1-23-2015

Executive Functioning is Related to Intraindividual Variability: A Response Time Distributional Approach

Peter V. Stewart
Indiana University of Pennsylvania

EXECUTIVE FUNCTIONING IS RELATED TO INTRAINDIVIDUAL VARIABILITY: A RESPONSE TIME DISTRIBUTIONAL APPROACH

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Psychology

Peter V. Stewart

Indiana University of Pennsylvania

August 2014

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Psychology

We hereby approve the dissertation of

Peter V. Stewart

Candidate for the degree of Doctor of Psychology

_____ _____
David J. LaPorte, Ph.D.
Professor of Psychology, Advisor

_____ _____
Dasen Luo, Ph.D.
Professor of Psychology

_____ _____
William M. Meil, Ph.D.
Professor of Psychology

_____ _____
Michael D. Franzen, Ph.D.
Chief, Psychology and Neuropsychology
Department of Psychiatry
Allegheny General Hospital

ACCEPTED

_____ _____
Timothy P. Mack, Ph.D.
Dean
School of Graduate Studies and Research

Title: Executive Functioning is Related to Intraindividual Variability: A Response Time Distributional Approach

Author: Peter V. Stewart

Dissertation Chair: Dr. David J. LaPorte

Dissertation Committee Members: Dr. Michael D. Franzen
Dr. Dasen Luo
Dr. William M. Meil

Intraindividual variability refers to short-term fluctuations in performance that may be indexed as trial-to-trial variability on response time (RT) tasks. A growing body of evidence suggests that such variability may be a sensitive indicator of the integrity of prefrontal cortical regions. This possibility was tested in a broad clinical sample by investigating the relationship between intraindividual variability on the Wisconsin Card Sorting Test and a large battery of executive functioning tasks using mathematical models of RT variability. The executive functioning tasks predicted intraindividual variability with a large effect size. The relationship between intraindividual variability and cognitive control tasks was marginally higher than the intercorrelations of the tasks themselves. Thus, intraindividual variability demonstrates considerable convergent validity with more traditional neuropsychological assessment techniques. Trial-to-trial variability was specifically related to measures of set shifting and this relationship was only partially mediated by more basic attentional functions. The findings of this study are consistent with the interpretation that intraindividual variability in response time may index the consistency with which individuals are able to regulate various cognitive control subprocesses involving attentional control, active engagement of response sets, and the ability to suppress task irrelevant information in the pursuit of a self-selected

goal. The primary neurostructural substrates of RT variability may lie within prefrontal cortical regions including ventrolateral, dorsolateral, and anterior cingulate cortex. Future research may set the stage for a new clinical measurement tradition based on the application of intraindividual variability data.

ACKNOWLEDGEMENTS

This project represents the culmination of three years of effort; three years during which the path forward was not always forthcoming or easy to tread. Persevering through that time period would not have been possible without the tireless support of my family, especially my mother Susan Stewart and my father Dr. Charles Stewart. I am also indebted to my significant other and closest companion Laura Gilman whose patience and ability to remain supportive while listening to my rants about likelihood maximization and rostro-caudal prefrontal cortical organization were truly impressive. I would also like to thank the supportive supervisors at the Dayton VA Medical Center who demonstrated patience and helped remind me that sometimes, it's good enough.

Dr. Michael Franzen made this project possible through introducing me to the subject matter and providing on-site resources throughout the data collection process, thank you Dr. Franzen. I also owe a tremendous debt of gratitude to my dissertation chair and advisor Dr. David LaPorte, for not only providing his assistance during the completion of this project, but for believing in me and helping to shape me throughout my graduate career. It is unlikely that I ever would have been able to enter the profession of clinical neuropsychology without his mentorship and for that I am so grateful.

During the planning of this project, a series of conversations with Dr. Dasen Luo convinced me to "go all out" and push the limits of my capabilities. 10,000 data points and countless hours in front of my statistics workbench later, I can say that I was glad I took his suggestions. Thank you Dr. Luo. Finally, I would like to acknowledge all those who put themselves through 5 - 6 years of suffering in graduate school so that they may join in the struggle against human suffering.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Historically, neuropsychology research has focused on mean levels of performance to the relative exclusion of other methodological approaches (MacDonald, Nyberg, & Bäckman, 2006). This is unfortunate given significant evidence that intraindividual performance variability both within and across occasions is substantial in magnitude (Nesselroade & Salthouse, 2004), can be reliably measured (Hertzog, Dixon, & Hultsch, 1992), and may obscure the detection of meaningful differences by measures of central tendency (Rabbitt, Osman, Moore, & Stollery, 2001). Indeed, some researchers have noted the substantial problem variability poses to classical theories of measurement and typical assessment practices (Lindenberger & von Oertzen, 2006). The study of variability is not merely a theoretical curiosity; a substantial body of work indicates that variability data can contribute meaningfully to clinical diagnosis (Kaiser et al., 2008; Leth-Steensen, King Elbaz, & Douglas, 2000), the prediction of important outcomes like mortality (Shipley, Der, Taylor, & Deary, 2006), and the quantification of individual differences on neuropsychological tasks (Heathcote, Popiel, & Mewhort, 1991) above and beyond the mean.

Intraindividual variability in response time is also related to a variety of clinical conditions including attention deficit hyperactivity disorder (Hervey et al., 2006), closed head injury (Stuss, Murphy, Binns, & Alexander, 2003), dementia (Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000), mild cognitive impairment (Bielak, Hultsch, Strauss, MacDonald, & Hunter, 2010), schizophrenia (Rentrop et al., 2010), and other forms of mental illness (Bunce, Handley, & Gaines, 2008). Some authors have attributed

increased variability to impairment in executive functioning (West, Murphy, Armilio, Craik, & Stuss, 2002) and for good reason: the magnitude of performance variability on response time tasks parallels the degree of frontal involvement in dementia subtypes (Murtha, Cismaru, Waechter, & Chertkow, 2002), is associated with damage to the frontal lobes (Stuss et al., 2003), and appears to be meaningfully related to other measures of this cognitive domain (Lövdén, Li, Shing, & Lindenberger, 2007). Studies utilizing both structural and functional neuroimaging have converged on the interpretation that the prefrontal cortex is a key determinant of variability in response time (MacDonald et al., 2006).

There are few investigations of the relationship between intraindividual variability and well-established measures of executive functioning, despite evidence that certain types of performance variability may be sensitive measures of the efficiency of executive control processes (Bellgrove, Hester, & Garavan, 2004). The vast majority of existing studies restrict their analyses to discrete clinical syndromes (e.g., dementia of the Alzheimer's type, focal brain lesions), which limits the generalizability of findings. Furthermore, most clinical investigations make use of relatively crude indices of variability, such as the intraindividual standard deviation (iSD), in combination with regression based statistical control approaches (Ram & Gerstorf, 2009). Such techniques may be statistically inappropriate and theoretically problematic (Salthouse & Berish, 2005; Schmiedek, Lövdén, & Lindenberger, 2009; Slifkin & Newell, 1998). This study uses contemporary statistical techniques to investigate the relationship between intraindividual variability and executive functioning in the service of clarifying the practical applications of intraindividual variability in a clinical context.

CHAPTER 2

LITERATURE REVIEW

A Taxonomy of Variability

Before discussing the theoretical, methodological, and clinical significance of intraindividual variability, it is first necessary to define what is meant by this term. Variability is not a unitary construct and at its broadest level, variability can exist between persons, between measures, or between occasions (Cattell, 1966). According to this division, performance variability may be defined in three primary ways: between individual variability (diversity), within individual across-trial variability (dispersion), and within individual across-occasion variability (consistency; Stuss, Pogue, Buckle, & Bondar, 1994). Each type of variability carries with it a unique history of application in the behavioral sciences and a set of practical and methodological considerations.

Variability existing between persons is *inter*-individual variability and is typically used to index differences in individual performance. This sort of variability is traditionally measured by the use of mean level comparisons of performance on one measurement occasion and constitutes the great bulk of cognitive and neuropsychological research (MacDonald et al., 2006). In the taxonomy presented above, this sort of variability is known as diversity. Variability existing within persons is known as *intra*-individual variability and may be measured by comparison of performances on different tasks (or trials) on a single measurement occasion, or as fluctuations in performance on one task across occasions (MacDonald et al., 2006). This type of variability encompasses both dispersion (when measured across trials) and consistency (when measured across occasions). Although it has been historically neglected, the study of intraindividual

variability holds with it tremendous promise in helping to explicate individual differences, development, and human behavior in general (Nesselroade & Ram, 2004). This type of variability is characteristic of all self-organizing dynamic systems (Li, Huxhold, & Schmiedek, 2004) and given that humans are self-organizing dynamic systems (Ford, 1994), variability is inherent in all aspects of human behavior and cognition.

Intraindividual variability itself may exist in a number of different forms. The notion of "intraindividual dynamics" encompasses this type of within person variation as it exists across time, in both univariate and multivariate forms (Li et al., 2004). The process of *becoming* (Ford, 1994) or *developing* (Li et al., 2004) refers to change that is typically enduring and takes place across months, years, or decades (e.g., neuromaturation, aging, complex skill acquisition). In contrast, *being* (Ford, 1994) or *functioning* (Li et al., 2004) takes place over a comparably shorter time span (e.g., days, hours, minutes). This type of variability has been called inconsistency, "lack of processing robustness", "wobble", or "lability" (Hultsch, Strauss, Hunter, & MacDonald, 2008) and refers to processes such as rapid fluctuations in moods, attention, or neuromodulatory conditions.

Intraindividual variability in functioning (i.e., short term within person variations in functioning) can be both adaptive and maladaptive (Ram & Gerstorf, 2009). Within intraindividual dynamics, Li and colleagues (2004) discuss three types of adaptive variability, typically observed on tasks that are amenable to strategy use. Plasticity (large gains as a result of learning after exposure to a task), diversity (using a multitude of different exploratory strategies in order to optimize performance during task acquisition),

and adaptability (quickly recovering by adapting when optimal functioning is challenged) can all be adaptive responses to novel situations such as completing a neuropsychological task. In contrast, random processing fluctuations around a given maximum level of functioning indicate a non-adaptive form of variation (MacDonald, Li, & Bäckman, 2009).

Much of the research on this type of maladaptive intra-individual variability has focused on within person variation on response time tasks. The measurement of variability in response latencies on such measures provides a precise index of rapidly changing internal processes and their effects on behavioral output. Extremely short-term performance variations such as those observed on consecutive trials of a response time task are likely subserved by different underlying sources of systematic variation than those that occur on the scale of days or months (Ram, Rabbitt, Stollery, & Nesselroade, 2005). The study of both types of variability (i.e., variability over shorter or longer periods) has proved to be fruitful, but there is evidence to suggest that endogenous sources of variation (e.g., those that underlie individual differences in brain functioning) are best captured over short intervals such as trial-to-trial fluctuations in performance on cognitive tasks (MacDonald et al., 2009).

This sort of short-term processing lability seems to be almost invariantly associated with injury (Stuss et al., 1994) and disease (Holtzer, Verghese, Wang, Hall, & Lipton, 208). It is also indicative of poorer cognitive functioning (Bielak et al., 2010), especially executive functioning (de Frias, Dixon, Fisher, & Camicioli, 2007; Murtha et al., 2002; Stuss et al., 2003; West et al., 2002). For these reasons among others, the study

of short-term, trial-to-trial variability in response time may be particularly informative about endogenous (i.e., neurological) influences on test performance.

A Threat to the Measurement of Central Tendency

What are the reasons we might want to consider an intraindividual approach to studying variability? This represents a distinct departure from the prevailing emphasis of the field. The literature is replete with examples of the knowledge we have gained by measuring diversity with measures of central tendency. Although we have learned a great deal from this method, when within person variation is large, the analysis of mean levels of performance can lead to incorrect estimates of between group differences (Hultsch & MacDonald, 2004). Regarding response time tasks specifically, some authors have gone so far as to assert that mean levels of performance are little more than gross summary statistics that are actually the products of variability themselves (Rabbitt et al., 2001). Much in the same way that classical Newtonian physics has proved useful yet inadequate to completely describe the physical universe (Hawking, 1998), it is quite possible that an exclusive emphasis on between person variation indexed by measures of central tendency does no better when attempting to describe many neuropsychological phenomena. In some situations in fact, Newton's system makes wildly inaccurate predictions and fails to adequately describe known phenomena (Hawking, 1998); the same may be said of summary statistics such as the mean under certain circumstances (Nesselroade, 2002).

There are a number of conditions under which the mean may be misleading. Consider an example from the physical sciences such as the analysis of light spectra involved in photosynthesis. Although it is possible to determine the mean wavelength of light absorbed by chloroplasts, this does little to help us understand the underlying

distribution of light energy involved in this crucially important process. Due to the unique chemical characteristics of chlorophyll molecules, the distribution of light absorbed by these compounds is bimodal in nature (specifically clustering around approximately 450nm and 675nm; Campbell & Reece, 2010) and is thus poorly described by its mean. Similarly, the mean may represent a minority of the population when considering presidential approval ratings, because they are likely to be distributed bimodally in correspondence with political party (Balota & Yap, 2011). Both of these examples show the problems inherent in attempting to represent certain types of distributions with summary statistics such as the mean.

Reaction time probability density functions are another type of distribution that is poorly described by measures of central tendency because they are almost always positively skewed (Luce, 1991). When analyzing highly skewed data, using statistics such as the mean can be inappropriate for any number of reasons and can mask important differences in response patterns (Andrews & Heathcote, 2001; Miller, 1991). Balota & Spieler (1999) illustrated this phenomena using mathematical modeling to capture the entire shape of response time probability density functions on three different word recognition tasks.

Figures 1a and 1b demonstrate a situation in which relying on measures of central tendency may be appropriate. In this scenario, the overall shape of the two distributions is quite equivalent but they are shifted with respect to the mean and here, the mean paints a relatively complete picture of the data. In other circumstances though, such an analysis may be inappropriate. Comparing Figure 1a with Figure 1c shows a comparable shift in

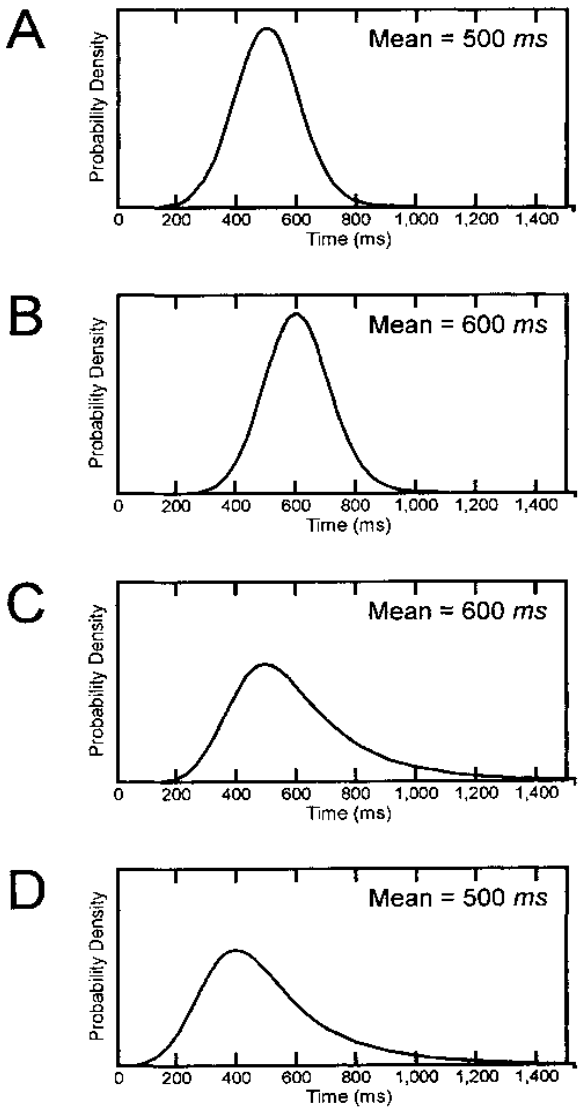

Figure 1. Simulated Distribution Shapes (adapted from Balota & Yap, 2011,p. 161).

means that is caused by the inclusion of a small percentage of observations in the tail of distribution cthat have little effect on the modal portion and a disproportionate effect on the mean. Similarly, there can be opposing effects on the modal and tail portions of a distribution that are significant yet cause no observable changes in the mean of the probability density function (Figure 1a vs. Figure 1d). When using measures of central tendency alone, researchers could conclude that the manipulation was of no consequence

to participants' performance when in fact, there was actually a significant effect. All of the patterns documented in this figure have now been observed empirically (Balota & Yap, 2011).

In other circumstances, there may simply be a lack of correspondence between the mean and other variables under investigation, for whatever reason. The worst performance rule in intelligence testing serves as a useful example of such a phenomenon. As the theory goes, on multi-trial reaction time tasks, an individual's worst performance is more predictive of general intelligence (i.e., g*)* than their best performance or average performance (Alderton & Larson, 1994). This theory has been the subject of some investigation in the intelligence testing literature and has been overwhelmingly supported by most of the studies that have examined it (Coyle, 2003). According to this theory, analysis of an individual's average performance would be far less informative than examining their worst performance. Although the mechanisms underlying this lack of correspondence have yet to be entirely delineated, they are hypothesized to involve factors such as lapses in working memory or individual differences in "neural oscillations" (Coyle, 2003). Both of these processes have been implicated in intraindividual variability (Kaiser et al., 2008; Li & Lindenberger, 1999).

If substantial intraindividual variability exists on a cognitive measure (either from trial-to-trial or from one testing occasion to the other) and does not simply represent random measurement error, this raises serious concerns about the accuracy of approaches that do not take this variability into account. The concept of a "true score" for instance implies a fixed quantity of a given variable that may differ between individuals or over long periods of time, but which is relatively invariant within an individual on a specific

testing occasion (Allen & Yen, 2001). When measuring such a fixed quantity, mean levels of performance may be a perfectly acceptable indicator of underlying cognitive abilities. The presence of substantial within person variability on the other hand raises serious questions about the practicality or tenability of true score values or of classical test theory itself (Lindenberger & von Oertzen, 2006). A true score should not vary from moment to moment if it can be reliably measured by mental tests that may take several hours to complete.

In the neuropsychological context, the existence of substantial within person variability in test performance might undermine the typical assumptions we make about the relationship between test scores and the neurological substrates that underlie an individual's performance (Salthouse & Berish, 2005). If neuropsychological tests correspond to underlying cortical structures that are putatively invariant from one moment to the next, how could we account for substantial variability within one testing occasion if it is not due to measurement error? Though neuromodulatory conditions may change rapidly, these changes are presumably in the service of maximizing functional capacity around some asymptote of optimal performance (hence preserving stability). Substantial within person variability could also result in a number of problems at the pragmatic level. For example, in the presence of substantial variability an individual might be classified as intact on one testing occasion and impaired on another occasion; a serious problem given the substantial legal and personal consequences that often follow a neuropsychological evaluation. If test scores correspond to the presumably static latent traits they purport to measure, then substantial within person variability in test

performance represents a conundrum that deserves explanation (Nesselroade & Salthouse, 2004).

An Alternative to the Measurement of Central Tendency

Recent work suggests that intraindividual variability exists ubiquitously and is quite substantial in magnitude (Nesselroade & Salthouse, 2004). In this study, the authors examined the performance of 204 adults from 20 to 91 years of age on a series of perceptual motor tasks. They took measurements on the same tasks across three occasions within a two-week period. The results indicated that intraindividual variability (as averaged both between sessions and across sessions) accounted for approximately 37% to 53% of the between persons variability. Moreover, those who varied more within a session also varied more between sessions, leading these authors to conclude that variability may be interpreted as a "trait-like" characteristic of an individual. This study is only one example among many consistent with the notion that there is significant intraindividual variability both within sessions and between sessions (Hultsch et al., 2000; Li, Aggen, Nesselroade, & Baltes, 2001).

It is possible that intraindividual variability simply represents random sources of measurement error. If this were the case, we would expect variability from one occasion to be unrelated to that on other occasions (which is definitional of error from a classical test theory point of view; Allen & Yen, 2001). The "trait-like" nature of the data presented above suggests that this is not the case. Other studies have also indicated that intraindividual variability can be measured reliably over long periods of time (Hertzog et al., 1992). And that intraindividual variability within a measurement occasion is

comparable to and correlates highly with that between occasions (Rabbitt et al., 2001). In other words, measures of dispersion correlate highly with measures of consistency.

The possibility that within person variability simply represents random error would also preclude it from having meaningful relationships with criterion variables; however, the literature suggests that this is also untenable. Intraindividual variability demonstrates significant correspondence with many variables of interest including parameters of physical health (Wu et al., 2011), information processing robustness (Li et al., 2004), and cognitive performance (Lövdén et al., 2007). Response time variability has also been associated with a variety of pathological conditions including disease (Bruce, Bruce, & Arnett, 2010) and traumatic brain injury (Stuss et al., 1994).

It has been argued that the relationship of variability parameters such as the iSD or intraindividual coefficient of variation (ICV) to meaningful outcomes is simply an artifact resulting from their high correlation with average (i.e., mean level) performance (Christensen et al., 2005; Salthouse & Berish, 2005). In numerous investigations though, powerful relationships have been found even after partialing out such effects (Dixon et al., 2007; Hultsch et al., 2000; MacDonald, Hultsch, & Dixon, 2003). It should be noted that the statistical methodology used by some of these authors to disentangle variability from mean levels of performance has been criticized by others (Salthouse & Berish, 2005), and may be inadequate in certain circumstances (Schmiedek et al., 2009); nevertheless, such debates do not warrant automatic dismissal of these findings. Other studies have also found meaningful relationships between intraindividual variability and important outcomes when statistically controlling for mean levels of performance. Perhaps the most striking example comes from the work of Macdonald, Hultsch, and

Dixon (2008) who found that intraindividual response time variability predicts impending death as many as 15 years in advance independent of cardiovascular health, demographic indicators, level of cognitive performance, and mean response time. Intraindividual variability also precedes and predicts impending cognitive decline before it is evident in mean levels of task performance (Lövdén et al., 2007). Several studies examining the factor structures of experimental tasks measuring response time latency vs. response time variability have also documented a divergence between these two outcome measures in different clinical populations (Kelly, 2000; Kremen, Seidman, Faraone, Pepple, & Tsuang, 1992; Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991).

With the above caveats, several conclusions can be drawn from the present review. First, intraindividual variability in response time on various types of cognitive tasks is sizable and can be reliably measured. Second, intraindividual variability seems to follow a characterological or trait-like association, such that those who are more variable on one occasion continue to be more variable on other occasions. Third, the degree of variation across trials (dispersion) is comparable to and related with that across occasions (consistency). Fourth, intraindividual variability cannot be dismissed simply as a statistical artifact and appears uniquely predictive of important outcomes including death. Having established intraindividual variability in response time tasks as a valid, reliable, and potentially unique measure, the next question becomes one of significance: How has the study of intraindividual variability contributed to our understanding of neuropsychological phenomena and clinical conditions?

Intraindividual Variability In Clinical Conditions

Hundreds of papers in the neuropsychology and cognitive science literatures have investigated the relationship of intraindividual variability in response time to a diverse array of phenomena including intelligence (Deary, Der, & Ford, 2001), semantic priming (Balota, Yap, Cortese, & Watson, 2008), neuropsychological test performance (Heathcote et al., 1991), aging (Bielak et al., 2010; Lövdén et al., 2007; MacDonald, Nyberg, Sandblom, Fischer, & Bäckman, 2008), and everyday problem solving (Burton, Strauss, Hultsch, & Hunter, 2009). With the exception of aging (which in itself may be considered a type of central nervous system pathology; Bäckman, Nyberg, Lindenberger, Li, & Farde, 2006), the current effort is restricted to patterns of intraindividual variability in response time tasks as they relate to various pathologies of the central nervous system. In that vein, this section presents studies from research on aging, mild cognitive impairment, dementia, traumatic brain injury, schizophrenia, attention deficit hyperactivity disorder, and general psychiatric disturbance as they are informative to the study of intraindividual variability in general.

Across the life span, intraindividual variability in response time tasks forms an approximately U shaped function (Williams, Hultsch, Strauss, Hunter, & Tannock, 2005) that parallels neuromaturation and development (particularly that of the frontal lobes; Macdonald et al., 2006). Increasing variability also parallels concomitant decrements in cognitive functioning throughout senescence (Li et al., 2004). This finding suggests that intraindividual variability in certain types of response time tasks is indicative of compromised neurologic integrity (Li & Lindenberger, 1999), and studies have found that intraindividual variability is associated with specific and general neurologic

compromise (Burton et al., 2006). Researchers in the aging literature have indexed intraindividual variability in response time using a variety of tasks, including reaction time measures of varying complexity (Hultsch et al., 2000), episodic recognition tasks (MacDonald et al., 2008), and indicators of postural sway (Li et al., 2004).

Intraindividual variability on all such measures (either from trial-to-trial or across occasions) seems to be negatively correlated with cognitive performance on a multitude of tasks, especially putative measures of executive functioning (de Frias et al., 2007). This is consistent with other evidence revealing that intraindividual variability is most pronounced on tasks with significant executive task demands (West et al., 2002). In studies examining individuals across time periods of up to three years, when variability is high, cognitive performance is correspondingly low (Bielak et al., 2010). In this study, the authors interpreted this as "evidence for within-person coupling" of cognitive and motor domains; the coupling relationship seems to be stronger for fluid cognitive domains such as processing speed, memory, and reasoning (Bielak et al., 2010). Evidence from this literature supports the notion that intraindividual variability in response time is an ecologically valid indicator of functioning in domains such as problem solving, and bears a relationship to this neuropsychological skill above and beyond the mean (Burton et al., 2009). In older individuals, intraindividual variability is also a reliable harbinger of risk of mortality from all causes (Shipley et al., 2006).

The incidence of mild cognitive impairment is generally thought to bear a positive relationship with advancing age (Panza et al., 2005). Inconsistency in response time distinguishes between cognitively impaired, non-demented adults and control subjects, and is also predictive of those who will go on to develop mild cognitive impairment over

periods of up to five years (Bielak et al., 2010). It should be noted that in this study, mean rate of responding was a comparable predictor of mild cognitive impairment. Other large-sample studies have found similar results (i.e., that variability in responding does not uniquely predict cognitive outcomes over and above mean levels of performance; Christensen et al., 2005), possibly due to statistical and methodological differences. For example, in elderly individuals, measures of consistency may be more predictive of diagnostic status than measures of dispersion, and the relationship of variability to cognitive impairment may be more robust in older vs. younger individuals (Christensen et al., 2005). This finding is still a matter of debate, and other studies have found unique contributions of variability vs. mean levels of responding on measures of simple reaction time (Dixon et al., 2007). In older individuals (70-102 years), inconsistency on measures of perceptual speed precedes and predicts decline in mean levels of performance on the same measures, as well as measures of ideational fluency (Lövdén et al., 2007).

