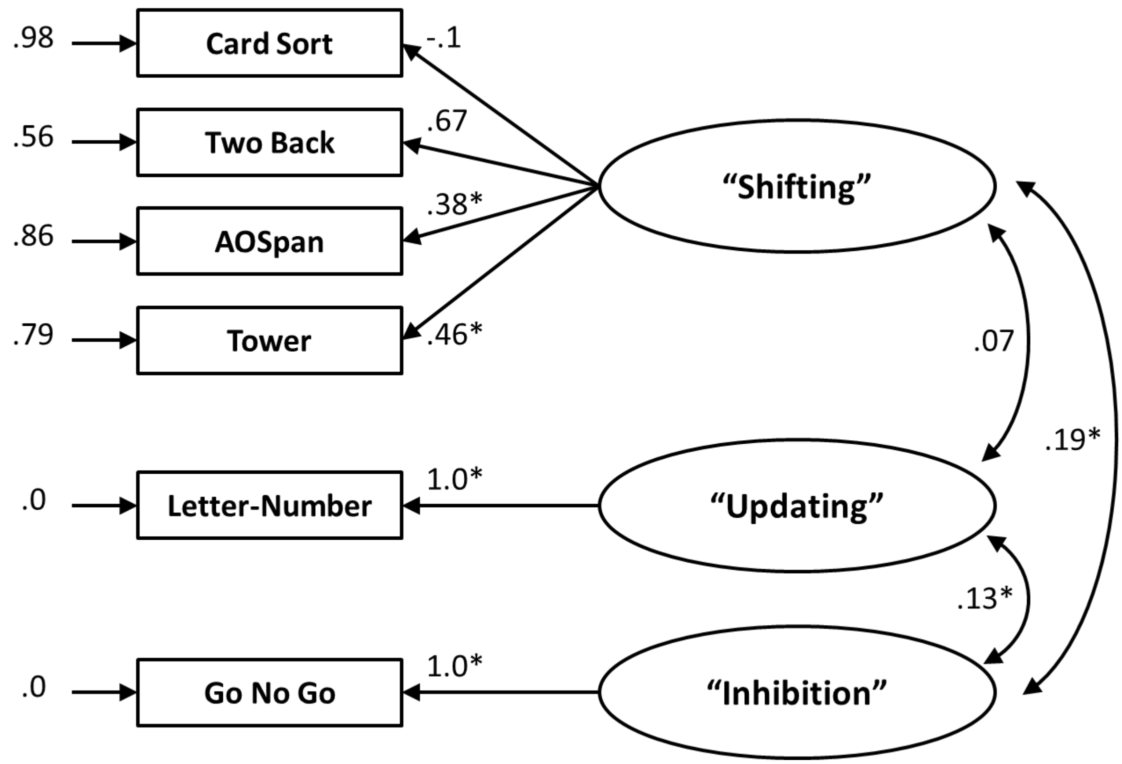

Figure 11. Factor loadings for the Faneros (2011) model. Factor loading significance determined by a Wald test.
*$p < .05$.

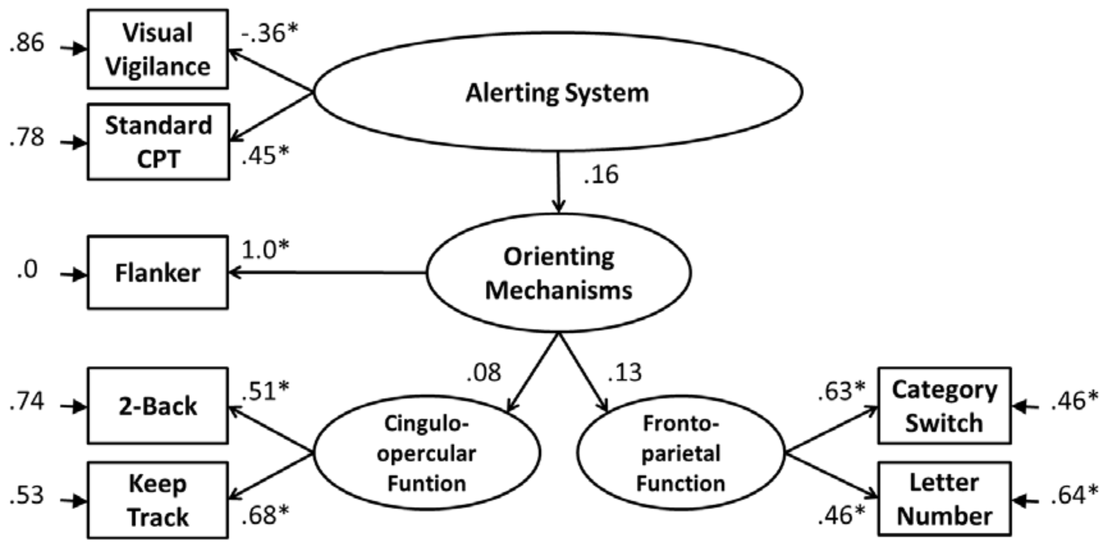

Figure 12. Factor loadings for the Peterson and Posner (2012) model. Factor loading significance determined by a Wald test.
*$p < .05$.

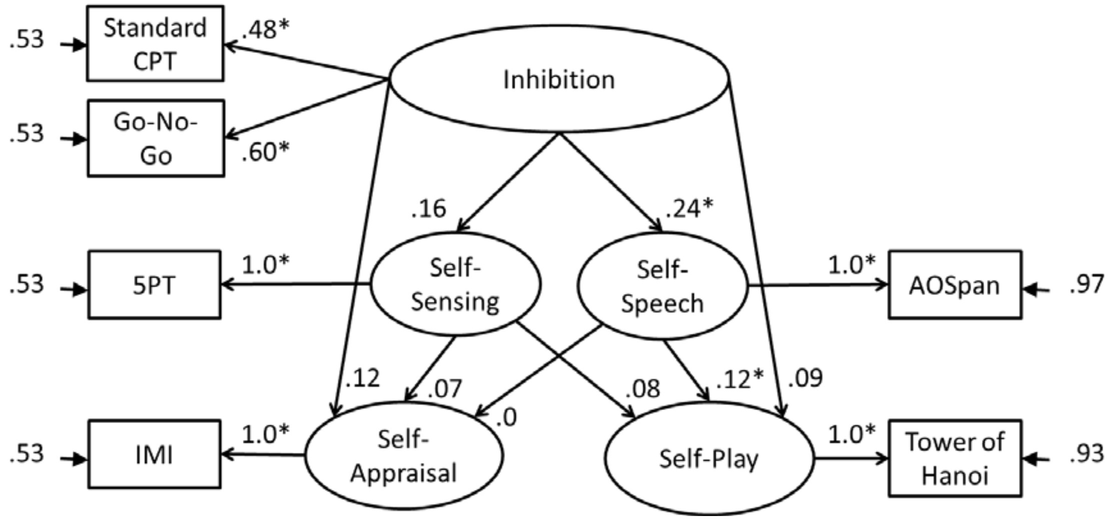

Figure 13. Factor loadings for the Barkley (2012) model. Factor loading significance determined by a Wald test.
*p < .05.