The rate of conversion from mild cognitive impairment to dementia varies across studies but has been documented around 5% - 10% in community samples and 10% - 15% in clinical samples over a one-year period (Christensen et al., 2005). Work within this area has further clarified the relationship of response time variability to neurologic dysfunction. In a study comparing healthy adults, adults with Parkinson's disease, and adults with Alzheimer's disease, both clinical groups showed greater variability in response time than the control group on all measures across four testing sessions (Burton et al., 2006). In this effort, the authors used a variety of different response time tasks including simple reaction time, choice reaction time, word recognition reaction time, and story recognition reaction time. Their findings lend support to the idea that

intraindividual variability is related to general nervous system dysfunction. A dose-response relationship also emerged such that those who were more impaired (as assessed by Mini Mental Status Exam and Intelligence Quotient discrepancy scores) were also more inconsistent across all tasks. A specific pattern of variability between clinical groups was also apparent, such that those with Alzheimer's disease were more variable than those with Parkinson's disease (especially on more complex tasks), an effect that remained after controlling for group differences in severity of impairment. This lends support to the idea that inconsistency is also related to specific patterns of neurologic dysfunction.

Research on intraindividual variability has also deepened our understanding of various primary progressive dementing conditions. As mentioned previously, the magnitude of variability seems to parallel the degree of frontal involvement in dementia subtypes; those affected by a frontal variant of frontotemporal dementia demonstrate greater fluctuation over time and greater between subject variability than those with Alzheimer's dementia or healthy controls (Murtha et al., 2002). In this study, frontal involvement was not associated with greater variability from trial-to-trial (as opposed to session-to-session), although the conclusions drawn from this work must be tempered due to the relatively non-executively demanding nature of the tasks these authors used as measures of dispersion (i.e., simple and choice reaction time) and the small sample size they studied.

In addition to disease processes such as dementia, intraindividual variability also accompanies insult to the central nervous system due to closed head injury (Stuss et al., 1994). While these authors reported no obvious factor (e.g., severity of injury) associated

with the degree of variability, other studies have found dose-response relationships between intraindividual variability in reaction time and severity of impairment (as measured by the Halstead - Reitan Neuropsychological Battery impairment index; Collins & Long, 1996). Recent efforts have clarified these conflicting findings, and trial-to-trial variability seems to be specifically associated with damage to the frontal lobes (with the exception of ventromedial/orbitofrontal regions), but not general brain damage (Stuss et al., 2003). This finding lead Stuss and colleagues (2003) to conclude that, "the frontal lobes control intraindividual performance variability." Mean levels of task performance were not significantly correlated with measures of trial-to-trial variability in this study, however, the sample studied represents a restricted clinical group and this relationship cannot be assumed in other populations. The pattern of variability observed in brain injured subjects bears a complex relationship with task characteristics (Stuss et al., 2003), and level of recovery (Collins & Long, 1996).

Intraindividual variability in response time has been extensively investigated in schizophrenia, and classic descriptions of this condition emphasize irregularities and inconsistencies in performance (Kraepelin, 1919). Investigators in this area have studied dispersion in response times primarily using continuous performance tests (Kaiser et al., 2008) and Go/No-go tasks (Rentrop et al., 2010). In this population, simple measures of variation such as the iSD and ICV are increased in magnitude relative to control subjects and also bear a relationship to task accuracy, although the incremental validity of such measures over that of mean response times is somewhat more equivocal (Kaiser et al., 2008). More complex metrics of variability derived from ex-Gaussian response time distributional analysis (which is discussed extensively in the following sections) are more

promising. Distribution parameters have been found to differentiate high functioning schizophrenic subjects from normal control with high effect sizes, and this relationship is independent of mean levels of performance (Rentrop et al., 2010). In this study, a measure of work capability (the Osnabruck Work Capabilities Profile) was also positively correlated with increased size of the slow tail of schizophrenic participants' response time distributions, giving more support to the idea that indices of variability may be ecologically valid measures.

Schizophrenia has been associated with pronounced dysfunction of the prefrontal cortex (Callicott et al., 2003). There is also evidence that the disorder likely involves widespread cortical disconnectivity, not just regional impairment (Andreasen, Paradiso, & O'Leary, 1998; Friston & Frith, 1995). In either case, high levels of intraindividual variability in schizophrenia have been interpreted as the result of increased prefrontal noise due to inefficient neural processing; a conclusion bolstered by both electrophysiological and functional imaging data (Winterer & Weinberger, 2004; Winterer et al., 2004). This increase in prefrontal noise may represent an "intermediate phenotype" between schizophrenics and those who are not symptomatic but have a high genetic liability for the disorder (Winterer et al., 2004). More research on this topic is required before reaching more than a tentative conclusion, as other studies have not documented increased intraindividual variability in the relatives of schizophrenic individuals (Birkett et al., 2007).

Intraindividual variability in response time has been researched in numerous other clinical conditions including attention deficit hyperactivity disorder, autism spectrum disorders, and general psychiatric disturbance (Bunce et al., 2008; Geurts et al., 2008;

Leth-Steensen et al., 2000). All of these clinical groups demonstrate greater response time variability than normal controls. Moreover, response-time distributional analyses have shown that certain parameters indexing intraindividual variability are highly diagnostic of attention deficit hyperactivity disorder (at least in males with this disorder; Leth-Steensen et al., 2000). Variability within these clinical conditions has been interpreted as a result of impairment in various executive processes such as inhibitory control (Geurts et al., 2008), monitoring and updating information, and mental set switching (Bunce et al., 2008).

The Measurement of Intraindividual RT Variability

Traditionally, researchers have relied on summary statistics to capture the effects of experimental manipulations on response time distributions or to quantify such distributions in general. The iSD is by far the most popular summary measure of variability (Ram & Gerstorf, 2009) and has been used in most of the studies reported above. Despite popularity of this statistic, there is a growing body of evidence that it may be an inadequate measure. It is well known that increased iSDs on response time tasks may simply reflect changes in mean response time: those who have longer response latencies also show more variability in performance at both the intraindividual and interindividual levels (Hale, Myerson, Smith, & Poon, 1988; Salthouse, 1993). Other researchers have not found significant relationships between variability and mean levels of performance (Stuss et al., 2003; West et al., 2002), yet the vast majority of studies detect a high degree of covariation. Derived from the standard deviation, the iSD is prone to the "lawful" relationship between the mean and standard deviation of response time distributions (Wagenmakers & Brown, 2007).

To deal with the iSD's association with mean response time latency, researchers have turned to several different approaches. One method is to index variability as the ICV, which is nothing more than an individual's iSD divided by their mean response time. Accordingly, if an individual's iSD increases simply as a result of corresponding increases in mean latency, the ICV will remain constant. Mathematically, this is a form of data transformation that is roughly equivalent to a logarithmic transformation (Murtha et al., 2002). Some authors have called for variability researchers to cite both these statistics to clarify their results (Kaiser et al., 2008; Stuss et al., 2003) and in some studies, effects that are found when indexed by the iSD have been rendered no longer significant in analyses using the ICV (e.g., Birkett et al., 2007).

A somewhat more sophisticated method of statistical control involves using regression analyses to partial out the effects of mean response time before computing the index of variability (Williams et al., 2005). In addition to removing linear covariates of the mean, this approach also allows for the control of other confounds such as practice, accuracy, and time of day (Allaire & Marsiske, 2005). This method of statistical control has been criticized because it does not control variance in the mean when examining relationships with the iSD; rather, it controls the variance in variables that are correlated with the scores from which the mean is derived (Salthouse & Berish, 2005). In other words, it controls for the association of the mean with other variables but it does not control for the relationship between the mean and the iSD directly.

Furthermore, the use of the ICV and regression-based approaches relies on a potentially flawed assumption: that of a linear relationship between an individual's mean reaction time and the index of variability. While some work has supported the notion of a

linear relationship between these variables (Wagenmakers & Brown, 2007), other efforts have found that this relationship can be violated in empirically obtained samples (Schmiedek et al., 2009). In the latter study, Schmiedek and colleagues (2009) used variance heterogeneity multilevel models to represent reaction time data, finding that the relationship between means and variances itself varies reliably - both across individuals, and across age groups. This obviously contradicts the assumption of invariant linearity presupposed by the use of the ICV or linear regression based statistical control approaches. If the relationship between means and measures of variability is non-linear, it cannot be adequately controlled by statistical approaches based on the general linear model without significant modifications.

A second criticism of summary statistics such as the mean, iSD, or ICV (regardless of what has been partialed out of the data before the computation of these measures) is more fundamental and regards the suitability of these indices for describing the shape of a distribution of reaction time data. Anyone who has read an introductory statistics textbook is well aware of the fact that certain distributions (e.g., the normal or Gaussian distribution) are described completely by the mean and standard deviation while others are not. Positively skewed distributions (or any non-symmetrical distribution for that matter) are not adequately represented by statistics such as the standard deviation, as this measurement does not index the inherent asymmetry of the data (Balota & Yap, 2011). Substantial evidence exists that response time distributions are nearly universally positively skewed and that this positive skew generally increases with task complexity (Luce, 1991); these observations have been so consistent across different tasks and subjects that some authors have likened them to the status of "laws" of behavior

(Wagenmakers & Brown, 2007). Mathematically, distributions with a definite lower boundary and an unlimited upper boundary (e.g., response time distributions) are nearly always positively skewed.

One reason that researchers have relied so heavily on the use of the iSD and similar statistics despite these problems is that the study of perceptual-motor processes (such as those involved in response time tasks) has been largely dominated by information theory (Shannon, 1948). According to this viewpoint, both errors and variability emanate from a motor command (the signal) that is contaminated by noise from the channel through which it is transmitted (Slifkin & Newell, 1998). The amount of variability as measured by the standard deviation of the distribution of an observed variable (e.g., a response time distribution) is assumed to directly correspond to the amount of noise in the underlying behavioral process (Fitts, 1954). A related assumption of information theory is that observations that deviate from the mean of a signal-distribution represent random errors, and that indices of the magnitude of this variability directly correspond to the amount of white Gaussian noise in the signal (Slifkin & Newell, 1998). This assumption has rarely been questioned, but recent work suggests that it may be the exception rather than the rule. Emerging evidence suggests that variability may have meaningful patterns that do not simply represent noise, but that the pattern of variability is important information in and of itself (Balota & Yap, 2011; Leth-Steensen et al., 2000; Rentrop et al., 2010; Slifkin & Newell, 1998).

The substantial positive skew of response time data suggests that such distributions are not adequately represented by a Gaussian function, the iSD, or ICV, and by extension, that the noise itself is not white (i.e., “errors” are not completely

independent of each other/uncorrelated). These problems have lead to a variety of more sophisticated alternative methods for the analysis of response time data. One approach involves the use of computationally explicit models that simultaneously estimate measures of central tendency and variability, the most popular of which is the diffusion model (Ratcliff, 1978). The basic form of this model has been adapted in ways that make it more practically useful given the constraints of real-world experimental situations (see Wagenmakers, vaan der Maas, & Grasman, 2007), yet many paradigms do not contain sufficient data for the adequate application of this methodology. It is also only appropriate on dichotomous response time tasks and cannot be used to model more complicated task paradigms.

Another technique involves the fitting of mathematical functions to response time distributions in order to more precisely quantify their characteristics without losing valuable data (van Zandt, 2000). Among the various distributions that have been utilized, the ex-Gaussian fits response time data "surprisingly well" (Ratcliff & Murdock, 1976). It adequately describes the entire shape of the data in a way that measures of central tendency cannot (Balota & Spieler, 1999). Promising new techniques (i.e., quantile maximum likelihood estimation) have recently become available to extract reliable distribution parameters from small samples of observations, and this process is especially robust and efficient for the ex-Gaussian distribution (Heathcote, Brown, & Mewhort, 2002). Response time distributional analysis makes it possible to analyze specific task manipulations or conditions on standard cognitive paradigms with a level of precision that was not possible a decade ago. Recent efforts have begun to link ex-Gaussian parameters with neuropsychological constructs germane to the present study (e.g.,

working memory and abstract reasoning), aiding the interpretation of these parameters in various applied situations (Schmiedek, Oberauer, Willhelm, Süß, & Wittman, 2007; Tse, Balota, Yap, Duchek, & McCabe, 2010). This methodology has been used to shed light on a variety of neuropsychological and cognitive phenomena, establishing its clinical as well as theoretical significance (Balota et al., 2008; Geurts et al., 2008; Heathcote et al., 1991; Leth-Steensen et al., 2000; Rentrop et al., 2010).

At its most basic level, the ex-Gaussian function represents the convolution of a normal (Gaussian) and an exponential distribution. This function is characterized by two parameters Mu and Sigma (μ and σ, respectively) corresponding to the mean and standard deviation of the normal part of the distribution and Tau (τ), which reflects the mean and standard deviation of the exponential part. The mean of an empirically obtained distribution is constrained to be the sum of μ and τ, while the variance is equal to $\sigma^2 + \tau^2$. As the third moment of the distribution, the skewness is represented by $2\tau^3$, although τ is a more reliable measure of skewness than is the third central moment of the distribution (Heathcote, Brown, & Cousineau, 2004). It is this parameter that has been used effectively to distinguish between individuals with attention deficit hyperactivity disorder and control subjects with impressive effect sizes (Leth-Steensen et al., 2000), clarify the mechanisms underlying Stroop test performance (Heathcote et al., 1991), and differentiate between schizophrenics and control subjects (Rentrop et al., 2010). As previously mentioned, the Tau parameter was also significantly negatively correlated with a measure of work performance in the latter study, giving preliminary support for the idea that this component of the ex-Gaussian distribution may be an ecologically valid measure of functioning. The interpretation of these parameters is not entirely

straightforward, yet there is widespread consensus on the appropriateness of this statistical technique for characterizing response time distributions (Matzke & Wagenmakers, 2009).

Schmiedek et al. (2007) contributed greatly to the literature describing the interpretation of ex-Gaussian parameters. Using structural equation modeling, they were able to show that a latent Tau parameter (defined by reaction time tasks of varying complexity) was a strong predictor of key executive abilities such as working memory and reasoning. These authors also showed that this parameter is directly related to the drift rate parameter of Ratcliff and Murdock's (1976) diffusion model, which symbolizes the efficiency of the information accumulation process and individual differences in processing efficiency (Schmiedek et al., 2007). This description is highly consistent with Li and colleagues (2004) definition of intraindividual performance variability as an indicator of "processing robustness." Their results, which were obtained in a sample of normal controls, have been found to generalize to clinical populations, including healthy older adults and those with mild dementia of the Alzheimer's type (Tse, Balota, Yap, Duchek, & McCabe, 2010).

Descriptive plots are often employed in conjunction with the parametric approaches described above. These graphical analyses represent an important complement to distributional analyses, as they directly display the changes in response time distributions that are being modeled. A distinct advantage to these methods is that they make no underlying assumptions about the theoretical distribution representing the data or the cognitive processes from which data are derived; they are purely descriptive. The results of these graphical analyses should confirm the results of distributional

analyses. Relevant techniques include delta, hazard, quantile, and Vincentile plots (Balota & Yap, 2011). Quantile plots are considered below in some detail, given that they are utilized in the present study.

Quantile plots visually depict differences between the response time distributions of experimental groups; for example, old vs. young participants, cognitively impaired individuals vs. normal controls, or individuals with high vs. low working memory. These plots may be understood to represent the average response time distribution of participants in a given group (i.e., a "super subject;" Tse et al., 2010). To form quantile plots, the RTs of a given participant are first rank ordered from lowest to highest. Quantiles (i.e., points that divide an individual's response time distribution into a pre-specified number of bins each containing an equal proportion of observations) are then calculated for each participant, and these points are averaged across subjects within a particular condition or group (Heathcote, 2000). For example, if 5 quantiles are desired, then one calculates points below which 10%, 30%, 50%, 70%, and 90% of an individuals observations fall. The intervals between contiguous quantiles each contain $1/q$ observations (where q is the number of quantiles). In the example above, each interval would contain 20% of a given individual's observed response times (i.e., $1/q = 0.2$). The number of quantiles should represent a balance between maximizing the resolution of analyses (i.e., a greater number of quantiles) and reliability in the presence of outlying observations (i.e., a fewer number of quantiles). As a heuristic, researchers often use 10 quantiles (e.g., Balota et al., 2008; Tse et al., 2010).

Figures 2 - 4 (adapted from Balota & Yap, 2011) represent idealized effects of changes in specific parameters of the ex-Gaussian distribution as they appear in quantile

plots. In Figure 2, all parameters are held constant with the exception of a 100msec increase in Mu, which results in separation of the quantile bands. In terms of a probability density function, this effect would result in simple shifting of the entire distribution without any changes in the shape of the distribution.

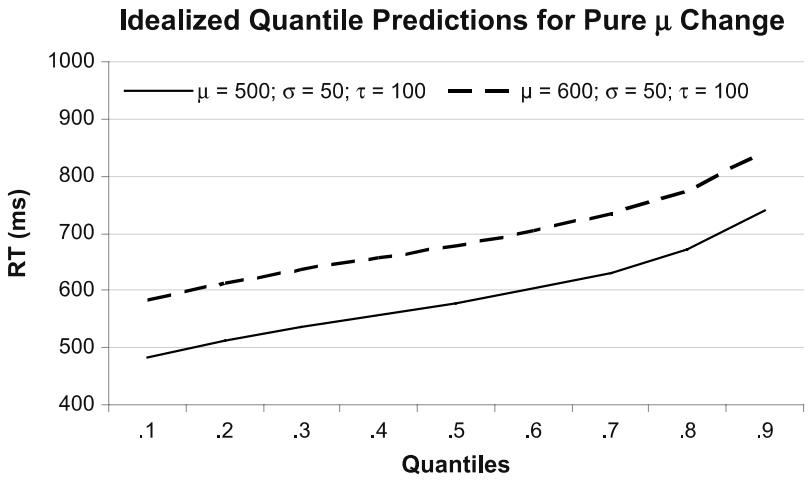

Figure 2. Effects of 100msec change in Mu (Adapted from Balota & Yap, 2011, p. 163).

Figure 3 represents the effect of a 100msec change in Tau with all other parameters held constant. Notice the increasing effect of changes in Tau on the slow tail of the distribution. In terms of a probability density function, an isolated effect in Tau would result in elongation of the right side and right tale of the distribution (similar to increased skewing).

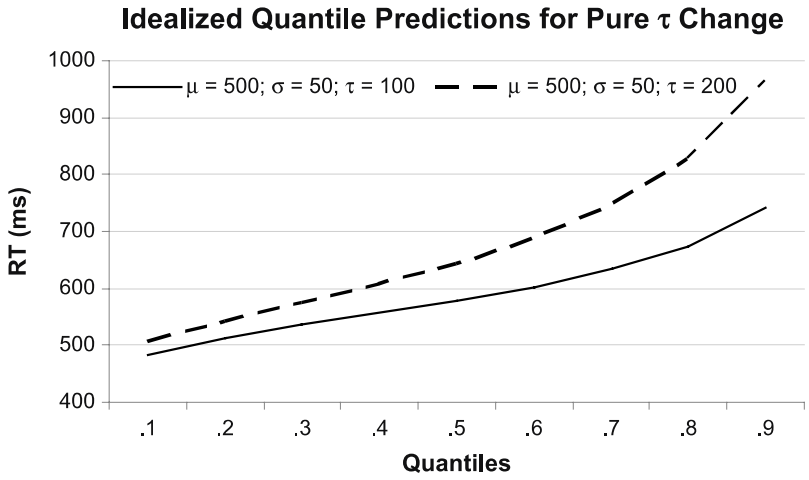

Figure 3. Effects of 100msec change in Tau (adapted from Balota & Yap, 2011, p. 163).

Finally, Figure 4 represents the idealized effect of a 100msec change in both Mu and Tau, resulting in separation of quantile bands in addition to increasing slope at more extreme observations for the line with the higher Tau parameter. In terms of a probability density function, changes in both Mu and Tau would result in distributional shifting as well as a change in the right side (i.e., exponential component) of the distribution.

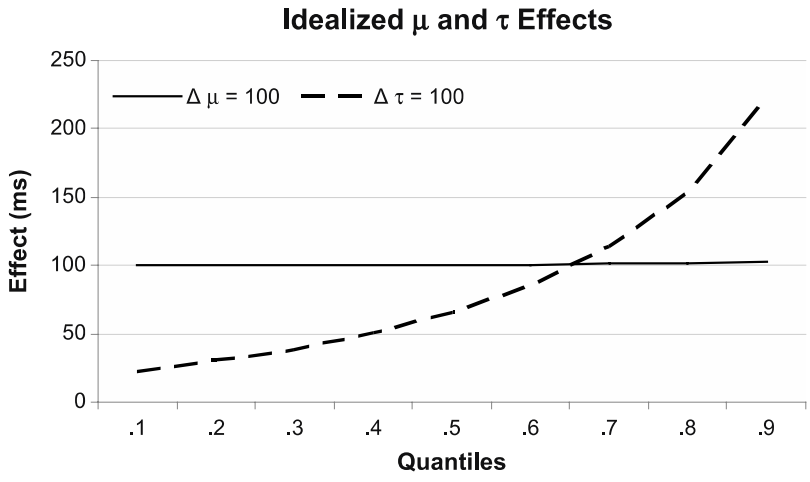

Figure 4. Effects of 100msec change in both Mu and Tau (Adapted from Balota & Yap, 2011, p. 163).

Estimation of Distribution Parameters

Until recently, continuous maximum likelihood (CML) estimation was the method of choice for parameter estimation in response time distributional research (van Zandt, 2000). This procedure requires approximately 100 observed response times in order to reliably extract distribution parameters. Heathcote and collages (2002) recently proposed quantile maximum likelihood (QML) as an alternative that requires far fewer observations for reliable parameter extraction. QML is a unique estimation method that extracts parameters on the basis of quantiles rather than raw data. To clarify, the likelihood of a sample is defined as the joint probability of the sample of observations according to the theoretical distribution model employed (Heathcote et al., 2004). The likelihood differs depending on the assumptions underlying the estimation method. CML estimation for the ex-Gaussian distribution assumes that the data are independent, identically distributed ex-Gaussian observations, and the likelihood for the data are described by the following equation (adapted from Brown & Heathcote, 2003):

$$L(x_1,x_2,\ldots,x_n \mid \mu,\sigma,\tau) \propto \prod_{i=1}^{N} \exg(x_i,\mu,\sigma,\tau).$$

QML nests the model assumed by CML, yet makes weaker assumptions about the model representing the data, allowing it to function in circumstances that would be inappropriate for CML estimation. As described by Brown & Heathcote (2003):

> QML estimation begins with the specification of a fixed set of probabilities, $\{p_i : i = 0...m\}$, where $p_1 = 0, p_i < p_{i+1}$, and $p_m = 1$, such that $p(t < q_i) = p_i$ where the set $\{q_i : i = 0...m\}$ are quantiles. (p. 486).

This method leads to the likelihood function below (adapted from Brown & Heathcote, 2003):

$$L(x_1,x_2,...,x_n \mid \mu,\sigma,\tau) \propto \prod_{i=1}^{m} \left(\int_{q_{i-1}}^{q_i} \exp g(x_i,\mu,\sigma,\tau) dt \right)^{n_i}$$

While more computationally intensive, the advantage to such an approach is increased precision and accuracy, and parameters estimated using this method have been shown to have smaller bias and variance than those estimated using CML (Cousineau, Brown, & Heathcote, 2004; Heathcote et al., 2002). Using QML also enables the researcher to obtain reliable parameter estimates from data sets as small as 40 observations, a serious advantage given the limitations encountered when trying to fit empirically obtained response time data (Brown & Heathcote, 2003). In contrast, using continuous maximum likelihood estimation calls for a minimum of 100 observations for a reliable fit (but see Geurts et al., 2008; Leth-Steensen et al., 2000 for possible exceptions to this heuristic). This fitting technique is implemented in Brown & Heathcote's (2003) Quantile Maximum Probability Estimator (QMPE v2.18) software. QMPE produces estimates of Mu, Sigma, and Tau, as well as estimates of the standard errors of these three parameters, log-likelihood values, and parameter intercorrelations. It also automatically tests model fit by calculating expected values for the quantiles and comparing them with those observed in the data.

The Executive Functions

There is little consensus regarding the term "executive function." As with the concept of intelligence, there is evidence that the executive functions are characterized by

both unity and diversity (Miyake et al., 2000). Early theories of executive functioning implicated the existence of a "central executive" that coordinated a series of component processes (Baddeley, 1986). Cytological architectonic studies at the beginning of the 20th century (e.g., Brodmann, 1909) paved the way for what is now a voluminous and contentious literature that suggests a significant amount of functional specialization within the prefrontal cortex and it's associated networks. The current evidence supports an understanding of the executive functions as "a collection of anatomically and functionally independent but interrelated attentional control processes" (Stuss & Alexander, 2007, p. 901). Theories abound and researchers have advanced numerous models based on neuropsychological and functional neuroanatomical properties (Goldman-Rakic, 1995; Petrides, 1995; Miyake et al., 2000; Stuss & Alexander, 2007; Tekin & Cummings, 2002). The following review presents several related theories of frontal lobe functioning that reflect contemporary neuropsychological, cytoarchitectonic, electrophysiological, and functional neuroimaging work.

At a very general level, the executive functions may be understood as "higher-level cognitive functions involved in the control and regulation of lower cognitive operations" (Stuss & Levine, 2002, p. 416). Skills associated with the ventral prefrontal cortex (VPFC) may be differentiated functionally and anatomically from those subserved by the dorsolateral prefrontal cortex (DLPFC), a distinction supported by the evolutionary theory of cortical architectonics (Pandya & Yeterian, 1996). The DLPFC is characterized by the archicortical trend, which originates in the hippocampus and is involved in spatial and conceptual reasoning processes. By contrast, the VPFC is part of the paleocortical trend, which is extensively interconnected with limbic nuclei involved

in affective processing and emerges from the caudal orbitofrontal cortex (Stuss & Levine, 2002). The executive functions explicitly measured by clinical neuropsychologists are almost exclusively those that are subserved by the dorsolateral prefrontal cortex and the associated structures of the archicortical trend.

There is evidence for a significant amount of functional specificity within the lateral cortex on the basis of functional, cytoarchitectonic, and anatomic distinctions (Owen, 1997). The ventrolateral prefrontal cortex (VLPFC; primarily the inferior frontal gyrus, Brodmann's areas 47, 44, 45) and the DLPFC (primarily the middle frontal gyrus, areas 9 and 46) in particular have been the subject of significant research (D'Esposito, Postle, & Rypma, 2000). In the domain of working memory, this has lead to several models of functional specialization that divide the cortex similarly yet differ with respect to the properties attributed to each region. Goldman-Rakic (1995) concluded that the DLPFC differs from the VLPFC with respect to the type (i.e., modality) of information that is being processed. According to this viewpoint, the DLPFC is responsible for working memory as it relates to spatial information and the VLPFC applies working memory functions to non-spatial information. While considerably more sophisticated, this theory mirrors aspects of classic models (e.g., Baddeley, 1986) that are no longer in favor.

More contemporary efforts have continued to support the notion that these regions are distinct albeit not with respect to the type of information they process. For Petrides (1995), the ventrolateral and dorsolateral areas of the prefrontal cortex are differentiable on a process specific basis. Information is initially received from posterior association areas by the VLPFC, which also performs organizing operations on information as it is

held in working memory. The VLPFC thus represent a first stage of working memory that permits the active organization of responses on the basis of consciously retrieved information from posterior storage systems. The DLPFC is recruited during a second stage of working memory that involves the active manipulation and monitoring of that information (Petrides, 1995). In other words, first stage functions bring information "online" within the VLPFC and second stage functions work with that information in the DLPFC. Contemporary neuroimaging data has largely supported this process-dissociable theory of executive control within the human prefrontal cortex (D'Esposito et al., 2000; Owen, 2000).

Petrides' (1995) hierarchical model parallels more than a century of clinical wisdom, as practicing neuropsychologists frequently conceptualize executive functions existing along a gradient of complexity from simple (e.g., retrieval, sustained attention, simple performance monitoring) to more complex (e.g., abstract reasoning, complex strategy generation, mental manipulation; A. Byrd, personal communication). This hierarchical conceptualization is also reflected in the construction of modern test batteries designed to evaluate the executive functions, which introduce more complex task demands in a staged fashion in order to allow for dissociation of specific neuropsychological processes (e.g., Delis, Kaplan, Kramer, 2001).

Building on the dual stage model, Christoff & Gabrielli (2000) have advocated the inclusion of a third stage of processing subserved by the frontopolar cortex (roughly corresponding to Brodmann's area 10) and a complementary rostro-caudal aspect of prefrontal cortical organization. The frontopolar cortex is differentially recruited on the basis of whether a particular cognitive task requires *internally generated* information or

externally generated information. When information exists (or has recently existed) externally, the DLPFC may be sufficient to evaluate and manipulate that information. By contrast, the frontopolar cortex is activated accompanying the addition of increased demands requiring the evaluation and manipulation of *internally generated* information. This model was derived from current neuroimaging literature that describes the performance of subjects completing episodic memory and abstract reasoning tasks.

These cognitive models share substantial overlap with conceptual frameworks used by practicing clinical neuropsychologists. A neuropsychological evaluation of "frontal lobe" functions typically considers four discrete areas: frontal lobe language functions, memory control functions, working memory functions, and anterior attention functions (Stuss & Levine, 2002), among others (e.g., abstract reasoning, perceptual synthesis/organization, planning)

Neuropsychological assessment of frontal language functions involves an examination of activation vs. formulation (Stuss & Levine, 2002). Formulation may be evaluated qualitatively during the clinical interview, as evidenced by the patient's ability to succinctly, pertinently, and relevantly respond to examiner inquiries. The inclusion of extraneous details, circumstantiality, and simplification of sentence forms suggest the presence of impairment. Requiring the patient to fluently generate a list of words in response to a semantic category (e.g., animals, actions) or a particular letter of the alphabet serves as a measure of activation (e.g., FAS; Strauss, Sherman, & Spreen, 2006). You indicated this is wrong, how should I rewrite it? I'm not sure which part is wrong. I typically think of semantic fluency as retrieval/lexical access and phonemic fluency as

ability to generate a retrieval strategy. Stuss and Levine (2003) list them as measures of "activation."

Executive control of memory involves the strategic coordination, elaboration, and interpretation of associations stored by medial-temporal/hippocampal regions (Stuss & Levine, 2002). Semantic clustering of words within associative learning paradigms (as observed in the California Verbal Learning Test; Delis, Kramer, Kaplan, & Ober, 2000) and placement of information in temporal and spatial contexts are of particular clinical importance (Wheeler, Stuss, Tulving, 1997). Retrieval of information from semantic networks and the application of retrieval strategies may also be assessed via verbal fluency (particularly phonemic fluency; Baldo, Schwartz, Wilkins, & Dronkers, 2006).

From a neuropsychological perspective, working memory involves holding information in abeyance in order to mentally manipulate it. This skill is measured on common batteries of full scale intellectual functioning through tasks such as repeating digits backward, resequencing random digits in ascending order, and completing mental arithmetic problems (Wechsler, 2002). Popular paradigms within the variability literature include reading span, computation span, and rotation span (e.g., Schmiedek et al., 2007; Tse et al., 2010).

Standard neuropsychological assessment of anterior attention functions involves evaluation of attentional switching, selective attention, and sustained attention (Stuss & Levine, 2002). Prototypical tasks used to measure attentional switching include the Wisconsin Card Sorting Task (WCST) and Trail Making Test Part B (TMT-B). This cognitive ability is more commonly referred to as "shifting" within the research literature (Miyake et al., 2000). Sustained attention, on the other hand, is assessed through the use

of continuous performance tests (e.g., Conners' Continuous Performance Test; CPT) or other bland tasks requiring "top down" modulation of basic attentional processes (Lezak, 2004). Measures of selective attention such as the Stroop test require the ability to prioritize particular stimulus characteristics (i.e., color naming) while suppressing other competing information (i.e., word reading) selectively.

Tasks measuring the neuropsychological domains of working memory and anterior attention functions can be broadly construed as tests of cognitive control that depend upon the partially dissociable networks of the DLPFC and VLPFC (Miller & Cohen, 2001; Stuss & Levine, 2002; Stuss, 2011). These tasks are particularly relevant to the study of variability, given that neuroimaging, latent variable, and clinical research suggest that intraindividual response time variability is a sensitive marker of the efficiency of executive control processes (Bellgrove et al., 2004; Bunce et al., 2007; Tse et al., 2010). Neuropsychological measures of cognitive control are thought to draw on multiple processes including task switching, response inhibition, error detection, response conflict, and working memory (Chase, Clark, Sahakian, Bullmore, & Robbins, 2008).

The Stroop test, Wisconsin Card Sorting Test, Trail Making Test Test (Part B), and COWAT represent prototypical cognitive control tasks and have been used in lesion mapping studies to further clarify the brain regions underlying this executive ability (Gläscher et al., 2012). Data from this revolutionary work have largely supported Petrides (1995) dual stage model and lends further evidence to the hierarchical rostro-caudal organization of the prefrontal cortex hypothesized by Christoff & Gabrielli (2000). Gläscher's (2012) work represents the first study of its size causally relating neuropsychological test performance to specific prefrontal cortical networks that

subserve cognitive control. These neuropsychological measures are thus well positioned for investigations of intraindividual response time variability as it relates to this cognitive construct.

Executive Functioning and Intraindividual Variability

A variety of neural substrates thought to subserve the executive functions have been implicated in the regulation of response-time consistency. These include the robustness of neuromodulatory control (Li & Lindenberger, 1999; Li, Lindenberger, & Sikström, 2001), overall white matter volume (Walhovd & Fjell, 2007), and fluctuations in the functional connectivity of various brain regions (Kelly, Uddin, Biswal, Castellanos, & Milham, 2008; Prado, Carp, & Weissman, 2011). At the behavioral level, significant evidence also suggests that executive dysfunction may give rise to increased performance variability. Numerous authors have invoked this supervisory cognitive skill as a framework for the interpretation of their findings (Bunce et al., 2008; Geurts et al., 2008; Stuss et al., 2003).

Considering the multifactorial nature of the executive functions, increased intraindividual variability could be the result of impairment or inefficiency in any one of a number of component processes. Attributions of intraindividual variability to specific executive functions include “lapses of intention” (i.e., prospective memory), fluctuations in executive control (West et al., 2002), increased prefrontal broadband noise due to inefficient neural processing (Winterer et al., 2004), and inadequate sustained attention caused by a failure of top-down executive processes (Kaiser et al., 2008). All of these descriptions are consistent with the interpretation that response time variability may be specifically related to cognitive control.

Evidence from structural and functional brain imaging techniques has given credence to this hypothesis. In one study, frontal white matter hyperintensities in older adults were related to increased intraindividual variability on reaction time tasks, but not on tests of psychomotor speed, memory, and global cognition (Bunce et al., 2007). This suggests that variability in response time is a discrete phenomenon that does not simply reflect general cognitive dysfunction or motor control problems (at least in some populations). In this study, Bunce and colleagues (2007) also found that white matter hyperintensities in non-frontal brain regions were not associated with increased intraindividual variability, strengthening the assumption that consistency may specifically map onto prefrontal areas of the cortex.

The first study of intraindividual variability and the functional neuroanatomy of executive control has helped to elucidate this association (Bellgrove et al., 2004). In this work, increased variability in responding on go trials of a Go/No-go task was associated with increased bilateral frontal brain activation, lower inhibitory success, and slower responding. These results are consistent with the idea that individuals who demonstrate greater variability have an increased need for executive control (resulting in greater activation of frontal/prefrontal circuitry) and that this translates to poor performance on other measures of this ability. The authors concluded that intraindividual variability could serve as a sensitive index of the efficiency of executive control processes (Bellgrove et al., 2004).

Genetics research further implicates the prefrontal cortex as a key determinant of response time variability. The catechol-*O* methyltransferase gene is responsible for dopamine catabolism in the prefrontal cortex through conversion to 3-methoxytyramine

and exists in two broad variants in human beings (Mattay et al., 2003). Val carriers of the gene have lower levels of this important neurotransmitter in prefrontal areas than Met carriers; they also show deficient performance on measures of various executive abilities (Goldberg et al., 2003). Carriers of the Val allele of the catechol-*O*-methyltransferase gene exhibit greater intraindividual variability than Met carriers on continuous performance tasks (Stefanis et al., 2005). They also demonstrate inefficient patterns of prefrontal neural activation on working memory measures, and have an increased incidence of conditions associated with heightened performance variability (e.g., schizophrenia, Egan et al., 2001; and attention deficit hyperactivity disorder, Eisenberg et al., 1999).

Neuropsychological efforts also suggest that intraindividual variability may be related to executive functioning and cognitive control. Greater variability is associated with poorer ideational fluency (Lövdén et al., 2007), weaker reasoning and retrieval (Bielak et al., 2010), and is negatively correlated with various measures of fluid intelligence (Alderton & Larson, 2002; Li et al., 2001; Ram et al., 2005). Fluid intelligence is a broad cognitive ability, yet this construct has been linked to behavioral measures of executive functioning and frontal lobe volume (Schretlen et al., 2000). The fact that variability is increased on moderately to highly executively demanding tasks is also consistent with the hypothesis that executive integrity is necessary for performance consistency (Bielak et al., 2010; West et al., 2002).

Several conclusions can be drawn on the basis of the present review. First, the relationship between the executive functions and performance variability is fairly robust. Second, there is a convergence of evidence from structural and functional neuroimaging,

neuropathological, and neuropsychological efforts supporting the presence of this relationship. Third, work suggests that performance consistency and executive ability may both be the products of genetic and environmental factors. The literature thus begs the question: do measures of intraindividual variability deserve a place in the test batteries of practicing neuropsychologists?

The Wisconsin Card Sorting Test (WCST)

The WCST may be ideally suited as a measure of response time variability. This task was originally construed as a "simple" and "objective" measure of problem solving and decision making in 1948 (Berg, 1948). Since that time, it has been variably conceptualized as a measure of attentional switching (Everett, Lavoie, Gagnon, & Gosselin, 2001), working memory (Goldman-Rakic, 1994), and sustained attention (Smith et al., 1998). Regardless of these various interpretations, the WCST has been a dominant paradigm for the measurement of executive functioning since its inception and over 75% of neuropsychologists report using it as a standard part of their test battery (Butler, Retzlaff, & Vanderploeg, 1991). The WCST has also been a topic of considerable research with over 600 papers published on this instrument, 80% of them within the last 16 years (Greve, Stickle, Love, Bianchini, & Stanford, 2005).

Data from numerous neuroimaging studies support the notion that the WCST taps working memory function in dorsolateral and frontopolar areas of the prefrontal cortex (Berman et al., 1995). The magnitude of activation of the dorsolateral prefrontal cortex is also positively related to levels of task performance on the WCST (Ragland et al., 1997). Stuss and colleagues (2003) found that lesions to this area were particularly important in effecting intraindividual variability on a series of response time tasks. In non-brain

injured individuals however, to effect significant variability in response time researchers have used more executively demanding tasks (West, 1999). A recent meta-analysis of neuroimaging studies on the WCST supported the notion that this instrument is an attentionally demanding executive task (Buchsbaum, Greer, Chang, & Berman, 2005), even after giving participants training practice (Berman et al., 1995).

Factor analytic studies have further characterized the latent cognitive constructs measured by the WCST. Most have found that a three-factor solution best describes the test and that the only statistically sound factor is the first one (Greve et al., 2005). This factor has been observed in studies with frontal patients, schizophrenic patients, and normal controls. It has been variably labeled "perseveration," "abstract thinking ability," "concept formation," "undifferentiated executive function," and "flexibility" (Somsen, Van der Molen, Richard Jennings, & van Beek, 2000).

Attempts to further specify the multitude of cognitive processes involved in this complicated task have resulted in examinations of various stages of task performance. A simple model that has been proposed implicates set-switching and set-maintenance as key components (Kieffaber et al., 2006). These two different aspects of executive functioning generally reflect rule-search and rule-application, two distinct processes that grossly map onto specific task demands (Somsen et al., 2000). After correct feedback and rule attainment, subjects must maintain their mental set and continue to respond correctly until an unannounced change in the sorting rule. Following negative feedback, subjects must disrupt their ongoing pattern of responding and begin searching for a new rule until they have determined the next sorting principle.

The distinction between rule-search and rule-application is an important one. As elaborated by Li and colleagues (2004) in their taxonomy of intraindividual dynamics, certain types of variability can be adaptive. Plasticity (functional diversity to attain large gains in performance following task exposure) and diversity (the set of exploratory behaviors and strategies used for performing a complex cognitive task) are two examples. Such behaviors are typically observed on complex tasks amenable to strategy use, like the WCST. On the other hand, fluctuation (characterized by instability of optimal performance) around an asymptote of optimal functioning represents a maladaptive type of variability. During rule application functioning is already optimal because the sorting principle has been attained, and variability during this phase of the task would represent precisely this sort of fluctuation or lack of processing robustness.

Limitations of our Current Knowledge

The previous review reveals several limitations of the current literature that are particularly germane to neuropsychologists who wish to incorporate intraindividual variability into their clinical practice. First, most of the studies of intraindividual variability reported above have not extensively examined the relationship of this construct with commonly used measures of executive functioning. The majority of efforts have focused on predicting outcomes using intraindividual variability data (e.g., Christensen et al., 2005; Holtzer et al., 2008; Lövden et al., 2007; MacDonald et al., 2008) or using intraindividual variability as a basis for between group comparisons (e.g., Burton et al., 2006; Geurts et al., 2008; Bunce et al., 2008). In those studies that do assess higher-order cognitive constructs directly relevant to the practice of clinical neuropsychology (e.g., de Frias et al., 2007; Schmiedek et al., 2007; Tse et al., 2010), the

batteries employed are frequently esoteric to American neuropsychologists or are insufficiently comprehensive. Failure to include clinical measures of abstract reasoning (a higher-order frontopolar aspect of executive functioning) is particularly common. For example, de Frias and colleagues (2007) operationalized executive functioning using only the Stroop task, Trail Making Test (Part B), and Digit Ordering Test. As higher-order executive skills (e.g., abstract reasoning may be functionally distinct from more basic executive abilities), this is a glaring omission. Moreover, no study has examined the relationship of intraindividual variability in response time to the WCST.

Second, studies that do employ commonly used neuropsychological tasks are often conducted in restricted clinical samples that do not reflect the diversity encountered in routine clinical practice. Different patterns of intraindividual variability exist within different populations; it is thus uncertain whether associations found in one clinical group generalize to others. This is an important limitation given that the appropriate application of psychometric methods requires normative data to generalize. Individuals presenting for neuropsychological evaluations in applied settings often demonstrate multiple comorbid conditions, psychiatric impairment, and are unlikely to be reflected in the well-screened samples employed in most published investigations.

Third, most studies index variability using simple or complex dichotomous response time tasks as criterion variables. Other common paradigms include episodic memory tasks (Hultsch et al., 2000), greater than two choice reaction time tasks (e.g., Schmiedek et al., 2007; Stuss et al., 2003), and Simon tasks (e.g, Tse et al., 2010). These measures are designed explicitly for research and are not commonly employed in routine

batteries used for clinical purposes. Thus, increased intraindividual variability on these tasks is of uncertain clinical application.

Fourth, the majority of neuropsychological research on intraindividual variability and executive functioning has employed relatively crude outcome measures such as the iSD, ICV, or other summary statistics that paint an incomplete picture of the data. As reviewed in detail above, the use of these statistics may be misleading. Results from studies employing these measures are of dubious validity and may simply reflect differences in mean response latency. It is not possible to determine the precise location within the response time distribution from which significant effects emanate using summary statistics, and thus fine-grained analyses are not possible. Finally, these studies rarely report graphical analyses that would allow for more careful inspection of significant results.

A Rationale For the Present Study

On the basis of the present review, it can be concluded that intraindividual response time variability is likely to be a sensitive and novel indicator of the integrity of the prefrontal cortex and the efficiency of cognitive control processes. Furthermore, the relationship between intraindividual variability and the executive functions appears to be robust and holds across studies of traumatic brain injury (Stuss et al., 2003), dementia (Murtha et al., 2002), aging (Lövdén et al., 2007), and psychiatric disturbance (Bunce et al., 2008). In the few studies examining intraindividual response time variability in normal populations, increases are also associated with poorer performance on measures of fluid intelligence (Nesselroade & Salthouse, 2004) and executive constructs such as working memory and reasoning (Schmiedek et al., 2007; Tse et al., 2010). This evidence

is consistent across neuroimaging (Bellgrove et al., 2004), genetic (Bellgrove et al., 2005; Stephanis et al., 2005) and neuropsychological studies (Bielak et al., 2010, Lövden et al., 2007).

In the setting of these findings, an investigation using clinical methods to explore intraindividual variability as it relates to executive functioning may inform the practice of clinical neuropsychology. The primary aim of this study is to evaluate the relationship between executive functioning (specifically, cognitive control) and intraindividual variability in response time. In doing so, this research makes use of Petrides' (1995) dual stage model of working memory and Christoff & Gabrielli's (2000) hierarchical rostro-caudal model of prefrontal cortex functioning as explanatory frameworks. Several research hypotheses are advanced in the service of this goal.

First, it is hypothesized that greater levels of cognitive control will result in lower levels of intraindividual variability. Cognitive control has been previously operationalized in neuropsychological studies through the use of the WCST (perseverative errors), Stroop (color-word trial), TMT-B (seconds to completion), and COWAT (total words generated to the letters FAS; Gläscher et al., 2012). Lower levels of performance on this battery of tests should be associated with greater levels of intraindividual variability (as parameterized by Tau). It is also likely that individuals who take with longer mean response latencies (as parameterized by Mu) will also exhibit poorer test performance.

Second, greater levels of simple sustained attention will be associated with lower levels of intraindividual variability. If intraindividual variability in response time is related to higher-order aspects of executive functioning, according to Petrides' (1995)

dual stage model, it should also be related to first stage processes. First stage processing involves comparatively simpler cognitive skills such as “comparisons between, or judgments about the occurrence or non-occurrence of remembered stimuli” (Owen, 2000, p. 34). This characterization describes the CPT, and this hypothesis will be investigated through examining the relationship of the ex-Gaussian parameters with various CPT variables thought to reflect first stage attentional processes. The Tau parameter (and perhaps Mu) should increase with lower levels of performance on the CPT.

Third, it is important to test the differential contributions of first stage vs. second stage functions to the prediction of intraindividual variability. Contemporary theories of frontal lobe functioning suggest that this region of the brain may be understood as an interconnected network of cytoarchitectonic regions that subserve partially differentiable cognitive functions (Christoff & Gabrielli, 2000; Petrides, 1995). From this perspective, basic attentional processes (i.e., first stage processes) such as maintaining an internal goal state and holding information in working memory are required in order to perform more complex operations such as mentally manipulating that information, monitoring, or shifting between different task sets (i.e., second stage processes). Hypothesis three tests whether individual differences in basic first stage attentional processes are sufficient to account for any observed relationships between intraindividual variability and cognitive control. Specifically, it is hypothesized that the relationship between cognitive control and Tau will be partially mediated by basic attentional functions.

The fourth and final hypothesis is derived from evidence that intraindividual variability is related to abstract reasoning (Schmiedek et al., 2007); a putative higher-order frontopolar executive function associated with highly abstract activities (Christoff

& Gabrielli, 2000). Specifically, it is hypothesized that intraindividual variability in response time will be associated with abstract non-verbal reasoning as operationalized by the number of categories completed on the WCST. To test this hypothesis, a series of regression equations will be used to evaluate the relationship between categories completed on the WCST and the three ex-Gaussian parameters. Tau (and perhaps Mu) should be positively associated with the number of card sort categories completed. A series of task specific analyses follow these hypotheses in order to clarify study results and determine the plausibility of alternative explanations.

This investigation seeks to address several of the limitations of the literature that are reviewed above. With respect to the first limitation, this study employs a comprehensive battery of executive function tasks that are used by neuropsychologists in routine clinical practice. Indices of intraindividual variability in response time will be compared to cognitive control tasks including TMT-B, COWAT, the Stroop, and indices from the WCST. The results of a large-scale lesion mapping study (Gläscher et al., 2012) provide causal evidence of the brain regions that subserve these cognitive tests, allowing for relatively confident interpretation of the results within a neuroanatomical framework. This work is also informed by several decades of neuroimaging and traditional lesion studies that further support the attribution of specific task demands to functional areas of the prefrontal cortex.

The design of the study allows direct tests of the hypothesis that first stage processes may mediate the relationship between cognitive control, abstract reasoning, and intraindividual variability. All of the cognitive control measures can be broadly construed as “second stage” tasks within dominant models of working memory and prefrontal

cortex function (D'Esposito et al, 2000; Petrides, 1995; Owen, 2000). Including the CPT incorporates simpler "first stage" indices that permit a hierarchical mediation analysis of intraindividual variability. Variables from the WCST that measure abstract reasoning (i.e., Categories) reflect the "third stage" postulated by Christoff & Gabrielli (2000), permitting tentative conclusions about the relationship of intraindividual variability to the rostro-caudal organization of the prefrontal cortex. These measures are of known clinical utility, have been extensively researched, and their interpretation is relatively clear, which make the findings of this study highly relevant to the practice of clinical neuropsychology.

Regarding the second limitation of the literature, the participants of this study are drawn from the general population of outpatients at a major northeastern hospital. This sample represents intraindividual variability as it exists clinically. The exclusion criteria are minimal in order to best capture variability phenomenon as they occur in routine practice. If intraindividual variability does bear a meaningful relationship to various executive functions, this relationship can be confidently generalized across a broad range of different syndromes and comorbidities on the basis of this study.

With respect to the third limitation of the literature, the present study indexes variability using the Wisconsin Card Sorting Test, the most commonly employed measure of executive functioning in the neuropsychological armamentarium (Stuss & Levine, 2002). This task is well suited as a variability measure given that inconsistency is pronounced on executively challenging tasks in most populations (MacDonald et al., 2009; West, 1999, West et al., 2002). There is also evidence that the performance of this task activates prefrontal circuitry; specifically, the frontopolar and DLPFC (Berman et

al., 1995; Christoff & Gabrielli, 2000; Gläscher et al., 2012). Performance consistency depends upon intact dorsolateral prefrontal cortical regions (Bunce et al., 2008; Bellgrove et al., 2004; Stuss, 2003), providing preliminary evidence that the WCST may serve as an ideal variability model with important experimental and clinical implications.

In previous efforts, there has been concern about the extent to which response time variability reflects degeneration of peripheral motor processes vs. more central cognitive abilities (Hultsch et al., 2000). The WCST minimizes this concern because the motor processes involved in this task have been dissociated from its more centrally demanding executive components (Marenco, Coppola, Daniel, Zigun, & Weinberger, 1993). Furthermore, performance of the WCST is not primarily dependent upon motor functioning. In contrast to simple motor tasks, the WCST predominantly activates the right anterior dorsolateral prefrontal cortex; thus, the variability observed cannot simply be interpreted as instability in perceptual-motor performance.

With respect to the fourth limitation revealed by the literature review, the present study attempts to overcome many of the statistical shortcomings of previous neuropsychological research on intraindividual variability. Employing a variety of response-time distributional techniques to index and graphically represent variability avoids the pitfalls inherent in summary statistics (such as the iSD or ICV). Previous similar efforts have documented strong correlations between various ex-Gaussian distribution parameters and measures of working memory and reasoning (Tse et al., 2010; Schmiedek et al., 2007), two abilities tapped by the WCST.

In sum, four specific hypotheses are tested in this investigation. The primary hypothesis is that greater levels of intraindividual variability will be associated with

lower levels of performance on the cognitive control tests. Second, greater levels of simple attention will accompany lower levels of performance on the CPT (a test of basic attentional abilities). Third, the relationship between intraindividual variability and cognitive control will remain significant after controlling for basic attentional functioning. Fourth, intraindividual variability will be related to abstract non-verbal problem solving.

CHAPTER 3

METHODS

Participants

Participants were drawn from the general population of outpatients presenting for neuropsychological evaluations at a large urban general hospital in western Pennsylvania. Archival records available within the section of neuropsychology were reviewed to determine whether participants met selection criteria, and all suitable protocols were considered for inclusion. Eligible participants were between 18 and 65 years of age. This particular age range was chosen to minimize age related variability, as the focus of this work is restricted to intraindividual variability caused by various forms of neuropathology. Evidence suggests that response time variability follows a U-shaped pattern in relation to age, with more pronounced variability towards the beginning and end of life (Williams et al., 2005). In addition to participants outside this age range, individuals with a diagnosis of intellectual disability were also barred from inclusion. Electronic WCST protocols were then reviewed to determine whether they met the following criteria: no more than 30% “other” responses (which might indicate poor task engagement or failure to follow instructions; Somsen et al., 2000) and a minimum of 30 correct responses (necessary for reliable extraction of distribution parameters; Brown & Heathcote, 2003).

A detailed record review was conducted for candidatess who met these criteria, in order to determine whether they had also completed the Stroop color-word test, COWAT, CPT, and TMT A & B. The results of this record review generated an initial sample of 191 individuals. Visual inspection of WCST protocols suggested that a cutoff of 30%

"other" responses might not be sufficient to eliminate individuals with suspect levels of task engagement. More stringent criteria were applied by converting the number of "other" responses to *z*-scores and screening them using a bonferroni adjusted decision criterion of a =0.01 (z = 3.883). This resulted in the identification of 3 subjects with a number of "other" responses suggesting marked impairment, confusion, or poor task engagement. These records were subsequently omitted from analyses. An additional 4 participants were dropped due to administrative errors during data collection. The final sample included 184 individual participants. Demographic characteristics and psychometric test performance of participants appears below in Table 1.

Table 1

Demographic Characteristics and Psychometric Test Performance Raw Scores

Variable	*n*	Mean	SD	Range
Age	184	42.26	13.84	18 -64
Education	184	13.81	2.52	7 - 24
CPT[a]				
Omissions	183	7.05	11.56	0 - 66
Standard Error	183	7.38	3.43	2.64 - 25.09
Variability	183	11.32	7.81	2.30 - 52.84
WCST				
Categories[b]	184	4.38	1.99	0 - 6
Perseverative Errors[a]	184	20.08	14.99	3 - 73
Correct Responses[b]	184	73.45	15.02	33 - 105
Total Errors[a]	184	39.08	23.17	7 - 95
Accuracy[b]	184	0.67	0.16	0.26 - 0.91
COWAT[b]	183	35.18	11.37	11 - 64
Stroop color-word[b]	183	35.26	11.09	13 - 65
Trail Making Test B[a]	184	85.50	44.62	33 - 300

Note: CPT = Conners' Continuous Performance Test - Second Edition; WCST = Wisconsin Card Sorting Test - Second Edition (Research Version).
[a]higher values indicate lower performance.
[b]lower values indicate lower performance.

The study sample consisted of 93 males and 91 females from 18 to 64 years of age. On average, participants reported 13.81 years of education, ranging from as few as 7 to as many as 24 years. A wide variety of diagnostic groups were represented among the sample participants, who were most commonly determined to have comorbid psychiatric and neurologic disorders. Complete test batteries were available for 181 of the 184 participants. Of the remaining 3 participants, one did not complete the CPT, one did not complete the COWAT, and one did not complete the color-word trial of the Stroop test.

Overall test performance for the sample was determined by scoring the neurocognitive tests using published normative data. For the CPT and WCST, T-scores for all variables were calculated using their respective scoring programs. CPT data were scored with reference to Conners' (1994) general population sample. The WCST was scored using Heaton et al.'s (1993) age and education corrected normative data. COWAT and TMT-B raw scores were first transformed into age and education corrected *z*-scores using Mitrushina's (2005) meta-analytic normative data and were then converted to *t*-scores with a mean of 50 and standard deviation of 10. Golden & Freshwater's (2004) normative data were used to calculate *t*-scores for the color-word trial of the Stroop test.

On the CPT inattention variables (i.e., omissions, hit RT SE, hit RT SE variability), participants scored on average 1 SD lower than the normative sample, indicating that their performance was low average in comparison to members of the general population. A similar pattern was observed on the other study variables, as overall sample performance was approximately 1 SD lower that of the reference groups on all neurocognitive tests (with the exception of TMT-B). Performance on this measure was 0.65 SDs lower than the reference group, perhaps due to the relatively greater variability in performance on TMT-B observed in the general population.

Across-participant consistency in performance on the neurocognitive tests varied widely depending upon both index and instrument. The greatest ranges of performance were observed for CPT inattention variables and TMT-B. With respect to CPT omissions, participants' *t*-scores varied from 20 to 240; hit RT standard error *t*-scores ranged from 29 - 112, and hit RT variability *t*-scores from 30 - 96. Finally, TMT-B *t*-scores ranged from 30 - 124. The standard deviations of these varibles are all greater than 10,

suggesting that the the performance level of study participants was somewhat more variable than the relevant normative samples for these instruments. On the basis of these data, a series of univariate and multivariate outlier screening procedures were employed (as outlined in detail, below) in order to determine whether these scores represented outlying data points or neuropsychologically meaningful variation in observed performance.

Descriptively, observed levels of performance ranged from superior to extremely low on all cognitive tasks. Overall, these data suggest that the sample adequately reflects a wide range of different clinical diagnoses and neurocognitive skill levels. The pattern of data suggests that this sample adequately reflects the diversity of performance levels encountered in clinical practice, and that the results of analyses reported below are likely to generalize across a wide variety of diagnostic groups and levels of ability. Descriptive statistics for demographically corrected neurocognitive test performance characteristics of the sample are depicted below in Table 2.

Table 2

Demographically Corrected Psychometric Test Performance (T-Scores)

Variable	*n*	Mean	SD	Range
CPT[a]				
Omissions	183	64.18	36.78	40 - 240
Standard Error	183	59.18	14.51	29 - 112
Variability	183	57.65	12.69	30 - 96
WCST[b]				
Categories[c]	184	-	-	-
Perseverative Errors[a]	184	43.27	11.83	20 - 81
Correct Responses[c]	184	-	-	-
Total Errors[a]	184	41.65	10.24	20 - 66
Accuracy[c]	184	-	-	-
COWAT[b]	183	42.26	10.06	20 - 68
Stroop color-word[b]	183	41.97	10.82	22-70
Trail Making Test B[a]	184	56.46	16.51	30 - 124

Note: CPT = Conners' Continuous Performance Test - Second Edition; WCST = Wisconsin Card Sorting Test - Second Edition (Research Version).
[a]higher values indicate lower performance.
[b]lower values indicate lower performance.
[c]Published normative data permitting conversion to t-scores not available.

Measures

The Wisconsin Card Sorting Test – Computer Version II (WCST-CV2).

The WCST (and its analogue the WCST-CV2) is an attentionally demanding, complex measure of abstract non-verbal problem solving with a long tradition of use in research and clinical practice (Heaton et al., 1993; Somsen et al., 2000). The basic format of the task requires the examinee to match a stimulus card to one of four response cards on the basis of varying stimulus features (i.e., color, form, or number). After 10 correct responses, the sorting principle is then changed (unbeknownst to the examinee). The administration continues until the examinee either successfully completes six categories

or sorts all 128 cards. The specific version utilized in the present study is a computerized administration based on Heaton's (1981) scoring system that records the time from stimulus presentation to a correct or incorrect response (i.e., response time) as well as more standard clinical variables. Response time is calculated with centisecond resolution, representing a somewhat coarser measurement than typically employed in studies of intraindividual response time variability (which frequently employ instrumentation capable of at least 1 millisecond resolution).

Reliability data for this instrument are somewhat variable in response to different populations and methodologies (see Nyhus & Barceló, 2009 for a review). The most recent set of norms provided in the manual by Heaton and colleagues (1993) lists generalizability coefficients ranging from .39 to .72, with an average of .57 and a median of .60. By some standards these can be considered moderately reliable, and by other standards moderate to good. Heaton's scoring system permits 15 different indices to be derived, of which four have been chosen for the current analysis (i.e., categories completed, perseverative errors, total number of correct responses, total number of incorrect responses).

The validity of the WCST has been examined through investigations using clinical methods, functional neuroimaging, and cortical lesion mapping. Early lesion studies indicated that WCST performance is more affected by lesions to the DLPFC (Milner et al., 1963) and general frontal cortex (Heaton, 1981) than it is by non-frontal lesions. Individuals with diffuse non-frontal lesions, however, also perform poorly on this task (Heaton et al., 1993). More contemporary lesion mapping studies suggest that the WCST perseverative errors score is sensitive to the integrity of the rostral anterior

cingulate cortex and left medial superior frontal gyrus in addition to other non-frontal brain regions, areas found to be crucially important in flexibly switching between response or instruction sets (Gläscher et al., 2012).

Neuroimaging studies highlight the activation of the DLPFC and frontopolar cortex during the performance of the WCST. Bilateral increased activity in frontopolar cortex (Brodmann's Area 10) and DLPFC (Brodmann's Area 9/46) have been observed in FMRI studies of computerized versions of the task (Berman et al., 1995; Nagahama et al., 1996). Authors have speculated that frontopolar cortical activation in particular may be involved in feedback evaluation and hypothesis generation/evaluation, two mental processes specifically implicated in WCST performance (Christoff & Gabrielli, 2000). Others have suggested that area 9/46 is also activated when subjects must relate test feedback information to earlier events stored in working memory (i.e., when receiving positive or negative feedback; Monchi, Petrides, Petre, Worsley, & Dagher, 2001). Overall, the evidence highlights the substantial contributions of the DLPFC and frontopolar cortex to WCST performance, in addition to other non-frontal brain regions.

Conners' Continuous Performance Test – Second Edition (CPT-II).

Conners' CPT-II is one of many variants of the Continuous Performance Test, pioneered by Rosvold, Mirsky, Sarason, Bransome, & Beck (1956). Various forms of this test have a long history of use in schizophrenia research (Orzack & Kornetsky, 1966). Conners' version presents stimuli in the form of letters on a computer screen separated by interstimulus intervals that vary from one to four seconds. The examinee is instructed to press a key as quickly as possible in response to every letter beside X. This implementation requires a higher frequency of responding than previous iterations of the

CPT, during which the examinee is traditionally instructed to respond only to targets (i.e., X) and to ignore non-targets (Conners, 2000). Statistically, more frequent responding results in greater reliability for various performance indices. Three CPT performance metrics are utilized in the present study: omissions, hit RT standard error (hit RT SE), and hit RT standard error variability (hit RT variability).

Omission errors occur when the examinee fails to respond to a target stimulus, indicating a lapse of attention or failure to orient to the test (Conners, 2000). By contrast, hit RT SE reflects the standard deviation of an individual's log-transformed response times on target trials (collapsed across interstimulus intervals) and as such, it is a measure of response time variability from trial to trial. Hit RT SE variability indexes within-individual response consistency across occasions; specifically, the consistency of responding over the 18 sub-blocks of the task. Consequently, it represents an index of variability across occasions. Omissions, hit RT SE, and hit RT SE variability are all thought to reflect inattentiveness.

The split-half reliability coefficients for omission errors and hit RT SE are high (r = .94 and r = .87, respectively) as reported in Conners' original 520 case standardization sample (Conners, 1994). Split half reliability for hit RT SE variability is somewhat lower (r = .66). Test-retest stability coefficients for all three variables are lower than their reliabilities, ranging from r = .60 (hit RT variability) to r = .84 (omissions). Lower test-retest reliability may reflect the sensitivity of the CPT to subtle fluctuations in various neurophysiological substrates across occasions rather than unreliability per se as a psychometric instrument. The review up to this point gives substantial credence to such an interpretation.

In validity studies, the CPT-II has been found to be sensitive to neurologic dysfunction of various types (including psychiatric disturbance, attention deficit/hyperactivity disorder, and temporal lobe epilepsy; Strauss et al., 2006). Clinically, omissions, hit RT SE, and hit RT variability were found to differentiate between ADHD individuals and normal controls in Conners' original standardization sample (Conners, 1994). In this study, his "neurological" group was significantly less consistent and made significantly more omission errors than the ADHD group. Research has generally converged on an understanding of various CPTs as measures of sustained attention that depend upon intact cortical, subcortical, and functional systems to include various regions of the frontal lobes.

The multitude of different CPTs and widely distributed nature of the attentional system, however, make definitive statements about the neurostructural bases of Conners' CPT problematic (Riccio, Reynolds, Lowe, & Moore, 2002). The three performance indices in the present study were chosen due to their conceptual similarity with hypothesized first-stage executive functioning processes rather than evidence indicating that they map on to brain regions that subserve this cognitive skill (Petrides, 1995). The CPT paradigm involves "comparisons between, or judgments about the occurrence or non-occurrence of remembered stimuli," synonymous with published descriptions of first-stage working memory (Owen, 2000, p. 34). Nonetheless, some fMRI studies have found evidence of ventral prefrontal activation during the performance of Conners' CPT, consistent with the interpretation that it does involve first-stage executive processing (Ogg et al., 2008).

The Stroop Color-Word Test (Stroop).

The Stroop test was formally adopted by Stroop (1935), but its roots can be traced to paradigms used by Cattell near the beginning of the 20th century (Mitrushina, 2005). It is a measure of selective attention, response inhibition, and conflict monitoring (aspects of cognitive control). The color-word trial requires the examinee inhibit a dominant word-reading response in favor of a novel color-naming response while reading a list of color-words. This is the most challenging task condition and has been found to differentiate between clinical groups with greater reliability than other test conditions (Strauss et al., 2006). The Stroop variant utilized in this investigation is the Golden version, which is one of the more common iterations employed in clinical practice (Golden & Freshwater, 2002). The performance index included in analyses is performance on the color-word trial, as evidenced by the total number of color-words correctly named during a 45-second time interval.

The test-retest reliability of the Stroop color-word score is adequate, with a reliability coefficient of .73 in both large-scale group administrations (N = 450) and smaller individual administrations (N = 30; Golden, 1975). Investigations by other researchers have generally confirmed these results (Franzen, Tishelman, Sharp, & Friedman, 1987). The Stroop test is subserved by working memory, as this cognitive ability seems to predict Stroop performance (Kane & Engle, 2003). There is some evidence that this instrument measures frontal/executive processes from studies of closed head injury; lesions to bilateral superior medial frontal areas of the cerebral cortex effect both slowness and a lower rate of accuracy (Stuss, Floden, Alexander, Levine, & Katz, 2001). Neuroimaging studies show right frontal activation during the color-word

interference condition (Bench et al., 1993). Poor performance on the color-word trial of the stroop test has been found accompanying left DLPFC damage in large-scale cortical lesion mapping studies, consistent with other evidence that underscores this region's role in response inhibition (Gläscher et al., 2012).

Controlled Oral Word Association Test (COWAT).

The COWAT is a measure of fluency. Performance is assessed via the total number of words generated in response to particular starting letters over a 60-second interval per trial. The present study uses phonemic fluency to the letters F, A, and S, the characters most commonly used for this task (Strauss et al., 2006). There are differences in levels of performance across different letters and the internal consistency between the letters F, A, and S is high, with a coefficient alpha of .83 (Tombaugh, Kozak, & Rees, 1999). The test-retest correlations for this measure are also high; in the investigation cited above, the authors found a reliability coefficient of .74 after more than five years had elapsed between administrations. This result has been confirmed in numerous studies, most documenting test-retest correlations in excess of .70 after varying test-retest intervals (Strauss et al., 2006).

Factor analytic studies of the tests validity have implicated attentional control and working memory as key contributors to intact performance on verbal fluency measures (Brocki & Bohlin, 2004). An analysis of the processes involved in the performance of phonemic fluency tasks supports this notion. Words with the same first letter are unlikely to be organized in any kind of semantic network; thus, phonemic fluency involves the suppression of previous responses, monitoring one's performance to avoid errors, and accessing new items (Azuma, 2004; Strauss et al., 2006). These are all attentionally

demanding processes, a contention supported by neuroimaging data showing activation of the left inferior frontal cortex and left dorsolateral frontal cortex during the performance of this task (Ravnkilde, Videbech, Rosenberg, Gjedde, & Gade, 2002). In the previous study, this finding lead the authors to conclude that phonemic fluency was a test relatively specific to the prefrontal cortex. Evidence from lesion mapping demonstrates the necessity of intact left frontoparietal cortex, anterior prefrontal cortex, and insula for phonemic fluency even when test performance is purified of verbal and memory ability (Gläscher et al., 2012).

Trail Making Test A & B (TMT A & B).

The Trail Making Test is a visual sequencing task that has been in use for nearly seven decades and debuted as part of the Army Individual Test Battery (1944). Since that time, it has entered the public domain, and this version of the test was used for the current study. On condition A (the single alternating sequence) the examinee must connect a series of numbers in order from 1 to 25 as fast as possible without making mistakes. On condition B (the double alternating sequence) they are instructed to switch between numbers and letters in order (i.e., 1 - A - 2 - B - 3 - C…) as fast as possible, without making any mistakes, until they reach the end of the sequence. The outcome measure is the total time to complete each sequence. TMT-B performance is used as a variable in the present study, given evidence that TMT-A relies on more peripheral sensory-motor processes (Arbuthnott & Frank, 2000).

The reliability of this instrument varies depending on the age range of test takers. In one study covering an age range of 15 to 83 years, test-retest reliability was high for part B (.89) following an 11 month inter-test interval (Dikmen, Heaton, Grant, & Temkin,

1999). Some authors have found conflicting results and overall, it seems that the test-retest reliability of this measure is generally high but that this figure may not be robust across different populations and test-retest intervals (Strauss et al., 2006). Interrater reliability for the TMT is high for both conditions (.90 for TMT- B; Strauss et al., 2006).

Recent investigations demonstrate correlations between TMT-B and other set switching tasks, despite allegations that TMT-B performance may be confounded by visuomotor processes (Arbuthnott & Frank, 2000). Conceptually, TMT-B is a prototypical measure of "set switching" or "shifting." Neuroimaging data are generally consistent with this interpretation, documenting lateralized dorsolateral and medial frontal activity on the left side of the cerebral hemispheres in the performance of condition B (Zakzanis, Mraz, & Graham, 2005). These regions are also activated in the performance of the Stroop task and WCST test, and they seem to be related to cognitive flexibility and visual motor functioning (Moll, Oliveira-Souza, Moll, Bramati, & Andreiuolo, 2002). The TMT is also sensitive to closed-head injuries and performance is inversely correlated with the severity of injury (Iverson, Lange, Green, & Franzen, 2002). Lesion mapping efforts suggest substantial overlap between cortical areas subserving performance on TMT-B and WCST perseverative errors, which share conceptually similar set-shifting task demands (Gläscher et al., 2012). Specifically, both of these tasks depend upon the integrity of the anterior cingulate gyrus, with TMT-B performance more strongly related to the fronto-rostral area of this structure.

Outlier Screening and Data Preparation

Response time data from the WCST (used to fit the ex-Gaussian distribution) and cognitive test scores were subject to extensive data screening in order to ensure parameter

extraction and fidelity to the underlying assumptions of statistical tests. RT data were gathered from computerized records of subjects' WCST performance. At the level of individual participants response times, only RTs on correct trials of the WCST were submitted to ex-Gaussian analyses (as is standard practice in response-time distributional analysis research). There are important theoretical ramifications of this decision. Because the WCST is a measure amenable to strategy use, it is possible that it captures both adaptive and maladaptive forms of variability as outlined by Li and colleagues (2004). The latter form of variability is expected following rule attainment (when examinees are sorting correctly), as fluctuation in response times may reflect variation from optimal levels of task performance and compromised "processing robustness." By using correct trials only, analyses are more likely to capture this maladaptive form of variability.

However, it is reasonable to assume that participants may have some correct sort trials during which they are engaging in rule search rather than rule application (i.e., they have not yet attained the sorting principle but respond correctly nonetheless). On these trials, variability would be expected to reflect adaptive processes (plasticity or diversity; Li et al., 2004). Adaptive variability has been found in a number of studies examining tasks that may benefit from practice or strategy use (Allaire & Marsiske, 2005). The presence of such adaptive variability is a potential confound in the present investigation. Unfortunately, the WCST paradigm does not allow for definitive detection and exclusion of these data points. Including these trials will represent a more stringent test of the relationship between intraindividual variability and executive functioning because adaptive forms of variability are expected to mask the effects of maladaptive variability.

Outlier detection and treatment was also a primary concern. When distributions are characterized by a defined lower bound and an undefined upper bound (e.g., response time distributions), they frequently display positive skew (Cousineau & Chartier, 2010). This makes the detection of outlying data points highly problematic to say the least and has stimulated a considerable amount of research in the response time literature. The use of cutoffs based on standardized scores (i.e., *z*-scores) is a practical method that is statistically sound (Ratcliff, 1992) and ubiquitously employed when fitting response time distributions. To implement this censoring method, the response times of individual participants are first transformed into standard scores (using the M and SD of each participants' response time distribution) and then a pre-specified cutoff is applied. Popular cutoff values range from 3SDs (e.g., Geurts et al., 2008; Rentrop et al., 2010; Tse et al., 2010) to 4SDs above individual participants' mean RT values (e.g., Leth-steensen et al., 2000; Schmiedek et al., 2007). Such censoring procedures typically result in an average loss of 1-3% of correct RT trials per participant.

The WCST is substantially different than most tasks featured in the variability literature, so a series of cutoff values were investigated to determine which threshold best fit the study data. RTs were first standardized on a participant-by-participant basis (i.e., using the M and SD of a given participant's response-time distribution). Applying cutoff values ranging from $z = 3$ to $z = 5$ resulted in a loss of 0.52 - 2.12% of correct RTs (as averaged across participants). Careful visual inspection of participants' RT distributions suggested that a cutoff of 4SDs represented the best compromise between removing very extreme response times and eliminating potentially meaningful data in the tail of individual RT distributions. Response times less than 200ms were also eliminated from

analyses given that they are likely to represent accidental key presses or other non-task related processes.

Outlier screening was also conducted at a variable level. With data that are assumed to follow a normal distribution, symmetry of the data is most salient and outlying data points are likely to be those that deviate from the mean by a particular margin. The results of exploratory analyses, however, suggested substantial departures from normality (as indexed on the basis of skew and kurtosis) for the CPT inattention variables, ex-Gaussian parameter estimates, TMT-B, and WCST perseverative errors. There are no universally accepted procedures to use in this circumstance (Cousineau & Chartier, 2011). Accordingly, more careful examination of the study variables was necessary to determine an optimal approach.

A decision criterion of $\alpha = 0.01$ was applied on a variable-by-variable basis to reveal potentially outlying data points. Many authors recommend such a conservative decision criterion, reflecting a bias towards keeping data (Cousineau & Chartier, 2011). The decision criterion was then subject to Bonferroni correction on the basis of sample size (i.e., $1- \alpha/2n$). Applying this criterion to the entire data set resulted in the identification of only 11 outlying data points, representing 0.054% of the total set, and no more than 1.6% of any given variable. As observed by Cohen, Cohen, West & Aiken (2003) "if outliers are few (less than 1 - 2% of *n*)… they are probably best left alone" (p. 128). Accordingly, these potentially outlying data points were retained when testing the study hypotheses. These univariate analyses were complemented by a series of multivariate outlier detection strategies to guard against unduly influential and outlying

data points in the regression analyses and are described in detail in their respective analysis sections below.

The main study hypotheses were evaluated using a series of hierarchical multiple linear regression analyses. The assumptions underlying this statistical method were considered at all levels of data analysis. Multiple regression requires a linear relationship between the criterion variable and predictor variables, normality of residuals, homoscedasticity, and independence of errors (Field, 2013). One might assume that the departures from normality described above represent a threat to statistical inference, yet neither univariate nor multivariate normality is required for the appropriate application of multiple regression (Howell, 2012). Nonetheless, substantial departures from normality can bias the results of analyses in a number of ways, particularly when criterion variables are highly asymmetric (Tabachnik & Fidell, 2001). Ex-Gaussian criterion variables all demonstrated significant departures from normality (on the basis of skewness) and unacceptably high levels of kurtosis (kurtosis values were Mu = 6.89, Sigma = 10.89, Tau =21.50). To avoid potential complications, Mu and Tau were subject to logarithmic transformations and Sigma was subject to a natural logarithmic transformation, both of which are frequently employed in the analysis of response time data (Ratcliff, 1992). Skewness and kurtosis values for the transformed variables were below 1.0 (with the exception of Sigma), indicating adequate approximation of a normal distribution (Meyers, Gamst, & Guarino, 2005). Visual analysis of normal Q-Q plots confirmed that transformations had the expected effects. Tests for violations of statistical assumptions (i.e., homoscedasticity, normality of residuals, and independence of errors) are reported in the results section below.

CHAPTER 4

RESULTS

Ex-Gaussian Fits

Raw response time data from the WCST were fit using the Quantile Maximum Probability Estimator QMPE (v. 2.18; Cousineau et al., 2004; Heathcote et al., 2002). This software allows the researcher to specify whether the ex-Gaussian distribution is fit on the basis of observed RT data (continuous maximum likelihood; CML) or quantiles calculated from those data (quantile maximum likelihood; QML). QML was used for all analyses. The program generates an exit code for each fit that provides important information about the degree to which data were adequately approximated by the likelihood solution, such as whether parameter convergence was obtained, whether standard error estimates are reliable, and if sample statistics (i.e., quantiles) could be calculated. According to the programmers:

> any exit code smaller than 32 means that both the parameter estimates and their standard errors and correlations are trustworthy…the parameter estimates themselves are probably useful as long as the exit code is smaller than 125 (Brown, Cousineau, & Heathcote, 2004, p. 5).

Quantiles were calculated from an average of 72.4 correct trials RTs on the WCST and parameter recovery was successful for all 184 participants. Parameter convergence occurred in fewer than 23 iterations for all fitting runs (250 were allowed). QMPE reported an exit code below 32 for 94% of fits, suggesting that most participants' response times were fit without incident. Of the remaining 6% of participants, the vast majority of exit codes revealed singularity of the hessian, leading to unreliable standard

errors (which are not necessary to test the study hypotheses). Overall, these data suggest excellent parameter recovery and appropriate application of the fitting methodology.

A series of visual analyses confirm the numerical data reported above. Simultaneously plotting the quantiles estimated by QMPE alongside those that characterize participants' actual WCST response time distributions allows one to evaluate the extent to which the estimated parameters reflect participants' actual response times. That is, the program provides an output of observed quantiles (i.e., actual data) and expected quantiles (i.e., estimated from the best fitting model). Graphing these quantiles yields a visual indicator of the degree of fit. In a model that fit perfectly, the expected quantiles would lie directly over the observed quantiles, such that the two sets of data points would be indistinguishable. Figure 5 presents simultaneous quantile plots for 6 randomly selected participants. Visual inspection of these plots suggests relatively high equivalence between quantiles derived from the likelihood solution and those observed in the data. There is clearly some variation in the degree of fit, but overall, there is a high degree of correspondence between estimated ex-Gaussian models and the observed data. This visual analysis suggests that the parameters extracted from participants' response times reflect the actual data quite well.

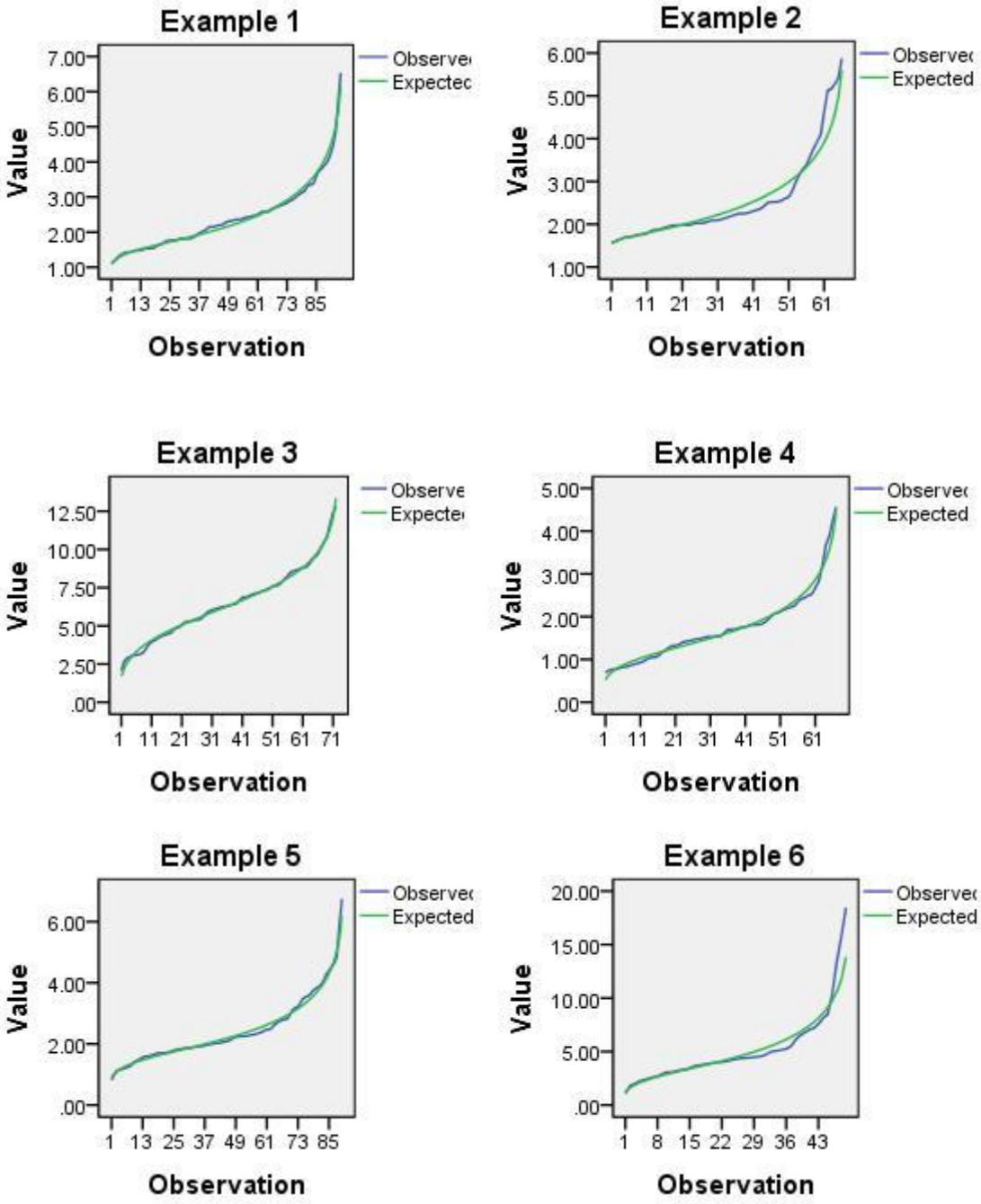

Figure 5. Observed vs. expected quantiles for a random sample of 6 participants. X axes reflect the number of quantile cut points per participant (in ascending order). Y axes display the response times corresponding to the quantile cut points in seconds.

QMPE also provides estimates of parameter standard errors and intercorrelations. As depicted below in Table 3, the standard error values of the ex-Gaussian parameters were somewhat variable, which is significant for interpreting the regression equations testing the primary study hypotheses. The various parameters of a multiple regression model are affected by the reliability of variables in the model, such that R^2 and Beta values are reduced in the presence of less reliable variables (Cohen, Cohen, West, Aiken, 2003). Beta values are most strongly affected by reliability of the criterion variable. Despite relatively similar values for Mu and Tau, the standard error of Tau is nearly twice that of Mu; as a result, this parameter may be at a disadvantage in subsequent correlation analyses. Differences in the variability of parameters across participants can also affect regression results. For example, the range of Sigma is considerably less than that of the other ex-Gaussian parameters, which is expected to attenuate its correlation with the neurocognitive predictor variables.

Table 3

Descriptive Statistics of ex-Gaussian Fits

Variable	*n*	Mean	SD
μ	184	1.6453	0.8883
σ	184	0.3112	0.327
τ	184	1.7741	1.5259
Number of RT Observations	184	72.40	14.65
Standard Error Mu	173	0.1519	0.1425
Standard Error Sigma	173	0.1386	0.1291
Standard Error Tau	173	0.2704	0.2409
$r_{\mu\sigma}$	173	0.6075	0.1842
$r_{\mu\tau}$	173	-0.5137	0.1775
$r_{\sigma\tau}$	173	-0.3632	0.1865

Note: $r_{\mu\sigma}$ = bivariate Pearson correlation between Mu and Sigma.
$r_{\mu\tau}$ = bivariate Pearson correlation between Mu and Tau
$r_{\sigma\tau}$ = bivariate Pearson correlation between Sigma and Tau.

The parameter intercorrelations in Table 3 have important ramifications for all analyses that involve the ex-Gaussian parameters. Parameter dependencies are a distinct disadvantage of response time distributional analyses that involve fitting descriptive models to RT distributions (see Cousineau, Brown, & Heathcote, 2004 for a detailed discussion of this problem as pertaining to QMPE). Essentially, trade offs during the parameter estimation process lead to correlated deviations of these estimates from true parameter values (Schmiedek et al., 2007). As a result, the observed intercorrelations are a combination of each parameter's true relationship with the other parameters and artifacts of the estimation process. These interdependencies can be overcome by specifying a measurement model of ex-Gaussian parameters (e.g., Schmiedek et al, 2007; Tse et al., 2012), but this methodology was not an option due to data restrictions. The

effects of parameter dependencies on conclusions that can be drawn from the following analyses are considered below in the discussion on study limitations.

Regression Analyses of Hypothesis 1

A series of hierarchical regression analyses were conducted to test the hypothesis that lower levels of cognitive control accompany greater levels of intraindividual variability. In these analyses, raw scores for WCST perseverative errors, TMT-B, COWAT, and Stroop color-word trial were used as predictors for the three ex-Gaussian criterion variables. Age was entered as a forced covariate in the first block of the models followed by a second block including the neurocognitive test scores. The order of terms in the model was selected via stepwise variable entry using $p = 0.05$ to enter and $p = 0.10$ to remove. Unfortunately, the absence of sufficient background information did not allow theoretical pre-specification of the expected order of neurocognitive test scores.

Moderating analyses were conducted to investigate interaction effects for all significant models including two or more cognitive test scores. To test moderation, significant cognitive control variables were first centered by subtracting their mean values from observed test scores. The product of the centered variables was then entered into a regression model as the third term (after the neurocognitive test scores from which it was derived). A significant interaction term indicates that the relationship observed in the regression model varies as a function of cognitive test performance. Interactions were then evaluated via 3-dimensional graphing and are interpreted below.

Before conducting all analyses, multivariate outliers were identified by computing a Mahalanobis distance measure (D^2) for the four neurocognitive test scores. D^2 provides an index of the distance between individual data points and the centroid (multivariate

mean) of the pattern mass in multivariate space. There were no data points with a D^2 suggesting that they significantly differed from the centroid at the prespecified level (α = .05). Unduly influential data points were detected on an analysis-by-analysis basis through computing Cook's Distance (CD_i), which is a measure of change in the regression coefficients that would occur if a particular case were eliminated from the analysis. As a very general rule of thumb, a CD_i of about 1 is generally considered large (Cook & Weisberg, 1982). No data points achieved a CD_i of greater than 0.20.

Fidelity to the assumptions of multiple regression was examined using a variety of methods. Linearity was determined via visual inspection of scatter plots and then subsequently confirmed by significant results of the main analyses (i.e., a significant linear regression effect implies adequate linearity of the data). Independence of errors is indexed by the Durbin-Watson statistic, with values significantly below 2.0 indicating positive autocorrelation of residuals and values above 2.0 revealing negative autocorrelation. This statistic ranged from a minimum value of 1.938 (for Sigma) to a maximum value of 2.169 (for Tau), suggesting that the assumption of independently distributed error terms was tenable. The assumption of homoscedasticity was evaluated via plots of residual vs. predicted values and was met for all analyses except those involving Sigma. Submitting error terms to the Shapiro-Wilk test showed that the assumption of normally distributed residual variance was supported for all parameters but Sigma. Violations of assumptions during the analysis of Sigma are discussed in that section below.

Bivariate Pearson correlations between the log-transformed ex-Gaussian parameters and cognitive-control variables are presented in Table 4. In general,

correlations with the cognitive control variables were higher for Mu and Tau than they were for Sigma, consistent with expectations on the basis of the literature review. The only variable significantly related to Sigma was TMT-B (r = .207, p = .005). The other ex-Gaussian parameters were meaningfully related to all of the cognitive control tasks. Mu bore a stronger relationship to Stroop test performance (r = -.456, p < .000) than Tau, while Tau's relationship to TMT-B (r = .571, p < .000), the COWAT (r = -.340, p < .000), and WCST Perseverative errors (r = .520, p < .000) was greater than Mu's. Visual analysis of the table suggests a small but consistent advantage to Tau in predicting performance on the cognitive control tests. Overall, these data suggest small to moderate relationships between Mu, Tau, and the cognitive control tasks, all of which are in the expected direction.

Table 4

Pearson Correlations Between ex-Gaussian Parameters, Cognitive Control Test Scores, and Age

	Age	μ	σ	τ	TMT-B	COWAT	Stroop	WCST
Age	-	.603**	.142	.386**	.250**	-.169**	-.349**	.229**
μ		-	.407**	.564**	.495**	-.299**	-.456**	.494**
σ			-	-.003	.207**	-.139	-.140	.121
τ				-	.571**	-.340**	-.408**	.520**
TMT-B					-	-.409**	-.510**	.479**
COWAT						-	.303**	-.301**
Stroop							-	-.331**
WCST								-

Note: μ = $base_{10}$ logarithm of Mu parameter of ex-Gaussian distribution, σ = $base_e$ logarithm of Sigma parameter of ex-Gaussian distribution, τ = $base_{10}$ logarithm of Tau parameter of ex-Gaussian distribution, TMT-B = Trail Making Test B (total seconds to completion), COWAT = Total words generated to the letters FAS, Stroop = Total words named on color-word trial of Stroop Test, WCST = WCST perseverative errors score.

*$p < 0.05$

**$p < 0.01$

For the analysis of response latency, the best fitting model was statistically significant ($F = 72.564$, $p < .000$) and accounted for 55.0% of the observed variance in Mu ($R = .742$). Mu was predicted by age, WCST perseverative errors, and TMT-B, such that increasing age and poorer performance on the cognitive tests resulted in greater overall response latency. Age received the strongest weight in the model, followed by WCST perseverative errors and TMT-B. Examining partial correlations revealed that age, WCST perseverative errors, and TMT-B each accounted for 32.60%, 10.76%, and 9.36% of the variance in Mu when controlling for the other predictors. R^2 change values indicate that the combined contribution of the cognitive tests accounted for 18.1% of the variance in Mu. Due to their strong intercorrelations however, semipartial correlations with Mu were somewhat lower in the final model. Age, WCST perseverative errors, and TMT-B uniquely explained 21.71%, 5.43% and 4.62% of the variance in Mu, respectively. The regression coefficients of the predictors as well as their squared partial and semipartial correlations with Mu are shown in Table 5.

Table 5

Results of Regression Analyses (Final Models)Predicting Mu, Sigma, and Tau From Cognitive Control Test Performance

	μ					σ					τ				
	R^2	*b*	β	$r_{partial}$	r_{semi}	R^2	*b*	β	$r_{partial}$	r_{semi}	R^2	*b*	β	$r_{partial}$	r_{semi}
	.550					.058					.451				
Age		.007	.485	.571	.466		.011	.113	.112	.110		.004	.225	.280	.216
TMT-B		.001	.248	.306	.215		.006	.002	.183	.181		.002	.376	.403	.326
WCST		.004	.267	.328	.233							.005	.289	.322	.252

Note: μ = $base_{10}$ logarithm of Mu parameter of ex-Gaussian distribution, σ = $base_e$ logarithm of Sigma parameter of ex-Gaussian distribution, τ = $base_{10}$ logarithm of Tau parameter of ex-Gaussian distribution, R^2 = multiple correlation coefficient, *b* = unstandardized regression coefficient, β = standardized regression coefficient, $r_{partial}$ = partial correlation coefficient between criterion and predictor, r_{semi} = semipartial correlation coefficient between criterion and predictor. Parameters from final model only.

For Sigma, the final best fitting regression was significant ($F = 5.528$, $p = .005$) but this time accounted for only 5.8% of the variance ($R = .241$). In this model, age received the strongest weight followed by TMT-B. Age uniquely predicted 1.21% of the variance in Sigma and TMT-B predicted 3.28% as evidenced by their squared semipartial correlations. Pertinent indices from this analysis are displayed in Table 5 but should be interpreted with caution given violations of homoscedasticity and normality of residuals. Visual inspection of the regression standardized residual vs. predicted values suggested potential heteroskedasticity, which would render significance tests for regression weights less reliable.

Normality of residual variance was also violated (Shapiro-Wilk = .792, $p < .000$) which frequently occurs in the presence of heteroskedastic variables. Violation of these assumptions is likely to be caused by the significantly positively skewed distribution of Sigma in the current sample. Non-normality was improved by the natural logarithmic transformation used in all analyses but could not be eliminated. Logarithmic, inverse, and inverse square root transformations were not effective at reducing Sigma's level of positive skew and kurtosis.

In the analyses of intraindividual variability, the final model was also significant ($F = 48.741$, $p < .000$) and accounted for 45.1% of observed variance in Tau. The structure of the model was similar to that for analyses of Mu, in that age, TMT-B, and WCST perseverative errors were all significant predictors of Tau. In this model, age received the strongest weight, followed by TMT-B and WCST perseverative errors. In contrast to the analyses of Mu however, the unique contribution of age was substantially less than that of the cognitive test scores. TMT-B and WCST perseverative errors jointly contributed to the prediction of 30.3% of the variance in Tau (as evidenced by R^2 change values). Semipartial correlations revealed that Age, TMT-B, and WCST perseverative errors each uniquely accounted for 4.67%, 10.63%, and 6.35% of Tau's variance, respectively. The interaction effect between TMT-B and WCST perseverative errors was significant for the overall model including Age, TMT-B, WCST perseverative errors, and the interaction effect of the latter term ($F = 39.194$, $p < .000$). In this model, the interaction effect achieved a Beta weight of -.153 ($t = -2.248$, $p = .026$). This interaction was explored graphically by removing the linear effect of age from all variables in the analysis and plotting the regression surface of a model predicting Tau

from the cognitive tests (Figure 6). As seen below, the moderate positive correlation between Tau and TMT-B became weaker at lower levels of WCST performance. In addition, the relationship between Tau and WCST performance became weaker and then reversed at the lowest levels of TMT-B performance. There are several alternative explanations for this interaction that are explored in detail below. One possibility is that individuals who fared extremely poorly on TMT-B demonstrated more consistent poor performance on the WCST as perseverative error scores increased. Another possibility is that individuals who performed poorly on either the TMT or WCST were all highly variable, introducing a restriction of range and instability in the regression weights.

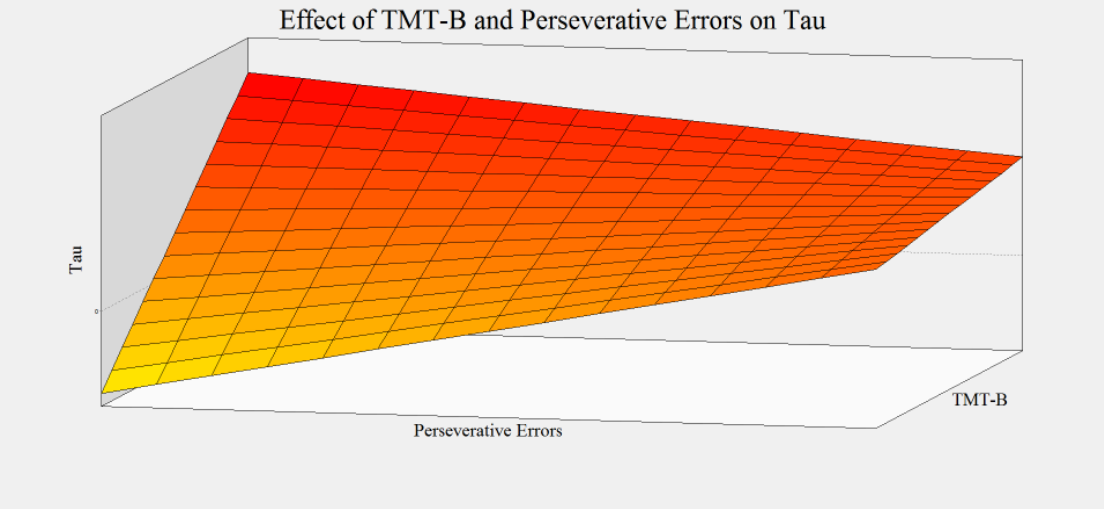

Figure 6. Regression surface of the linear model predicting Tau from Trail Making Test B, Wisconsin Card Sorting Test Perseverative Errors, and their interaction. The effects of age were removed from all variables prior to computing the regression equation.

Visual Analyses of Hypothesis 1

Quantile-quantile plots can provide a visual analysis of between group differences in RT distributions and clarify the results of statistical analyses (Balota et al., 2008). This methodology was used to graphically investigate the shape of RT distributions for

individuals with high vs. low cognitive control (as measured by the neurocognitive tests). These analyses allow for detailed evaluation of the significant correlations reported above by simultaneously examining the contributions of all three ex-Gaussian parameters across the entire response time distribution. It is important to note that these plots do not directly correspond to the regression analyses, however, because the effects of demographic variables were removed to make the results more meaningful.

Analyses were carried out on a test-by-test basis. To form these plots, participants were first ranked (using demographically corrected *t*-scores) in order of their performance on TMT-B, WCST perseverative errors, COWAT, and Stroop color-word trial. Ranking in this fashion is functionally equivalent to controlling for demographic variables in a statistical analysis. Groups of high vs. low cognitive control ability were formed by sampling participants in the highest and lowest quartiles of performance for each cognitive test. RT data for these participants was then re-fit using QMPE with a prespecified number of quantiles. Next, quantiles were averaged across participants within a given group. The resulting quantile bands represent group level RT distributions. Contrasting the two quantile bands allows one to index differences in the shape of RT distributions for those in the highest vs. lowest quartiles of ability on each cognitive task. The results of these graphical analyses appear below in figures 7 - 10.

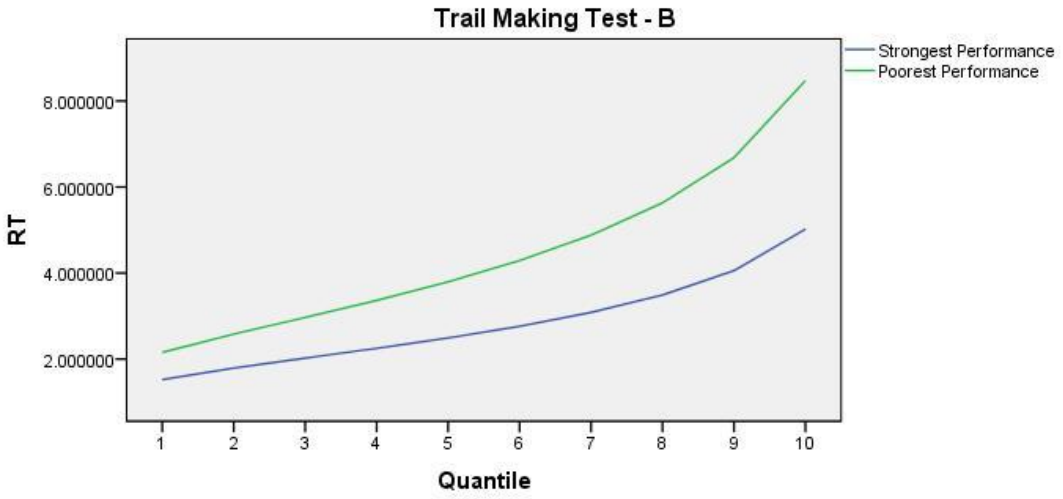

Figure 7. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on TMT-B. RT displayed in seconds.

Figure 7 depicts the shape of RT distributions for high vs. low scorers on TMT-B. The vertical distance between the two quantile bands demonstrates an effect in Mu, meaning that participants with stronger performance were, on average, also faster in completing the WCST. A Tau effect is also evident, as separation between the bands increases in an accelerating fashion at progressively more extreme response time values. That is, increased separation is observed in the later quantiles. Quantiles 8 - 10 of the poorest performing subjects (representing the slowest 20% of response times for each group) show particular acceleration, which suggests a maximal Tau effect for these quantiles. There is no obvious effect for Sigma, which would present as a sigmoid shape in the quantile band.

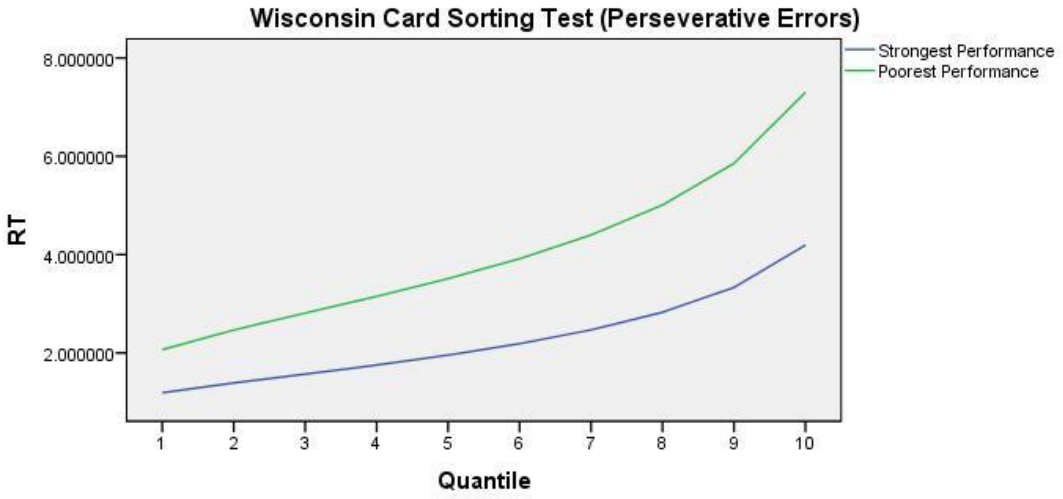

Figure 8. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on WCST (perseverative error score). RT displayed in seconds.

For the WCST (Figure 8), the effects in Mu and Tau are also apparent. Again, separation between the bands shows group differences in Mu. The effect in Mu for this cognitive test appears to be larger than for the TMT, consistent with the regression analyses that suggested a greater correlation between Mu and WCST performance than for Mu and TMT-B performance. The effect in Tau seems to be more distributed across quantiles for the WCST than the TMT, as acceleration is evident beginning at Quantile 6 or 7. There is no obvious effect in Sigma.

The Stroop test and COWAT were not significant predictors of Mu, Sigma, or Tau in the regression analyses, but bivariate correlations of these tests and the ex-Gaussian parameters were significant. Bivariate correlations of the Stroop were small (r = -.456, p < .000) for Mu, non-significant (r = -.140, p = .059) for Sigma, and small (r = -.408, p < .000) for Tau. The relationship between the COWAT and the ex-Gaussian parameters was negligible to small (r = -.299, p < .000) for Mu, small for Tau(r = -.340, p < .000) and non-significant (r = -.139, p = .061) for Sigma. The quantile plots below

are consistent with these correlational analyses, as shown in Figure 9 (Stroop) and Figure 10 (COWAT). Effects in Mu are evident from the separation of quantile bands and effects in Tau are apparent from increasing acceleration across slower quantiles. These effects appear to be almost the same magnitude as those observed in WCST and TMT-B performance. Because participants were ranked on the basis of demographically corrected *t*-scores (removing these effects of these variables in the visual analyses), it is possible that demographically corrected Stroop and COWAT scores might be significant predictors of Mu and Tau in regression analyses but failed to achieve significance in tests of hypothesis one due to the presence of demographically related error variance.

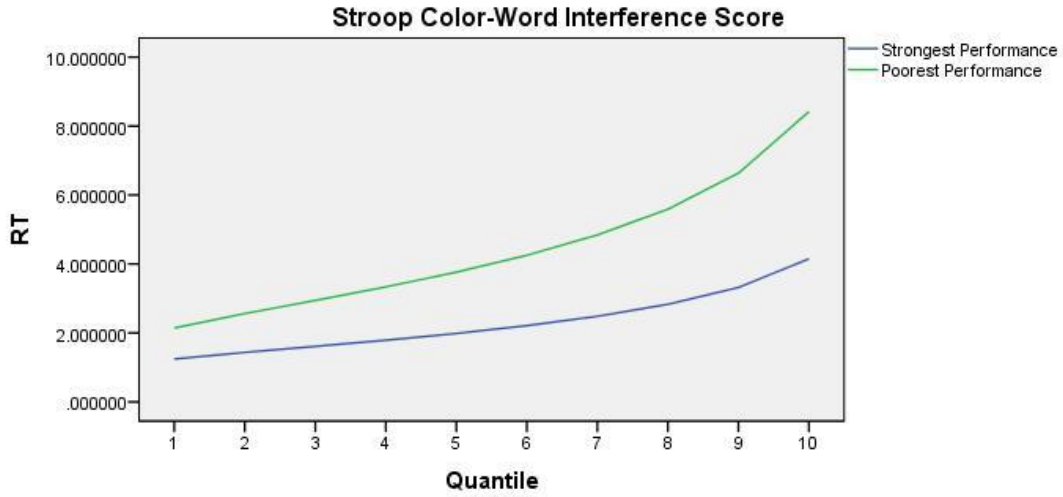

Figure 9. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on the Stroop color-word trial. RT displayed in seconds.

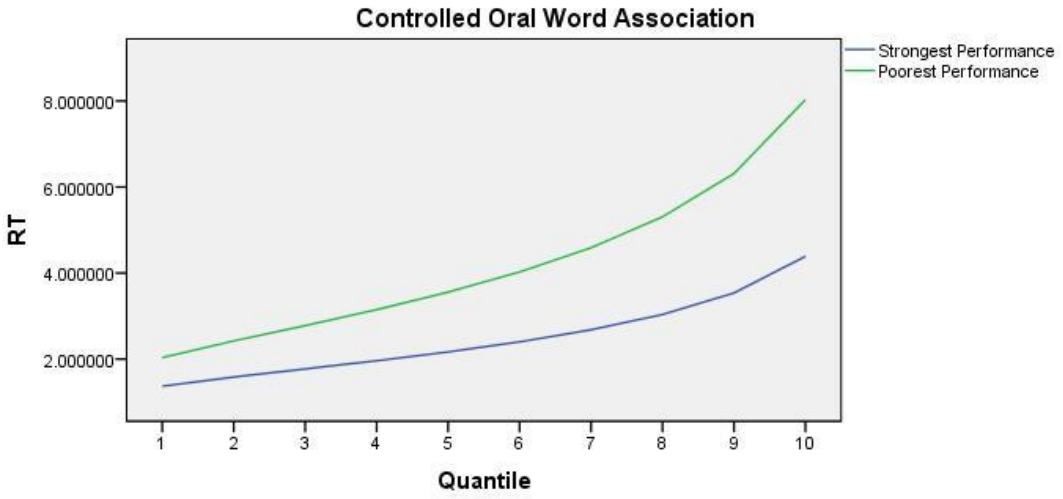

Figure 10. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on the COWAT. RT displayed in seconds.

Regression Analyses of Hypothesis 2

Hypothesis two states that intraindividual variability will be related to more basic "first stage" executive functions. These cognitive skills involve the effortful deployment of simple attentional resources, such as holding information in abeyance in order to make comparisons about the occurrence or non-occurrence of a stimulus (Owen, 2000). In the regression analyses testing this hypothesis, variables from the CPT (i.e., hit RT SE, hit RT Variability, and omissions) were specified as predictors of the ex-Gaussian parameters. Because hit RT SE represents a prototypical measure of intraindividual variability, tests of this hypothesis also provide an index of the extent to which variability on the WCST is related to more traditional measures of variability. Technically, hit RT SE is a measure of two-choice reaction time variability (as indexed by the intraindividual standard deviation; iSD). The general procedure for specifying the regression models was the same as in tests of hypothesis one. Age was entered as a forced covariate in the first block of the models and CPT scores were selected via stepwise variable entry (p = 0.5 to

enter, p = 0.10 to remove) during a second block. Interaction effects were tested using the same procedure as in tests of hypothesis 1. None of the interaction terms achieved significance in the final models.

Data preparation for these analyses proceeded in exactly the same fashion as in the tests of hypothesis 1. Screening on the basis of Mahalanobis distance resulted in the identification of 4 multivariate outliers at the .01 level that were removed from subsequent analyses. There were no unduly influential cases present in the data, as evidenced by a maximum CD_i of 0.23. Examination of Durbin-Watson values, predicted vs. residual plots, and Shapiro-Wilk tests of residual variance supported the tenability of statistical assumptions for all variables (with the exception of Sigma). The details of these violations during the analysis of Sigma are reported in the relevant section below. Bivariate correlations between the ex-Gaussian parameters, CPT variables, and age appear below in Table 6.

Table 6

Pearson Correlations Between ex-Gaussian Parameters, CPT Test Scores, and Age

	Age	μ	σ	τ	Omissions	Hit RT SE	Variability
Age	-	.603**	.142	.386**	.001	.040	-.122
μ		-	.407**	.564**	.165*	.149*	.023
σ			-	-.003	.022	-.018	-.052
τ				-	.363**	.369**	.208**
Omissions					-	.717**	.629**
Hit RT SE						-	.831**
Variability							-

Note: μ = $base_{10}$ logarithm of Mu parameter of ex-Gaussian distribution, σ = $base_e$ Logarithm of Sigma parameter of ex-Gaussian distribution, τ = $base_{10}$ logarithm of Tau parameter of ex-Gaussian distribution, Omissions = CPT Omisions, Hit RT SE = CPT Hit RT Standard Error, Variability = CPT Hit RT Standard Error Variability.
*$p<.05$
**$p< 0.01$

The correlations of the ex-Gaussian parameters and the CPT variables ranged from negligible (r = -.018, n.s.) to small (r = .369, p < .000). This set of bivariate correlations largely favored Tau. The correlation between Tau and omissions (r = .363, p < .000) as well as Tau and hit RT SE (r = .369, p < .000) were both small. Mu was negligibly correlated with these variables. In addition, Tau was the only ex-Gaussian parameter significantly related to hit RT SE variability (r = .208, p < .000). Sigma was not significantly related to any of the CPT variables. Intercorrelations between the CPT variables were generally high and suggest the presence of possible collinearity. Consequently, tolerance and variance inflation factor were calculated for all variables in the analyses and are reported below to assess the extent of any collinearity in the final models.

The final regression modeling response latency as a function of the CPT variables was significant ($F = 52.580$, $p < .000$) and successfully predicted 37.4% of the variance in Mu. Age and Omissions emerged as significant predictors of Mu. Age accounted for overwhelming portion of unique variance in the model (34.93%) with the additional modest contribution of Omissions (2.0%), as evidenced by the semipartial correlations of these two variables. In this regression, tolerance values were high (0.999) and variance inflation factor values were low (1.001), suggesting that multicollinearity did not pose a significant problem in the estimation of the final model.

Only age was significant in the model assessing Sigma and this variable accounted for a negligible portion of the variance in this parameter (1%). The low *t* values for the tests of regression coefficients associated with the CPT variables may be due, in part, to violations of underlying statistical assumptions. Heteroscedasticity can bias ordinary least squares estimates of the variance (and hence standard error) of regression coefficients, thus rendering tests for significance less reliable. This may result in either type 1 or type 2 errors in tests of regression coefficients. Alternatively, these results are consistent with previous reports that Sigma is only weakly associated with measures of attention (Tse et al., 2010), does not discriminate reliably between clinical groups (Rentrop et al., 2010), and is less affected by CNS pathology than Mu and Tau (Spieler, Balota, & Faust, 1996). With the caveat above, it is likely that Sigma's relationship to the CPT variables is either quite small or statistically insignificant.

In the model of intraindividual variability, age and hit RT SE significantly predicted 24.7% variance in Tau ($F = 28.823$, $p < .000$). Age received the strongest weight by a small margin followed by hit RT SE. These predictors accounted for similar

portions of variance due to their low intercorrelations, as Age uniquely accounted for 13.1% of the variance in Tau and hit RT SE uniquely accounted for 10.24% of the variance. This result demonstrates that intraindividual variability on the WCST, a complex abstract reasoning task, is related to variability on simpler tests of choice reaction time variability. Tolerance values were high (0.997) and variance inflation values were low (1.003) for both terms in the model, consistent with the interpretation that multicollinearity did not significantly affect the results. Omissions achieved statistical significance during the first step of the model but did not substantially contribute after the addition of hit RT SE (p = .078). This is somewhat surprising given previous associations between intraindividual variability and lapses of attention (e.g., West et al., 2002) and highlights the importance of task characteristics for the study of intraindividual variability. The results of this analysis are featured alongside those for Mu and Sigma in Table 7.

Table 7

Results of Regression Analyses (Final Models) Predicting Mu, Sigma, and Tau From CPT Test Performance

	μ					σ					τ				
	R^2	*b*	β	$r_{partial}$	r_{semi}	R^2	*b*	β	$r_{partial}$	r_{semi}	R^2	*b*	β	$r_{partial}$	r_{semi}
	.374					.010					.247				
Age		.009	.591	.599	.591		.009	.098	.098	.098		.007	.362	.385	.362
Omissions		.003	.143	.177	.142							-	-	-	-
Std. Error												.027	.321	.346	.320
Variability															

Note: μ = $base_{10}$ logarithm of Mu parameter of ex-Gaussian distribution, σ = $base_e$ logarithm of Sigma parameter of ex-Gaussian distribution, τ = $base_{10}$ logarithm of Tau parameter of ex-Gaussian distribution, R^2 = multiple correlation coefficient, *b* = unstandardized regression coefficient, β = standardized regression coefficient, $r_{partial}$ = partial correlation coefficient between criterion and predictor, r_{semi} = semipartial correlation coefficient between criterion and predictor, Omissions = CPT Omissions, Std. Error = CPT Hit RT Standard Error, Variability = CPT Hit RT Standard Error Variability. Parameters from final model only.

Visual Analyses of Hypothesis 2

Quantile plots of the CPT variables were constructed via the same methodology used to visually investigate hypothesis 1. In these analyses, cognitive test performance on the various CPT indices was ranked via demographically corrected *t*-score. The RT data of participants in the highest vs. lowest quartiles of performance were refit with QMPE and then averaged across participants to create quantile bands representing group level RT distributions. The results appear below in Figures 11 - 13.

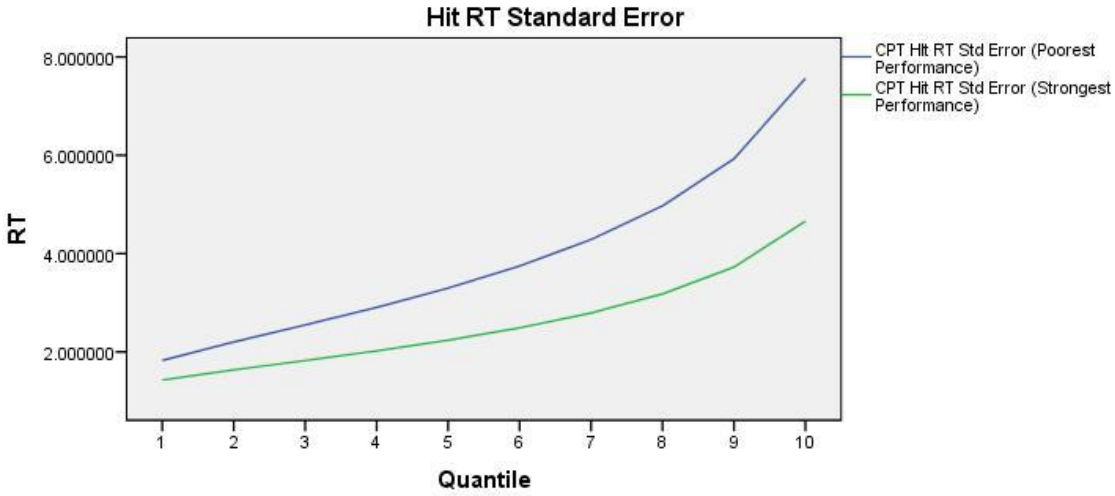

Figure 11. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on the CPT Hit RT Standard Error. RT displayed in seconds.

These plots generally confirm the relationships observed during statistical tests of hypothesis 2. For Hit RT SE, increasing distance between the quantile bands at more extreme values (i.e., acceleration) suggests a primary effect for Tau. Though there is some separation between the quantile bands in this plot (that might suggest an effect in Mu), it is minimal at faster RT values. The effect appears to be more widely distributed across the RT distribution than for the cognitive control variables, with acceleration evident at approximately the 6th quantile. There is a minimal effect of Mu, consistent with

the modest bivariate correlation of this parameter with Hit RT SE. There is no obvious effect of Sigma.

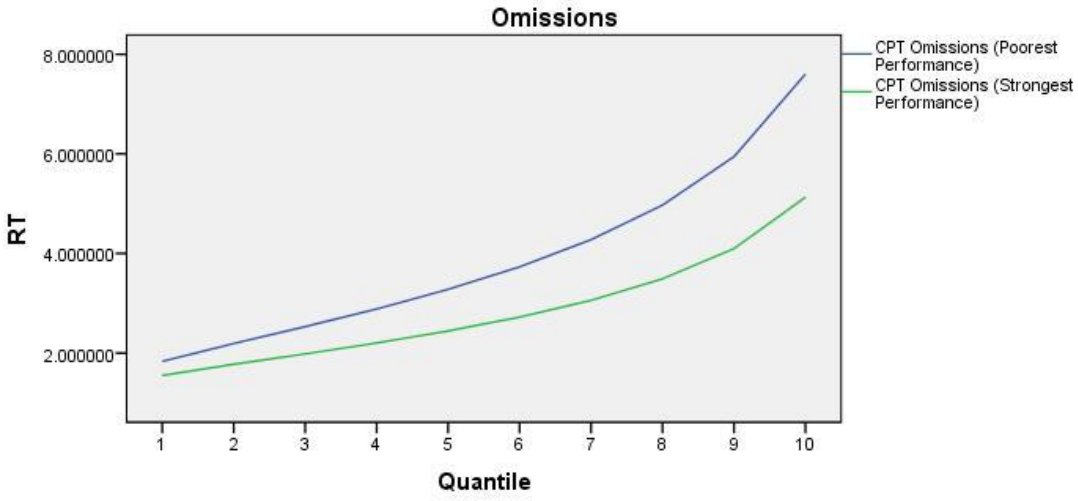

Figure 12. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on the CPT Hit RT Standard Error Variability (RT displayed in seconds).

For omissions (Figure 12) there is little visual evidence for an effect in Mu, which would present as separation in the quantile bands. Some positive acceleration (beginning perhaps at quantile 6 or 7) is evidence of an effect in Tau. There is no obvious effect for sigma. This plots highlights the fact that it is difficult to judge the point at which an effect becomes significant using visual RT distributional analysis, because the modest bivariate correlation of Mu with omissions is not readily apparent. The moderate correlation with Tau, however, is quite visible. In the setting of small effect sizes, definitively attributing an effect to Mu or Tau is further complicated by the statistical dependencies of these parameters when fit via quantile maximum likelihood estimation. Nonetheless, the plot is instructive in that it generally shows some degree of separation and acceleration (consistent with the statistical analyses) in the expected directions.

Finally, the plot for hit RT SE variability appears below in Figure 13. As with the plot for omissions, it is difficult to discern the precise nature of effects. Some positive acceleration appears in quantiles 7 to 10, consistent with an effect in Tau. The separation between the quantile bands is minimal and appears to be somewhat less than observed for the other CPT variables or the cognitive control tasks. Accordingly, it does not appear that Mu reliably differentiates between high vs. low performers ranked on the basis of this performance index. This result mirrors the findings of the correlational analyses and regression analyses, both of which failed to detect a significant relationship between Mu and hit RT SE variability.

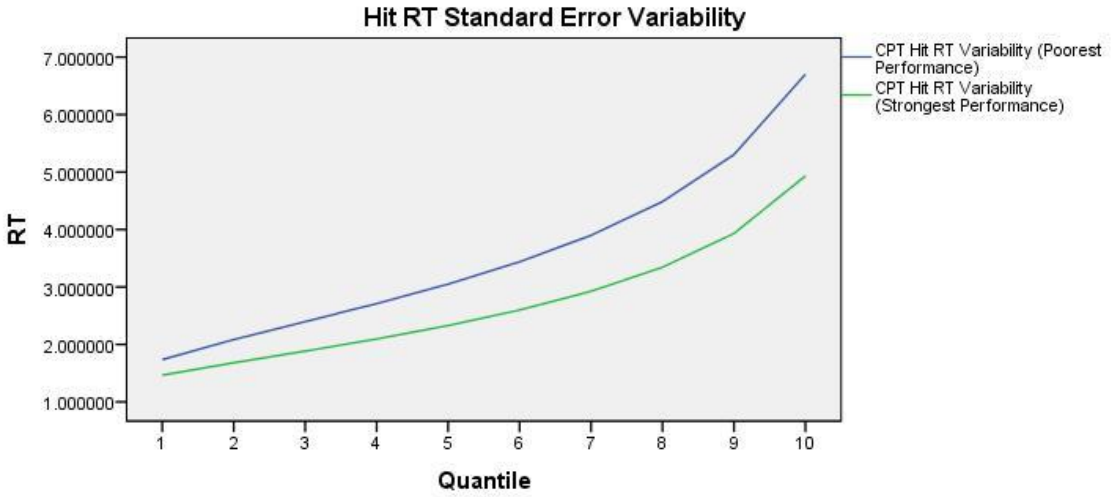

Figure 13. Quantile plot displaying the RT distributions of the highest vs. lowest quartiles of performance on the CPT Hit RT Standard Error Variability (RT displayed in seconds).

Regression Analyses of Hypothesis 3

It is possible that the significant relationships observed between response latency, intraindividual variability, and the cognitive control tasks were simply the results of individual differences in basic attentional functions rather than higher-order cognitive skills. A series of linear regression models were constructed to examine this possibility.

These models were specified in three blocks: Age was entered as a forced covariate in the first block, the CPT (i.e., stage one) variables were introduced via stepwise variable entry during the second block, and the cognitive control (i.e., stage two) tests were then entered in a third block (also employing stepwise variable entry). The resulting models test the relationships between the ex-Gaussian parameters and cognitive control tasks, controlling for the effects of basic attentional skill. Because Sigma was not significantly predicted by any of the stage one CPT variables, the results of any analyses under hypothesis three would be functionally identical to those under hypothesis one. As a result, this variable was not investigated under hypothesis three.

To facilitate comparisons with tests of hypothesis one and two, the four outlying cases from hypothesis two (identified on the basis of D^2) were eliminated from the following analyses. Fidelity to underlying statistical assumptions was evaluated by the same methodology employed in the other hypothesis tests. Durbin-Watson values ranged from 1.724 (for tests of Tau) to 2.087 (for tests of Mu), suggesting that independence of errors was tenable. Visual plots of standardized predicted values vs. residuals revealed relative homoscedasticity (with the exception of Sigma). Shapiro-Wilk tests for normality of residual variance were .997 for Mu (p = .968) and .995 (p = .752) for Tau. In sum, statistical assumptions were supported for all variables. None of the interaction terms were significant for any of the models.

Table 8

Pearson Correlations Between Cognitive Control Test Scores and CPT Variables

	Omissions	Hit RT SE	Variability	TMT-B	Stroop	COWAT	WCST
Omissions	-	.642**	.654**	.235**	-.208**	-.155*	.225**
Hit RT SE		-	.874**	.307**	-.321**	-.198**	.218**
Variability			-	.223**	-.207**	-.194**	-.154*
TMT-B				-	-.495**	-.402**	.491**
Stroop					-	.289**	-.332**
COWAT						-	-.300**
WCST							-

Note: Omissions = number of CPT omission errors, Hit RT SE = CPT hit RT Standard Error, Variability = CPT hit RT Standard Error Variability, TMT-B = Trail Making Test B (total seconds to completion), COWAT = Total words generated to the letters FAS, Stroop = Total words named on color-word trial of Stroop Test, WCST = WCST perseverative errors score.
*$p < 0.05$
**$p < 0.01$

Bivariate correlations between the cognitive control tests and CPT variables are depicted in Table 8. The CPT variables were significantly related to the cognitive control tests in the expected direction, with bivariate Pearson correlations ranging from r = -.154, (p = .040, hit RT SE variability and perseverative errors) to r = -.321 (p < .000; hit RT SE and Stroop). By standard conventions, these represent negligible to small effects. Correlations were modestly higher between hit RT SE and the cognitive control tasks than they were for the other CPT variables. As reported above under tests of hypothesis 2, intercorrelations were moderate to high between the CPT variables.

The final regression model predicting response latency from the cognitive control tests and CPT variables was significant (F = 50.558, p < .000) and accounted for 54% of

the variance in Mu. In addition to age, both perseverative errors and TMT-B emerged as significant predictors of Mu (in order of entry). The effect of omissions was significant during the second step of the model (p = .021), but failed to achieve significance following the addition of TMT-B and perseverative errors (p = .625) during the third step. This pattern of results suggests that the modest correlation observed between Mu and Omissions in the second step of the model was primarily the result of a small portion of variance that this performance index shared with TMT-B and perseverative errors. Thus, it does not appear that basic attentional functions mediate the relationship between response latency and cognitive control.

Examination of R^2 change values reveals that the block containing the cognitive control tests accounted for 16.2% of the variance in Mu controlling for the effects of stage one functioning. In the final model, age accounted for the majority of unique variance in response latency (22.09%), followed by perseverative errors (5.19%) and TMT-B (4.04%). The contribution of omissions was modest, as this measure of simple attention and orientation only accounted for .06% unique variance. Qualitatively speaking, poorer performance on TMT-B and perseverative errors resulted in greater response latency after controlling for basic attentional functioning.

The model predicting intraindividual variability was also significant (F = 37.33, p < .000) and accounted for 46.5% of the total variance in Tau. Age, hit RT SE, TMT-B, and perseverative errors emerged as significant predictors of Tau (in order of entry in the final model). As expected, increasing age and poorer performance on the neurocognitive tests was associated with increased variability. The cognitive control variables took up the greatest portion of unique variance in the final model, as they jointly accounted for

21.8% of the variance in Tau (on the basis of R^2 change values). The squared semipartial correlations demonstrate that perseverative errors uniquely accounted for the greatest proportion of variance in Tau (7.08%) followed by TMT-B (5.48%). The unique contribution of age was 5.2% and the contribution of hit RT SE was 2.79% of the unique variance in Tau. Thus, it appears the relationship between intraindividual variability and cognitive control is partially mediated by stage one working memory processes (as measured by the CPT).

In contrast to the analyses of mean response latency, controlling for stage one functioning decreased the observed correlations between intraindividual variability and cognitive control. The total unique contribution of the cognitive control variables was reduced from 30.3% unique variance in the model under hypothesis 1 (i.e., without controlling for stage one functions) to 21.8% unique variance in the model controlling for stage one functions. This suggests that a portion of the relationship between cognitive control and intraindividual variability is mediated by more basic attentional processes. Additionally, part of the relationship between hit RT SE and Tau may be due to method variance (discussed in detail below). The results of these analyses appear below in Table 9.

Table 9

Results of Regression Analyses (Final Models)Testing the Relationship Between the ex-Gaussian Parameters and Cognitive Control (Controlling for Basic Attentional Processes)

	μ					τ				
	R^2	*b*	β	$r_{partial}$	r_{semi}	R^2	*b*	β	$r_{partial}$	r_{semi}
	.540					.465				
Age		.007	.486	.569	.470		.004	.236	.298	.228
Stage One										
Omissions		.001	.026	.037	.025		-	-	-	-
Std. Error		-	-	-	-		.015	.177	.223	.167
Stage Two										
WCST		.004	.266	.319	.228		.005	.308	.341	.266
TMT-B		.001	.235	.284	.201		.002	.278	.304	.234

Note: μ = $base_{10}$ logarithm of Mu parameter of ex-Gaussian distribution, τ = $base_{10}$ logarithm of Tau parameter of ex-Gaussian distribution, R^2 = multiple correlation coefficient, *b* = unstandardized regression coefficient, β = standardized regression coefficient, $r_{partial}$ = partial correlation coefficient between criterion and predictor, r_{semi} = semipartial correlation coefficient between criterion and predictor, Omissions = CPT Omissions, Std. Error = CPT Hit RT Standard Error, WCST = WCST Perseverative Error Score, TMT-B = Trail Making Test B.

Regression Analyses of Hypothesis 4

Previous efforts have found significant relationships between intraindividual variability and abstract reasoning (Schmiedek et al., 2007). Hypothesis 4 tested this possibility through specifying a series of regression models that predicted the ex-Gaussian parameters on the basis of total categories completed on the WCST, perhaps the most direct measure of abstract non-verbal problem solving on this rule based categorization task (Heaton et al., 1993). The models were constructed by entering age as a forced covariate in the first step, followed by categories completed. To evaluate the extent to which any observed relationships were influenced by basic attentional processes, analyses were re-run by including the CPT variables in the second block of the models with age and categories entered in the first and third blocks, respectively.

Computing Mahalanobis Distance and Cook's Distance did not reveal the presence of any outliers for the primary tests of the relationship between the ex-Gaussian parameters and categories completed. Four outliers were identified on the basis of Mahalanobis Distance in the models controlling for stage one functioning, but they were not removed in order to allow direct comparisons between the models with and without these variables. Durbin-Watson values, which ranged from 1.924 (Mu) to 2.069 (Sigma), supported the assumption of independently distributed error terms. Normality of residual variance (assessed via Shapiro-Wilk tests of residuals) was tenable with the exception of Sigma (.793, $p < .000$) and Tau (.985, $p < .000$). For Tau, however, the corresponding Komolgorov-Smirnov value of .054 suggested that any violations of normality were quite modest ($p = .200$). Visual inspection of predicted vs. residual values revealed relative

homoscedasticity. Finally, there were no significant interaction effects in the final models.

Bivariate correlations between the CPT variables, age, and WCST categories completed were significant; however, they were generally low. Numerically, correlations ranged from r = -.210 (for categories and hit RT SE variability) to r = -.346 (for categories and hit RT SE). By standard conventions, these are negligible to small effects. As in other analyses, hit RT SE (a measure of choice RT variability) bore the highest relationship to higher order executive functioning (in this case WCST categories completed). These results suggest that the CPT variables may serve (at most) as partial mediators of the relationship between cognitive control and the RT parameters. These bivariate correlations appear below in table 10.

Table 10

Pearson Correlations Between Categories Completed Age, and CPT Variables

	Age	Omissions	Hit RT SE	Variability
WCST	-.214**	-.314**	-.346**	-.210**

Note: Omissions = CPT omission errors, Hit RT SE = CPT RT Standard Error, Variability = CPT hit RT Standard Error Variability, WCST = Wisconsin Card Sorting Test number of categories completed.
** $p < .01$

Both age and categories completed were significant predictors of response latency (F = 87.640, $p < .000$) and accounted for 49.2% of the variance in Mu. Increasing age and fewer categories completed resulted in greater response latency. As in other tests of Mu, age accounted for the greatest proportion of unique variance in the model (26.21%) followed by categories completed (12.82%). Controlling for stage one working memory (i.e., CPT omissions) had little effect on the overall model (R^2 = 49.4) or the variance proportions. Overall, these findings suggest that mean response latency on the WCST is

related to abstract problem solving, and that this relationship is not mediated by more basic attentional processes.

Intraindividual variability was also related to higher order abstract reasoning. Both age and categories completed significantly predicted Tau in the final model (F = 53.407, $p < .000$). As expected, older age and poorer performance on categories completed were associated with a higher level of inconsistency. The solution including these variables accounted for 37.1% of the variance in Tau, the majority of which was uniquely predicted by categories completed (22.28%). The semipartial correlation of age revealed that this variable uniquely predicted 7.56% of the variance in Tau. Controlling for the effects of stage one processing increased the predictive power of the total model (R^2 = .419), reduced the unique contribution of categories completed to 14.14% of the variance, and increased the proportion of variance that was uniquely predicted by age to 8.35%. In this final model, Omissions was the only significant stage-one variable ($p <$ .000). Though hit RT SE achieved significance during the third stage of the model, it was rendered insignificant following the entry of omissions. These findings reveal partial mediation of the relationship between intraindividual variability and abstract reasoning by basic stage one attentional functions (consistent with predictions on the basis of theory; e.g., Christoff & Gabrielli, 2000; Owen, 2000).

In the final model, omissions was the only significant stage one variable. By contrast, hit RT SE was significant when testing mediation of cognitive control. This may be due to the inclusion of the 4 outlying cases that were eliminated in the latter analysis. This is of little practical significance, given that omissions and hit RT SE are nearly interchangeable as predictors of intraindividual variability (on the basis of their bivariate

correlations with Tau, reported above). Sigma was not significantly related to abstract non-verbal problem solving. The results of the final analyses (excluding the CPT variables) are featured in Table 11 below.

Table 11

Results of Regression Analyses (Final Models) Testing the Relationship Between the ex-Gaussian Parameters and Abstract Non-Verbal Reasoning

	μ					τ				
	R^2	*b*	β	$r_{partial}$	r_{semi}	R^2	*b*	β	$r_{partial}$	r_{semi}
	.492					.371				
Age		.008	.524	.583	.512		.005	.282	.328	.275
WCST		-.039	-.367	-.449	-.358		-.063	-.483	-.511	-.472

Note: μ = $base_{10}$ logarithm of Mu parameter of ex-Gaussian distribution, τ = $base_{10}$ logarithm of Tau parameter of ex-Gaussian distribution, R^2 = multiple correlation coefficient, *b* = unstandardized regression coefficient, β = standardized regression coefficient, $r_{partial}$ = partial correlation between criterion and predictor, r_{semi} = semipartial correlation between criterion and predictor, WCST = WCST Categories Completed

Analysis of Task Specific Variables

In investigations of RT data, a variety of potential confounds must be ruled out before meaningful conclusions can be drawn. General slowing, levels of task engagement, unspecified task specific factors, and differences in response threshold all represent significant threats to inference in RT investigations. Analyzing specific task characteristics (e.g., accuracy) allows for conclusions about the plausibility of these scenarios. Bivariate correlations between the ex-Gaussian parameters, total number of correct responses, total number of incorrect responses, and accuracy on the WCST are reported below.

Average task accuracy was 67%. By contrast, task accuracies typically exceed .90% on choice reaction time tasks and other similar measures frequently employed by variability researchers. This suggests that the WCST is much more difficult than similar other tasks used to study variability. Accuracy was negatively related to Mu (r = -.522, p < .000), Sigma (r = -.156, p = .034), and Tau (r = -.567, p < .000). Increased response latency, greater dispersion in the modal portion of the distribution, and greater variability resulted in a higher proportion of incorrect responses, which is expected given the results reported above. Slower, more variable individuals also sorted a greater number of cards incorrectly and a lower number of cards correctly. For total correct responses, bivariate correlations were small for Mu (r = -.332, p < .000), negligible for Tau (r = -.240, p < .000), and non-significant for Sigma (r = -.025, p = .737). For total errors, correlations were moderate for Mu (r = .503, p < .000) and Tau (r = .557, p < .000) and negligible for Sigma (r = .158, p = .032).

Power Analysis

The power analyses conducted during the design of this study are not reported. Post hoc power analyses, however, were conducted with G*Power 3.0 (Faul, Erdfelder, Lang, & Bruchner, 2007). The actual power of main regression equations was computed by submitting α levels, observed effect sizes, sample size, and the number of predictors to analysis with G*Power. Results of these computations indicate excellent power for nearly all analyses involving Mu and Tau. In hypothesis one, tests of the deviation of ρ^2 from zero achieved an actual power of 1.0 for Mu and Tau. Due to the lower effect size, tests of Sigma achieved an actual power of .80, which meets the level recommended by Cohen (1988). Tests of R^2 increase (i.e., testing the power to detect whether terms should be added to the model) achieved a power of 1.0 for tests of Mu and Tau and .75 for Sigma.

The hypothesis two analyses were powerful for Mu and Tau, with tests of the deviation of ρ^2 from 0 achieving an actual power of 1.0. Due to the very small effect size (f^2 = 0.01) tests of the deviation of ρ^2 from 0 were underpowered for Sigma (1-β= .17). Tests of R^2 increase achieved an actual power of 1.0 for Mu, .99 for Tau, and .15 for Sigma. For analyses of hypothesis three, tests were powerful despite the presence of a greater number of predictors. The actual power achieved was 1.0 for Mu and Tau when testing the deviation of ρ^2 from 0 as well as R^2 increase. Sigma was not investigated under hypothesis three.

Tests of hypothesis 4 were similarly powerful with evaluation of deviation of ρ^2 from 0 and R^2 increase both achieving a power of 1.0 for Mu and Tau. Tests of Sigma achieved a power of .50 for deviation of ρ^2 from 0 and R^2 increase. The mediating

analyses for Mu and Tau included the CPT variables achieved power of 1.0 for testing the deviation of ρ^2 from 0 and R^2 increase. Sigma was not investigated in mediating analysis. Overall, these data indicate that the majority of analyses achieved power that substantially exceeded Cohen's (1988) recommendation of .80 for confidence in the experimental design. Achieving greater power during tests of Sigma would have required a prohibitively large sample due to the small observed effect sizes.

CHAPTER 5

DISCUSSION

Summary and Review

Intraindividual variability refers to individual changes that occur across trials, measurement occasions, or tasks. Such variability may represent enduring functional adaptations that enhance the ability of an organism to interact functionally with the environment (i.e., *developing;* Li et al., 2004). Alternatively, it may represent fluctuation around an asymptote of optimal performance in the face of challenges or inability to sustain an ongoing stream of adaptive behavior (*functioning*). Researchers have used individual differences in maladaptive forms of variability to predict impending death (Shipley, Der, Taylor, & Deary, 2006), improve the accuracy of dementia diagnosis (Holtzer et al., 2008), and differentiate clinical populations from normal controls (Leth-Steensen et al., 2000). In the setting of response time research, fluctuation in performance across trials serves as a harbinger of deficiency, as it predicts impairment on neuropsychological measures before it is evident in mean levels of performance (Lövdén et al., 2007).

There is now a significant body of evidence revealing that the integrity of various prefrontal cortical regions is necessary for performance consistency. Studies of focal brain lesions (Stuss et al., 2003), dementia subtypes (Murtha et al., 2002), functional neuroimaging (Bellgrove et al, 2004), structural neuroimaging (Bunce et al., 2004; Bunce et al., 2007), and genetic determinants (Stefanis et al., 2005) all converge on this interpretation. The literature suggests that variability may be a sensitive, novel measure

of executive control processes. Unfortunately, the scientific study of variability has had little impact on the mainstream practice of clinical neuropsychology.

The primary goal of this investigation is to bridge this gap through ascertaining the extent to which trial-to-trial variability in response time is related to cognitive control and higher order executive abilities, thereby establishing its convergent validity with other more traditional neuropsychological assessment techniques. This effort stands to contribute to our existing knowledge in several capacities. First, this study is one of a few to provide convergent evidence that intraindividual variability on commonly used assessment measures is related to other clinical tasks of executive functioning. Second, it is the first study to use a battery of cognitive control tests whose neuroanatomical basis has been clarified by the results of a large scale cortical lesion mapping study (Gläscher et al., 2012) and several decades of functional neuroimaging research, thereby allowing the explicit relation of findings to dominant neuroanatomical theories of working memory and executive functioning (Christoff & Gabrielli, 2001; Owen, 2000; Petrides, 1995). Third, it is one of a few studies to investigate intraindividual variability in a heterogeneous clinical population. Finally, it is the only study (to the author's knowledge) to examine intraindividual response time variability on the Wisconsin Card Sorting Task.

Four primary hypotheses were investigated during the course of this project. First, greater levels of intraindividual variability will be associated with lower levels of performance on cognitive control tests. Second, increased attentional control will predict lower levels of intraindividual variability. Third, basic attentional abilities will partially mediate the relationship between cognitive control and intraindividual variability. Fourth,

intraindividual variability will be related to higher-order aspects of executive functioning, such as abstract reasoning.

The following discussion section considers each of these hypotheses in detail through exploration of the study results and linking the findings to neuroanatomical data and theory about the higherarchical organization of the prefrontal cortex. Main findings are reviewed and alternative explanations are examined before a discussion of the theoretical and neuroanatomical significance of the findings. These sections primarily highlight the strengths of the study. Detailed consideration of methodological and statistical limitations follows. Next, suggestions for future research are offered. The project concludes with a brief statement about the implications of this investigation for the practice of clinical neuropsychology.

Review and Interpretation of Significant Results

In this study, the primary research hypotheses were supported. Intraindividual variability on the WCST was significantly related to cognitive control (as well as simple attention) and individual differences in basic attentional ability partially mediated this relationship. Intraindividual variability also appeared to be related to non-verbal problem solving, a higher-order aspect of executive functioning. Both statistical and visual analyses were consistent with these interpretations. Importantly, visual analyses were conducted with demographically corrected *t*-scores, which revealed that these significant results do not simply reflect the contribution of uncontrolled demographic variance. Demographically corrected *t*-scores were not used in the regression analyses because published normative data for the ex-Gaussian parameters is not available. Thus, utilizing

demographically corrected scores would have effectively removed the effects of age and education from the predictor variables but not the criterion variables.

With respect to the first study hypothesis, lower levels of performance on the cognitive control tests predicted increased intraindividual variability. TMT-B and perseverative error score emerged as the best (and only) significant predictors of variability after controlling for age. Interpretation of the correlation between Tau and perseverative errors must be tempered somewhat, however, due to the fact that both indices were derived from the WCST. The potential presence of content similarity might bias estimates of their true correlation. This possibility is examined in more detail below in the section on alternative explanations.

By Cohen's (1988) standards, the size of the overall regression effect predicting Tau was large (f^2 = .821) and the vast majority of unique variance in the model was predicted by the cognitive control tests (as opposed to age).Together, perseverative errors and TMT-B accounted for 30.3% of the variance in intraindividual variability. The effect size for the regression equation predicting Mu on the basis of cognitive control was even larger (f^2 = 1.22). For response latency however, the cognitive control tests contributed somewhat more modestly (18.1% combined) and age accounted for a greater portion unique variance in the model (21.7%). In comparing and contrasting the relative portions of variance for the models of Mu vs. Tau, one must bear in mind that the average correlation between these two parameters was r = -.51 (partially due to parameter depenencies). Thus, the methodology employed does not allow for definitive statements about precise variance portions. Nevertheless, cognitive control predicted nearly twice

the variance in Tau that it did in Mu, a finding that is consistent with previous investigations.

The use of commonality coefficients, which allow researchers to partition the portions of a regression effect that can be exclusively attributed to each predictor and each combination of predictors (Nimon, 2010) further clarifies these findings. The combined unique and common effects of the cognitive control variables accounted for 67.11% of the variance in the final regression model predicting Tau ($\gamma_{TMT\text{-}B}$ = .1064, γ_{WCST} = .0631, $\gamma_{TMT\text{-}BxWCST}$ = .1338; 30.3% of the overall variance in Tau). By contrast, these tests accounted for only 33.12% of the variance in the effect for Mu ($\gamma_{TMT\text{-}B}$ = .0460, γ_{WCST} = .0536, $\gamma_{TMT\text{-}BxWCST}$ = .0803; 18.1% of the overall variance in Mu). These findings reveal that the majority of variance in the regression for Mu reflects the contribution of age, which is expected on the basis of theory. Numerous investigations support the notion that general slowing is responsible for age related changes in response times (e.g., Myserson, Hale, Wagstaff, Poon & Smith, 1990; Zheng, Myserson, & Hale, 2000). In sum, both intraindividual variability and response latency are meaningfully related to executive tests measuring attentional set shifting; response latency, however, is is more significantly related to age.

An interaction effect emerged during tests of hypothesis one. Specifically, the relationship between intraindividual variability and the cognitive control tasks varied as a function of test performance (Figure 4). For a given neurocognitive test, its positive association with intraindividual variability decreased in strength with increasing levels of impairment on the other cognitive test. In the case of individuals with the poorest performance on TMT-B, the relationship between intraindividual variability and

perseverative errors actually reversed. In other words, individuals who performed poorly on either cognitive task might or might not have performed poorly on the other task; they were all more variable. The significance of these correlations was not investigated statistically due to sample size restrictions.

The importance of this finding is unclear as it may be due to several factors. First, because all individuals with poor performance on either cognitive task were considerably more variable than their unimpaired counterparts, a restriction of range may have ocurred, thus making this interaction effect less meaningful. If this were the case, the change in strength and direction of the relationship between variability and the cognitive control tasks would be a statistical artifact. Second, individuals who were impaired on TMT-B may have engaged in a more consistent but still comparatively ineffective response strategy on the WCST, accounting for the negative relationship between perseverative errors and intraindividual variability at impaired levels of performance on the TMT. This is less likely, however. The numerous findings reported above suggest that variability accompanies lower levels of skill across the spectrum of performance on the cognitive control measures.

The third and perhaps most interesting interpretation is that intraindividual variability may reflect deficiency in a general underlying neural substrate, making it sensitive to impairment yet not highly specific. This cross-task factor could account for impairment on either TMT-B or perseverative errors before a level of disability sufficient to impair performance on both tasks was reached. Other work attests to the fact that individuals may become more variable before impairment is evident in their mean level of performance (Lövden et al., 2007), which is consistent with this conceptualization.

In this scenario, TMT-B and perseverative errors would be more specific measures, which would account for the finding that some individuals who were impaired on one test were not impaired on the other. Variability would be reflective of a more general process, thus explaining why individuals who performed poorly on either task were more inconsistent. Previous investigations have found that RT variability reflects both the severity of cognitive impairment and the specific nature of the neurological disturbance (Burton et al., 2006), which is consistent with this possibility. Investigation of the extent to which variability on the WCST represents specific vs. general impairment (as well as sensitivity/specificity data for the detection of impairment) might be a fruitful area for future research.

Hypothesis two was also supported, as variability was related to more basic attentional functions measured by the CPT variables. Specifically, Tau was predicted by hit RT SE (a prototypical measure of RT variability). This performance index accounted for 10.24% of the variance in Tau and the overall effect size for the regression equation (predicting Tau on the basis of age and hit RT SE) was medium by Cohen's (1988) standards (f^2 = .328). It is worth mentioning that omissions indexed nearly the same portion of unique variance, as it accounted for 8.64% of Tau after rebuilding the regression equation with this measure in place of hit RT SE. The fact that hit RT SE was included in the final model instead of omissions may reflect the atheoretical stepwise variable entry procedure that was employed (rather than theoretically significant differences). Mu was also related to the CPT variables, but they accounted for a miniscule portion of variance in this parameter.

Methodologically, this finding demonstrates that trial-to-trial variability on the WCST is related to trial-to-trial variability on Go/No-go tests such as the CPT. Other research has documented minimal differences in variability across an assortment of relatively distinctive RT tasks (Burton et al., 2006) in certain populations, which supports the notion that it is a "trait-like" individual difference. For example, those who are more variable on one occasion are likely to be more variable on subsequent measurement occasions (Hultsch et al., 2000). In the present study, variability on one measurement occasion (on the WCST) predicted variability during another measurement occasion (on the CPT) within the same testing session. These results support the findings of Burton et al. (2006) and Hultsch et al. (2000). The correlation between Tau and hit RT SE is also consistent with the interpretation that variability (as measured by the Tau parameter) is, in part, analogous to intraindividual variability indexed by more traditional methods such as the iSD.

There is an important caveat to this finding, as both CPT hit RT SE and Tau may share method variance, given that they are both measures of trial-to-trail response time variability. This variance might include peripheral, non-cognitive contributions that could spuriously inflate the observed magnitude of the relationship between Tau and CPT hit RT SE. Despite some similarity, however, the instructions for the two tasks are significantly different. Whereas individuals are instructed to respond as quickly and accurately as possible during the CPT, they are instructed to sort the cards to the best of their ability without regard to time on the WCST. Moreover, researchers have found that measures of intraindividual RT variability correlate poorly with other measures of simple motor functioning (e.g., tapping speed; de Frias et al., 2007) and that group differences in

cognitive functioning are more predictive of intraindividual variability than are differences in peripheral motor abilities (Hultsch et al., 2000). Thus, the results may not be overly biased by similarity between the two tasks.

A substantial proportion of WCST RT variability was not explained by the CPT variables, which may be due to a number of factors. First, methodological differences in the way intraindividual variability was indexed between the WCST (i.e., Tau parameter) and CPT (i.e., iSD of log-transformed RT distribution) undoubtedly contribute to the lack of correspondence between these measures. To the author's knowledge, there are no previous reports of the extent to which Tau parameters and iSDs converge and as such, speculations about the significance of this methodological disparity would be reduced to conjecture. Quantifying the consistency between these different techniques is important to examine in future research. Second, as a measure of abstract non-verbal problem solving, the WCST imposes considerably greater executive and cognitive control demands on the examinee as compared to the CPT. This substantive difference might easily give rise to different patterns of variability. Researchers have found that previously significant effects may be rendered insignificant when they are investigated using functionally different intraindividual variability paradigms (Christensen et al., 2005). Nevertheless, the present findings add to the growing body of evidence that intraindividual variability on different tasks is both related to and distinct from intraindividual variability on other procedures.

At the level of theory, the CPT variables in this study are conceptualized as operationalizations of more basic attentional functions that are likely to reflect first-order working memory processes (Petrides, 1995; Christoff & Gabrielli, 2001). First-order

functions are those that maintain rules involving a small number of concrete item properties (Badre, 2005). For example, decisions about the color of a stimulus, the identity of a letter, or judgments about the occurrence or non-occurrence of a stimulus. The VMPFC appears to be the neuroanatomical basis for first-order cognitive control operations (Christoff & Gabrielli, 2000; Owen, 2000). This brain region is also associated with the orienting response, a "subprocess signaling the active orientation of attention towards potentially significant events" (Williams et al., 2000, p. 3011). Orienting is indexed by omissions on the CPT (Conners, 2000).

This level of executive functioning (i.e., first-order processes) has served as the basis of interpretation in previous studies of the relationship between Tau and higher-order cognitive abilities (e.g., Balota, Cortese, Sergeant-Marshall, Spieler, & Yap, 2004; Geurts et al., 2008; Rentrop et al., 2010; Tse et al., 2010; West et al., 2002; Leth-Steensen et al., 2000). Researchers have frequently made use of an attentional control framework to attribute their significant findings to lapses of intention (West et al., 2002), lapses of attention (Rentrop et al., 2010), or changes in the efficiency of attentional control systems (Tse et al., 2010).

The findings under hypothesis 2 are consistent with these previous results, yet the fact that basic attentional functioning (indexed by the CPT variables) did not account for the vast majority of the relationship between cognitive control and response time variability is not. Furthermore, intraindividual variability was only modestly related to Omissions (r = .363), a direct measure of failure to orient/attentional lapses. The complex nature of the tasks employed means that these results provide only a weak challenge to previous, more parsimonious attributions of increased RT variability to simple attentional

functions. Still, it is significant that a prototypical task of vigilance and attentional functioning failed to predict a greater portion of the variance in Tau.

Hypothesis 3 directly tested the possibility that differences in simple attentional abilities were sufficient to explain the observed relations between intraindividual variability and higher order executive functions. Controlling for first-order functioning (indexed by CPT hit RT SE) significantly affected the observed relationship between cognitive control and Tau. As expected, this relationship was reduced but it was not eliminated. With the effects of first-order functioning controlled, TMT-B and perseverative errors continued to account for 21.8% of the variance in intraindividual variability. By contrast, they accounted for 30.3% unique variance when hit RT SE was left out of the model. Thus, in this investigation basic (i.e., first-order) executive processes serve no more than a modest mediating function. This finding is particularly significant considering the fact that the CPT is not, strictly speaking, a pure measure of first-order functioning.

Despite the CPTs failure to mediate the majority of shared variance between Tau and the cognitive control tasks, it did make an appreciable contribution to the prediction of Tau in its own right. Hit RT SE predicted 10.2% of the variance in Tau during tests of hypothesis 2 and mediated 8.5% of the variance that was formerly attributable to the cognitive control tests during tests of hypothesis 3. An analysis of the cognitive skills required by the CPT sheds light on the abilities that these findings may reflect.

The CPT involves demands such as judging whether a given letter matches an X, conflict monitoring during incorrect trials, updating the contents of working memory as the test progresses, and inhibition on No-go trials, all of which represent more higher

order cognitive skills (Riccio et al, 2002). Indices directly measuring conflict trials (e.g., commissions) were omitted from these analyses in order to minimize the extent to which results were confounded by these other demands. Still, the CPT is considerably more involved than many procedures used to index first-order executive processes. This may be the reason that neuroimaging studies demonstrate engagement of the dorsomedial PFC and caudal PFC in addition to the VLPFC during inhibition trials of Go/No-go tasks such as the CPT (Ridderinkhof, van den Wildenberg, Segalowitz, & Carter, 2004). Depth-electrode recordings in non-human primates show that inhibition trials are accompanied by firing of cells in the principle sulcus (the monkey homologue of DLPFC, Brodmann's area 46; Sakagami, Tsutsui, Lauwereyns, Koizumi, Kobayashi, & Hikosaka, 2001). For Conners' CPT specifically, neuroimaging data reveals widespread activation of various portions of the PFC thought to reflect second-order aspects of executive functioning in addition to VMPFC (Ogg et al., 2008). The CPT is thus likely to be partially confounded with higher-order aspects of cognitive control, making analyses controlling for CPT performance only partially reflective of the differential contributions of first order vs. second order executive functions.

Finally, in tests of hypothesis 4, intraindividual variability was related to abstract reasoning, a putative third-order frontopolar executive ability (Christoff & Gabrielli, 2000; Owen, 2000; Petrides, 1995). The overall effect size for this analysis was large (*f2* = .589) and categories uniquely predicted 22.28% of the variance in Tau controlling for age. This contribution was reduced to 14.14% after controlling for first-stage functioning (CPT Omissions). Response latency was also related to abstract reasoning and the effect size for this overall regression effect was also large (*f2* = .969). In the latency models,

categories accounted for 12.82% of the unique variance in Mu. Controlling for first-order functioning (i.e., omissions) did not significantly reduce its contribution. The relationship between Tau and abstract reasoning is expected on the basis of previous literature. The findings with respect to Mu are also in line with the results of previous studies. In their analysis of ex-Gaussian parameters, Schmiedek and colleagues (2007) found $r = .38$ between a latent Mu variable and latent reasoning construct, but Mu was not a unique significant predictor in the setting of the other ex-Gaussian parameters. This possibility was not investigated in this study.

In summation then, intraindividual variability was related to cognitive control; specifically, tasks requiring set shifting. It was also related to more basic (primarily) first-order executive functions (indexed by the CPT). Tau and Mu were associated with abstract non-verbal problem solving, a purportedly higher-order rostral executive ability. In tests of mediation, first-order processes served no more than a moderate mediating function of the relationship between intraindividual variability and the higher-order cognitive skills. This suggests that intraindividual variability is related to specific aspects of higher order executive cognition and that this relationship is not explained by simple lapses of attention or basic fluctuations in attentional control.

Evaluation of Alternative Explanations

In RT research, it is important to examine more parsimonious methodological/statistical explanations of significant findings before making statements about the relations between RT parameters and higher order aspects of cognition. This section considers alternative explanations in detail and assesses their plausibility. Specifically, the results may be caused by at least four alternative factors.

First, the findings could simply be the by products of general slowing. If individuals with lower levels of cognitive control are generally slower, then they are more likely to occasionally produce extremely slow responses that would increase Tau values and account for the significant effects. Slowing is expected to effect all RTs equally. Consequently, the RT distributions of participants with low levels of cognitive control would simply reflect a linear transformation of the distributions of participants with high levels of cognitive control. Such a linear transformation would uniformally effect all three ex-Gaussian parameters (Schmiedek et al., 2007) and cognitive control would be an equally good predictor of Mu, Sigma, and Tau. The fact that Mu and Tau were the only parameters robustly related to TMT-B, perseverative errors, and categories provides clear evidence against such an interpretation. Although Sigma was significantly related to perseverative errors, this cognitive control test predicted only a minute portion of this parameter.

Another way to test whether the results are due to general slowing is to construct Q-Q plots of the RT distributions for individuals with high vs. low levels of cognitive control (Myerson, Adams, Hale, & Jenkins, 2003). The resulting regression line modeling the relationship between the quantiles of the two groups can be interpreted as an indicator of the extent to which they are simply linear transformations of one another. Nonlinear relationships between the quantiles attest to differential RT distribution shape as a function of group membership and are inconsistent with a general slowing interpretation.

To carry out these analyses, participants were ranked in order of their performance (on the basis of demographically corrected *t* - scores) on the cognitive control tests. High and low cognitive control were defined as the upper and lower

quartiles of performance for each of the assessment tasks. Quantiles were calculated for each participant individually and then averaged (within groups), to form a set of mean quantile points for each group and each cognitive test. Nonlinearity was evaluated using hierarchical regression with a linear effect entered during the first step of the models and a quadratic effect entered during the second step. The quadratic trends were significant for each of the cognitive control variables, with $p = .002$ for TMT-B, $p = 0.001$ for perseverative errors, $p = .007$ for the COWAT, and $p < .000$ for the Stroop test. These data reveal that RT differences between individuals with high vs. low performance on the cognitive control tasks are not the results of a linear tranformation and consequently, that general slowing does not account for the findings of this study.

Second, individuals with lower levels of cognitive control might be more cautious during responding due to difficulty with the task, making their responses slower and increasing the possibility that they would have a longer response. In this case, Tau would be positively related to accuracy because a more conservative response threshold is expected to increase performance (Ratcliff, 1978). Tau, however, was *negatively* related to task accuracy ($r = -.567$, $p < .000$), revealing that differences in response threshold do not explain the observed findings.

Speed accuracy tradeoffs are a related possibility that is important to consider. It may be the case that some individuals simply rushed through the WCST, which might adversely affect their performance and render significant results more reflective of task engagement than underlying levels of neurocognitive skill. This possibility was guarded against before any analyses took place, by excluding individuals with a higher number of responses to the “other” category (which suggests failure to adequately understand the

task or random responding). Furthermore, the analyses of Mu reveal that individuals who were faster on the WCST actually performed better on the cognitive control tasks, which rules out the possibility of a speed/accuracy trade off.

Third, variability was compared to performance indices on the task from which it was derived (i.e., the Wisconsin Card Sorting Test). As a result, there is a certain level of content similarity between Tau, WCST perseverative errors, and WCST categories. Some task-specific non-executive source of variance shared by these measures might be responsible for the observed relationships. That is, some individuals might simply have had difficulty with the Wisconsin Card Sorting Test for whatever reason. This could result in slower, more variable responding and lower levels of performance during this task that might explain the findings above. Although content similarity cannot be definitively ruled out, several findings suggest that it cannot independently account for the results reported above.

The fact that variability was meaningfully related to all of the cognitive control variables during tests of bivariate correlations (reported in Table 3) is inconsistent with the effects of content similarity. Furthermore, intraindividual variability on the WCST was also related to more classical variability measures (such as CPT hit RT SE) and these measures were, in turn, related to the cognitive control variables. There is thus a convergence of evidence across multiple measures and multiple measurement modalities suggesting that the relationship between intraindividual variability and cognitive control is a meaningful finding that must be explained.

Theoretical and Neuroanatomical Significance of Findings

The purpose for examining the relationship between Tau, basic attention, cognitive control, and abstract reasoning was to permit statements about how variability may relate to the rostro-caudal organization of the prefrontal cortex. The extent to which the higher-order relationships investigated in hypothesis one and four (i.e., cognitive control vs. reasoning) are differentiable, however, remains open to debate. Abstract reasoning was indexed by only one variable (categories) that in turn was highly correlated with a cognitive control task (perseverative errors), making the determination of differential contributions impossible under this study's analytic constraints. Thus, the relationship between Tau and abstract reasoning may simply reflect individual differences in cognitive control. As a result, the interpretation below primarily focuses on the relationship between Tau and the cognitive control variables.

Within a more general framework, comparing the simple (primarily first-order) executive functions indexed by the CPT variables to more complex functions (indexed by cognitive control and reasoning tasks) is informative and guides the interpretation of significant results below. This general distinction may be more reflective of the current state of our current knowledge about the hierarchical organization of the PFC. The preponderance of rostro-caudal theories agree that progressively anterior regions of the frontal cortex support increasingly complex abstract representations and processes, regardless of their specifics (Badre, 2008).

What might account for the findings of this investigation? At the level of task demands, both perseverative errors (Heaton et al., 1993) and TMT-B (Arbuthnott & Frank, 2000) may be construed as measures of attentional set shifting or cognitive

flexibility. This executive cognitive skill involves the flexible deployment of attentional resources, allowing individuals to switch between competing sets of cognitive operations. Within clinical neuropsychology, conceptualizations of set shifting have focused on the disengagement of an irrelevant task set and subsequent active engagement of a relevant task set (Miyake et al., 2000). Performing a new cognitive operation may also require one to overcome negative priming (i.e., proactive interference) from having previously executed a particular cognitive operation on the same stimulus set (Waszak, Hommel, & Allport, 2003). From this perspective, significant relations between intraindividual variability and the shifting tasks may indicate that variability reflects the control of attention, active engagement of response sets, and the ability to suppress irrelevant information in the pursuit of a self-selected goal.

Differences between the two cognitive control tasks are important to consider in further elucidating these results. TMT-B and perseverative errors are distinct in that the former requires shifting between visual stimuli while the latter is more prototypically executive (i.e., it requires intentional switching between internally generated sorting rules). Functionally different areas of the cortex are thought to subserve the shifting of visual attention vs. more executively oriented shifts, such as the conscious direction of attention on the basis of task instructions (Posner & Raichle, 1997). There is evidence from neuroimaging, however, that significant commonality exists between different types of shifting tasks regardless of whether they involve external visual material or more traditionally executive material. In addition to parietal and occipital regions, various anterior frontal regions appear to be activated during a variety of different shifting paradigms (Wager, Jonides, & Reading, 2004). This neural activation is in turn related to

regions subserving executive cognitive skills that have been associated with intraindividual variability. For example, there is "a striking degree of overlap between attention shifting-related regions and regions shown to be selectively responsive to executive processes in working memory" (Wager et al., 2004, p. 1690). The fact that Tau was significantly related to set shifting tests in this study is thus highly consistent with previous findings that it predicts working memory task performance with impressive effect sizes (Schmiedek et al., 2007; Tse et al., 2010).

Further evidence comes from a study of cortical lesion mapping, which provides causal evidence of the brain regions that underly various components of cognitive control. In this study, Gläscher and colleagues (2012) found that performance on both TMT-B and perseverative errors (purified of non-executive task demands) was related to directly overlapping regions of the left Anterior Cingulate Cortex (ACC). As these authors explain, this area is uniquely situated between frontal networks thought to representationally subserve valuation and cognitive control. The valuation and control networks have their neuroanatomical basis in areas of the dorsolateral and ventral PFC (Gläscher et al., 2012), which in turn, are related to second and first order working memory processes (respectively; Owen, 2000).

Taken together, these findings suggest that intraindividual variability on the Wisconsin Card Sorting Test may reflect fluctuations in the control of various working memory subprocesses involved in set-shifting, and that variability may have a neuroanatomical basis in portions of the ACC and frontal attention networks involving the VMPFC and DLPFC. Previous neuroimaging data found intraindividual variability in response time to be associated with activations of Brodmann's area 9 and 46 (Bellgrove

et al., 2004), both regions of the dorsolateral prefrontal cortex associated with working memory, cognitive control, and the control of attention (Banich, 2009).

There has been a surge of interest in the function of the ACC in recent years, with discussions centering around three key features of this structure as it effects behavioral control: the role of the ACC in motor control, the role of the ACC in cognition, and the importance of arousal/drive state for ACC engagement (Paus, 2001). Neuropsychologically, the rostral ACC has been related to uninstructed set-shifting and error detection (Braver, Barch, Grey, Molfese, & Snyder, 2001; Lie, Specht, Marshall, & Fink, 2006). The specific areas associated with TMT-B and WCST performance in Gläscher and colleagues' (2012) lesion mapping study appear to lie within the proisocortical region (Brodmann's area 24, 25) and the paracingulate gyrus (area 32). They encompass both subcallosal and supracallosal areas of the ACC. Interestingly, these areas are not traditionally associated with motor functions. The density of cortico-cortical connections of the lateral PFC with supracallosal areas 24 and 32 suggests that this particular region of the ACC may be involved in the functions of this structure that relate to cognition (Paus, 2001). In his influential and widely cited article, Paus (2001) attributes the primary function of the ACC as the willed control of behavior. Thus, intraindividual variability may reflect less efficient control mechanisms involved in the coordination of affective, cognitive, and motor information in the pursuit of intentionally selected goals.

The ACC is thought to integrate information from cognitive, affective, and motor channels with the help of ascending striatal dopaminergic neurons (Rushworth & Behrens, 2008). Alterations in dopaminergic neuromotransmission have been

documented in populations that also demonstrate increased RT variability, including patients with schizophrenia, children with ADHD, the elderly, and patient's with Parkinson's disease (MacDonald et al., 2009). Lower D2 receptor binding in the ACC appears to be systematically linked to intraindividual variability on both memory and executive functioning tasks as assessed by a PET-derived *in vivo* marker of D2 binding potential (MacDonald, Cervenka, Farde, Nyberg, & Bäckman, 2009). Additionally, neurocomputational models suggest that "DA dysregulation alters the signal-to-noise ratio of neural information processing, impairing neurons' sensitivity to afferent signals, leading to less distinct cortical representations, and ultimately resulting in increasing intraindividual variability and impaired cognitive functioning" (Macdonald et al., 2009, p. 801). These data are consistent with the results of this investigation, in that they suggest that performance consistency is dependent upon neuromodulatory and neurostructural substrates that are also related to cognitive control (particularly those abilities subserved by the rostral ACC).

An additional perspective is obtained by interpreting these results in the context of task difficulty. Within this framework, greater intraindividual variability in response time reflects the increased difficulty encountered by participants with low levels of cognitive control as they attempted the Wisconsin Card Sorting Task. It is already well established that more executively challenging tasks effect greater variability in normal populations (West et al., 2002). Task difficulty is also the most important factor for effecting increasing ACC activity in functional neuroimaging studies (Paus, 2001). Might lower levels of cognitive control increase the relative task difficulty of the WCST in the same

way, thus leading to greater variability? If this were the case, variability might represent an additional dimension of performance on the WCST that could be utilized clinically.

In an ingenious experiment examining the architecture of cognitive control in the human prefrontal cortex, Koechlin, Ody, and Kouneiher (2003) discovered that mean reaction time increases significantly and in a cumulative fashion with the addition of increased executive demands. To grossly simplify their results, increasing response time was accompanied by greater activation of lateral prefrontal cortical regions (including Brodmann's area 46). RT latency and activations of area 46 were greatest for the most episodically (i.e., executively) demanding block of the experiment and were not explained by increases in motor demands or working memory load.

Koechlin and colleagues (2003) did not explicitly examine the response time data produced by their experimental design, but the linear relationship between the mean and standard deviation of RT distributions (Wagenmakers & Brown, 2007) suggests that if they had, they might have also discovered increased RT variability. Accordingly, lower levels of executive reserve or the presence of impairment might effect greater response time latency and variability in clinical populations in the same way that increased executive task demands effect variability in normal populations. This possibility is intriguing because the greater resolution of response time measures (as compared to traditional assessment techniques) suggests that they might be more sensitive performance indices. Variability might thus represent an additional dimension of measurement, one that is easily extracted from commonly used computerized neuropsychological tasks but is (at present) largely neglected in clinical practice.

Limitations of This Investigation

The conclusions that can be drawn from this study are subject to several important limitations. These weaknesses are considered below as they relate to statistical, methodological, and practical limitations of the project.

At the level of statistics, this study (like all studies of ex-Gaussian parameters that do not employ structural equation modeling) is subject to the problem of parameter dependency. Implementation of the fitting routines in QMPE introduces correlated deviations of parameter estimates from their true values. These correlated error terms then serve as unwanted variance that confounds subsequent analyses of the ex-Gaussian parameters. Cousineau et al. (2004) discuss this problem in detail as it pertains to QMPE. The average intercorrelations of the ex-Gaussian parameters were previously reported in Table 3 in the interest of transparency with respect to this statistical issue.

The functional form of the fitting routines utilized by QMPE dictates that parameter dependencies are positive for Mu and Sigma, negative for Mu and Tau, and negative for Sigma and Tau. In this study, the intercorrelations of Mu and Sigma (r = .61) and Mu and Tau (r = -0.51) were particularly high. The fact that Mu and Sigma were the most highly correlated, yet Sigma largely failed to relate meaningfully to the predictor variables that Mu related to, suggests that the effects of dependencies were not catastrophic to the analyses conducted. Furthermore, Mu and Tau were negatively correlated with one another but their correlations with the other predictor variables were universally in the same direction. Thus, correlated error introduced by the parameter estimation process is unlikely to have overly influenced the study results. If anything, it is possible that such dependencies represented systematic "error" that decreased the degree

of association between Mu, Tau, and the predictor variables. If idealized Mu and Tau variables both correlated in the expected direction with the various behavioral indicators and their dependencies were negative, this would serve to decrease their association with the neuropsychological tasks. Nevertheless, the methodology employed here does not permit completely unambiguous statements about the relative contributions of the ex-Gaussian parameters.

A related issue concerns the reliability and variability of parameter estimates. These are also reported in Table 3. Specifically, reduced variability of Sigma relative to Mu and Tau may have disadvantaged this parameter in the correlational and regression analyses. The substantial positive skew of Sigma (which could not be eliminated via the use of data transformations) also translated to numerous violations of statistical assumptions. These violations could have increased or decreased Sigma's relationship with the various cognitive tasks. Such circumstances may be the norm in most published journal articles that do not provide sufficient data to evaluate the extent to which their analyses violated underlying assumptions; as such, they may not be overly concerning. With these caveats, the findings of this study with respect to Sigma are consistent with previous investigations, as this parameter has only been infrequently related to meaningful constructs.

Decreased reliability of Tau vs. Mu may be a more important issue, given that these parameters demonstrated similar magnitudes of relationship to theoretically meaningful variables in certain scenarios (e.g., tests of hypothesis 4). The standard error of Tau was nearly twice that of Mu, which is expected to attenuate its correlations with the various predictor variables. This may be the reason why the overall variance in Tau

associated with the cognitive control tasks was lower than in previous investigations using SEM, which is capable of measuring potentially more reliable latent Tau constructs that are devoid of measurement error.

This study made extensive use of atheoretical stepwise variable entry to form the regression equations used in hypothesis tests. This methodological decision was made in the setting of insufficient prior research to determine the expected order of variables *a priori*. Bivariate correlations between all the variables analyzed in this study are reported in order to allow the reader to come to their own conclusion about their relative importance, regardless of whether they appeared in the final regression models. The decision to use stepwise entry had its largest impact in Tests of hypothesis 1, as the Stroop test and COWAT were not included in final models despite their significant bivariate correlations with Mu and Tau. Unreported analyses testing various combinations of predictor variables through forced entry (conducted after tests of primary study hypotheses) suggest that the results of hypothesis 1 would have generally been the same even if the variables were entered in a different order.

Significant overlap between the predictor variables suggests that the use of SEM to form latent first-order and second-order executive factors might have been preferable to the regression-based approaches employed herein. This possibility was considered and abandoned due to sample size limitations, practical limitations, and the more stringent data requirements for this methodology. Certain theories modeling the rostro-caudal organization of the PFC have been tested and validated using SEM (e.g., hierarchy based on relational complexity, Christoff & Gabrielli, 2000; the cascade-control model;

Koechlin et al., 2003) and it would be intriguing to see how variability relates to a measurement model of PFC organization.

Methodologically speaking, this study employed a series of highly complex neuropsychological tasks that are commonly utilized in clinical practice. These tasks all contain multiple demands and do not permit the type of fine-grained dissociation of component processes necessary for definitive statements about the relationship of intraindividual variability to precisely specified executive task demands. Such an analysis would be useful to inform theory. The purpose of this project, however, is to establish the clinical importance of variability in the context of theory, rather than to primarily investigate or advance a theory. For this reason, links to neuroanatomical substrates and theories of frontal lobe functioning are offered tentatively rather than definitively. Future efforts should use experimental designs capable of disentangling the relative contributions of different executive sub-processes to response time consistency.

Perhaps most importantly, nearly half the variance in Tau was not successfully modeled by any combination of variables. The implication being that this measure of intraindividual variability may reflect many systematic sources of underlying variability that are unrelated to executive cognitive skills. The effect sizes for the overall regression equations predicting Tau were generally large, however. Previous investigation have reported higher correlations between Tau and latent working memory constructs (i.e., r = -.72, Schmiedek et al, 2010; r = -.90, Tse et al. 2010). Using statistical methodology that permitted more precise quantification of Tau (absent parameter dependencies and error variance) could have possibly reproduced similar results.

Recommendations for Future Research

This study adds to a growing body of evidence that suggests variability research may be directly applicable to clinical assessment. Previous results have demonstrated that RT variability data can predict and precede impairment before it is evident in mean levels of performance (Lövden et al., 2007), contribute incrementally to diagnostic decisions (Holtzer et al., 2008; Hultsch et al., 2000), and achieve greater sensitivity than more commonly employed measures of mean RT (Leth-Steensen et al., 2000). To establish intraindividual variability measures as useful neuropsychological tools, future research might examine which measures of RT variability would be most effectively developed for clinical use. It is likely that certain tasks are best suited to particular populations or assessment questions, the specifics of which could be clarified by studies simultaneously testing a variety of methods in broad clinical samples. Select publications provide preliminary evidence that the diagnostic utility of variability measures is partially dependent upon the task paradigm from which they are derived (de Frias et al., 2007).

It will also be important to demonstrate the extent to which intraindividual variability contributes meaningfully to the diagnostic process. Previous work indicates that variability data may increase the accuracy of classification in the setting of dementia (Dixon et al., 2007; Holtzer et al., 2008) and ADHD (Leth-Steensen et al., 2000). Establishing the sensitivity and specificity of intraindividual variability measures for the detection of different forms of neuropathology is an important area for future research.

The interaction effect uncovered during tests of hypothesis 1 raises the possibility that variability might differentially relate to underlying neurocognitive skills as a function of impairment level. Clarifying this possibility by evaluating variability at different levels

of cognitive performance might elucidate those contexts in which it would be most effecitvely utilized. For example, variability measures might prove to be a sensitive indicator of impending impairment that would be ideal in a screening context or for the detection of mild cognitive impairment (e.g., Lövden et al., 2007).

Theoretically, this study is the first to explicitly relate its findings to theories of rostro-caudal prefrontal cortical organization. The results suggest that intraindividual variability measures may be differentially related to functionally and neuroanatomically dissociable stages of executive functioning. Researchers might explicitly test this hypothesis through the use of carefully designed experimental paradigms capable of measuring various executive subprocesses in conjunction with RT variability. Such investigations could provide valuable data about how variability relates to the rostro-caudal organization of the prefrontal cortex. Combined with latent variability constructs (measured by structural equation modeling), such experiments might make definitive statements about the relationship between performance consistency and various subprocesses of cognitive control and executive functioning.

Finally, a great deal of previous intraindividual variability studies have focused on the use of intraindividual standard deviations, the coefficient of variation, or other summary statistics. These less sophisticated techniques appear to be sub-optimal to the use of ex-Gaussian parameters for measuring RT variability (particularly in the context of structural equation modeling). Nevertheless, many significant previous findings rely primarily on the use of these statistics. Exploring the extent to which ex-Gaussian parameters converge with these more traditional measurements could link previous findings with contemporary efforts that employ distribution fitting.

Implications for the Practice of Clinical Neuropsychology

To effectively utilize an assessment measure in the clinical context, neuropsychologists must have detailed knowledge about its reliability and validity. Previous efforts suggest that intraindividual variability in response time is substantial in magnitude (Nesselroade & Salthouse, 2004) and can be reliably measured (Hertzog, Dixon, & Hultsch, 1992). The results of this study demonstrate that it also enjoys convergent validity with commonly employed measures of executive functioning. Bivariate correlations of the ex-Gaussian parameter Tau with a battery of cognitive control tasks were marginally higher than the intercorrelations of the tasks themselves in a broad clinical sample. By these standards, variability seems to be at least as deserving of space in neuropsychological test batteries as these more traditional assessment methods.

Tests of primary hypotheses revealed that variability is more reflective of cognitive control abilities (as they are typically defined by practicing neuropsychologists) than simple attentional abilities. Moreover, the relationship between intraindividual variability and cognitive control was only partially mediated by more basic attentional skill. The results also provide preliminary evidence that variability may be related to putative "third stage" fronto-polar executive functions, such as abstract reasoning. Variability thus appears to be a potentially viable measure that might differentially index various partially dissociable facets of executive functioning.

The executive functions represent a complicated set of interrelated cognitive skills that allow individuals to meaningfully pursue external goals on the basis of internally and externally generated information (Badre et al., 2005; Miyake et al., 2000). Three

executive functions have been frequently hypothesized in the neuropsychological literature: set shifting, information updating/monitoring, and inhibition. Intraindividual variability in response time on the WCST appears to be specifically related to tests of set-shifting.

The WCST is the most commonly utilized clinical test of executive functioning (Heaton et al., 1993) and current evidence supports the notion that performance on this task is differentially affected by DLPFC lesions (Stuss & Levine, 2003). TMT-B is another measure widely thought to be affected by DLPFC damage. In this study, perseverative error score from the WCST and time to completion on TMT-B predicted intraindividual variability on the WCST with a large effect size (f^2 = .821). This combination of cognitive tests accounted for 30.3% of the variance in intraindividual variability on the WCST after controlling for age. These data are consistent with the interpretation that RT variability on the WCST reflects the active control of attention, engagement of response sets, and the ability to suppress task irrelevant information in the pursuit of a self-selected goal.

Some previous work by Stuss and colleagues (2003) provides lesion data that intraindividual variability seems to be causally related to DLPFC damage and the results of neuroimaging converge on this interpretation (e.g., Bellgrove et al., 2004). Intraindividual variability was selectively related to measures thought to index DLPFC functioning in this study. Combined with the results of previous investigations, these findings suggest that intraindividual variability may represent the efficiency or stability with which individuals are able to regulate various working memory subprocesses whose

neuroanatomical basis corresponds to frontal executive attention networks including the DLPFC, VLPFC, and ACC in addition to other non-frontal areas.

Further research is necessary before clinicians can confidently incorporate novel RT variability measures into their clinical batteries. Specifically, efforts utilizing full neuropsychological batteries in conjunction with variability measurements might help to clarify the construct validity of intraindividual variability in the clinical setting. Work examining patterns of variability across different tasks in distinct clinical subpopulations might add to our existing knowledge of how variability relates to underlying CNS pathology. Continued use of functional neuroimaging, cortical-lesion mapping, and lesion study methods may set the stage for a new clinical measurement tradition based on the application of intraindividual variability data.

References

Alderton, D. L., & Larson, G. E. (1994). Cross-task consistency in strategy use and the relationship with intelligence. *Intelligence*, *18*(1), 47–76. doi:16/0160-2896(94)90020-5

Allaire, J. C., & Marsiske, M. (2005). Intraindividual variability may not always indicate vulnerability in elders' cognitive performance. *Psychology and Aging*, *20*(3), 390–401. doi:10.1037/0882-7974.20.3.390

Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory* (1st ed.). Long Grove, IL: Waveland Press

Andreasen, N. C., Paradiso, S., & O'Leary, D. S. (1998). "Cognitive dysmetria" as an integrative theory of schizophrenia: A dysfunction in cortical-subcortical-cerebellar circuitry? *Schizophrenia Bulletin*, *24*(2), 203–218. doi:10.1093/oxfordjournals.schbul.a033321

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(2), 514–544. doi:10.1037/0278-7393.27.2.514

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, *22*(4), 518–528. doi:10.1076/1380-3395(200008)22:4;1-0;FT518

Army Individual Test Battery (1944): *Manual of directions and scoring*. Washington, DC: War Department, Adjutant General's Office.

Azuma, T. (2004). Working memory and perseveration in verbal fluency. *Neuropsychology*, *18*(1), 69–77. doi:10.1037/0894-4105.18.1.69

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. *Neuroscience & Biobehavioral Reviews*, *30*(6), 791–807. doi:10.1016/j.neubiorev.2006.06.005

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47 – 89). New York, NY: Academy Press.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*(5), 193–200. doi:10.1016/j.tics.2008.02.004

Baldo, J. V., Schwartz, S., Wilkins, D., & Dronkers, N. F. (2006). Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society*, *12*(06), 896–900. doi:10.1017/S1355617706061078

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. doi:10.1037/0096-3445.133.2.283

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry. *Current Directions in Psychological Science*, *20*(3), 160 –166. doi:10.1177/0963721411408885

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language*, *59*(4), 495–523. doi:10.1016/j.jml.2007.10.004

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*(2), 89–94. doi:10.1111/j.1467-8721.2009.01615.x

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

Bellgrove, M. A., Hester, R., & Garavan, H. (2004). The functional neuroanatomical correlates of response variability: Evidence from a response inhibition task. *Neuropsychologia*, *42*(14), 1910–1916. doi:10.1016/j.neuropsychologia.2004.05.007

Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, *31*(9), 907–922. doi:10.1016/0028-3932(93)90147-R

Berman, K. F., Ostrem, J. L., Randolph, C., Gold, J., Goldberg, T. E., Coppola, R.,…Weinberger, D. R. (1995). Physiological activation of a cortical network during performance of the Wisconsin Card Sorting Test: A positron emission tomography study. *Neuropsychologia*, *33*(8), 1027–1046. doi:10.1016/0028-3932(95)00035-2

Bielak, A. A. M., Hultsch, D. F., Strauss, E., MacDonald, S. W. S., & Hunter, M. A. (2010). Intraindividual variability is related to cognitive change in older adults: Evidence for within-person coupling. *Psychology and Aging*, *25*(3), 575–586. doi:10.1037/a0019503

Birkett, P., Sigmundsson, T., Sharma, T., Toulopoulou, T., Griffiths, T. D., Reveley, A., & Murray, R. (2007). Reaction time and sustained attention in schizophrenia and its genetic predisposition. *Schizophrenia Research*, *95*(1-3), 76–85. doi:10.1016/j.schres.2007.05.030

Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The reliability and internal validity of the Wisconsin Card Sorting Test. *Neuropsychological Rehabilitation, 8*(3), 243–254. doi:10.1080/713755573

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition and errors. *Cerebral Cortex*, *11*(9), 825–836. doi:10.1093/cercor/11.9.825

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, *26*(2), 571–593. doi:10.1207/s15326942dn2602_3

Brodmann, K. (1909). *Vergleichende lokalisationslehre der grosshirnrinde: In ihren prinzipien dargestellt auf grund des zellenbaues* [Localization in the cerebral cortex: The principles of comparative localization in the cerebral cortex based on cytoarchitechtonics]. Leipzig, Germany: Verlag von Johan Ambrosius Barth.

Brown, S., Cousineau, D., & Heathcote, A. (2004). Technical manual for QMPE v2.18: Fortran code to fit response time distributions. Retrieved from http://newcl.org/software/qmpe/QMLE2man_v218.pdf

Bruce, J. M., Bruce, A. S., & Arnett, P. A. (2010). Response variability is associated with self-reported cognitive fatigue in Multiple Sclerosis. *Neuropsychology, 24*(1), 77–83. doi:10.1037/a0015046

Buchsbaum, B. R., Greer, S., Chang, W.-L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting task and component processes. *Human Brain Mapping*, *25*(1), 35–45. doi:10.1002/hbm.20128

Bunce, D., Anstey, K. J., Christensen, H., Dear, K., Wen, W., & Sachdev, P. (2007). White matter hyperintensities and within-person variability in community-dwelling adults aged 60-64 years. *Neuropsychologia*, *45*(9), 2009–2015. doi:10.1016/j.neuropsychologia.2007.02.006

Bunce, D., Handley, R., & Gaines Jr., S. O. (2008). Depression, anxiety, and within-person variability in adults aged 18 to 85 years. *Psychology and Aging*, *23*(4), 848–858. doi:10.1037/a0013678

Burton, C. L., Strauss, E., Hultsch, D. F., & Hunter, M. A. (2009). The relationship between everyday problem solving and inconsistency in reaction time in older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *16*(5), 607–632. doi:10.1080/13825580903167283

Burton, C. L., Strauss, E., Hultsch, D. F., Moll, A., & Hunter, M. A. (2006). Intraindividual variability as a marker of neurological dysfunction: A comparison of Alzheimer's disease and Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, *28*(1), 67–83. doi:10.1080/13803390490918318

Callicott, J. H., Mattay, V. S., Verchinski, B. A., Marenco, S., Egan, M. F., & Weinberger, D. R. (2003). Complexity of prefrontal cortical dysfunction in schizophrenia: More than up or down. *American Journal of Psychiatry*, *160*(12), 2209–2215. doi:10.1176/appi.ajp.160.12.2209

Cattell (1966). The data box: It's ordering of total resources in terms of relational systems. In *Handbook of Multivariate Experimental Psychology*, R. B. Cattel (Ed.) pp. 67 - 128, Chicago, IL: Rand McNally.

Chase, H. W., Clark, L., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2008). Dissociable roles of prefrontal subregions in self-ordered working memory performance. *Neuropsychologia*, *46*(11), 2650–2661. doi:10.1016/j.neuropsychologia.2008.04.021

Christensen, H., Dear, K. B. G., Anstey, K. J., Parslow, R. A., Sachdev, P., & Jorm, A. F. (2005). Within-occasion intraindividual variability and preclinical diagnostic status: Is intraindividual variability an indicator of mild cognitive impairment? *Neuropsychology*, *19*(3), 309–317. doi:10.1037/0894-4105.19.3.309

Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, *28*(2), 168–186. doi:10.3758/BF03331976

Clark, L., Blackwell, A. D., Aron, A. R., Turner, D. C., Dowson, J., Robbins, T. W., & Sahakian, B. J. (2007). Association between response inhibition and working memory in adult ADHD: A link to right frontal cortex pathology? *Biological Psychiatry*, *61*(12), 1395–1401. doi:10.1016/j.biopsych.2006.07.020

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. London, UK: Routledge.

Collins, L. F., & Long, C. J. (1996). Visual reaction time and its relationship to neuropsychological test performance. *Archives of Clinical Neuropsychology*, *11*(7), 613–623. doi:10.1016/0887-6177(97)81255-3

Conners, C. K., & Staff, M. H. S. (2000). *Conners' Continuous Performance Test II* (CPT II V. 5). North Tonawanda, NY: Multi-health systems

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.

Cousineau, D., Brown, S., & Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers, 36*, 742–756. doi:10.3758/BF03206555

Cousineau, D., & Chartier, S. (2010). Outliers detection and treatment. *International Journal of Psychological Research, 3*(1), 58–67.

Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence, 31*(6), 567–587. doi:16/S0160-2896(03)00054-0

Cummings, J. L., & Miller, B. L. (2007). Conceptual and clinical aspects of the frontal lobes. In B. L. Miller & J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (2nd ed., pp. 12-21). New York, NY: The Guilford Press.

D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. *Experimental Brain Research, 133*(1), 3–11. doi:10.1007/s002210000395

De Frias, C. M., Dixon, R. A., Fisher, N., & Camicioli, R. (2007). Intraindividual variability in neurocognitive speed: A comparison of Parkinson's disease and normal older adults. *Neuropsychologia*, *45*(11), 2499–2507. doi:10.1016/j.neuropsychologia.2007.03.022

Deary, I. J., Der, G., & Ford, G. (2000). Reaction times and intelligence differences: A population-based cohort study. *Intelligence*, *29*(5), 389–399. doi:10.1016/S0160-2896(01)00062-9

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS).* San Antonio, TX: The Psychological Corporation.

Delis, S. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *The California Verbal Learning Test - Second Edition (Adult version).* San Antonio, TX: The Psychological Corporation.

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society: JINS*, *5*(4), 346–356. doi:10.1017/S1355617799544056

Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W. S., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. *Neuropsychology*, *21*(3), 381–399. doi:10.1037/0894-4105.21.3.381

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E.,…Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(12), 6917–6922. doi:10.1073/pnas.111134598

Eisenberg, J., Mei- Tal, G., Steinberg, A., Tartakovsky, E., Zohar, A., Gritsenko, I., … Ebstein, R. P. (1999). Haplotype relative risk study of catechol- O- methyltransferase (COMT) and attention deficit hyperactivity disorder (ADHD): Association of the high-enzyme activity val allele with adhd impulsive- hyperactive phenotype. *American Journal of Medical Genetics*, *88*(5), 497–502. doi:10.1002/(SICI)1096-8628(19991015)88:5<497::AID-AJMG12>3.0.CO;2-F

Ettenhofer, M. L., Foley, J., Behdin, N., Levine, A. J., Castellon, S. A., & Hinkin, C. H. (2010). Reaction time variability in HIV-positive individuals. *Archives of Clinical Neuropsychology*, *25*(8), 791–798. doi:10.1093/arclin/acq064

Everett, J., Lavoie, K., Gagnon, J. F., & Gosselin, N. (2001). Performance of patients with schizophrenia on the Wisconsin Card Sorting Test (WCST). *Journal of Psychiatry and Neuroscience*, *26*(2), 123–130.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi:10.3758/BRM.41.4.1149

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London, England: SAGE.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381–391. doi:10.1037/h0055392

Ford, D. H. (1994). *Humans as self-constructing living systems* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Franzen, M. D., Tishelman, A. C., Sharp, B. H., & Friedman, A. G. (1987). An investigation of the test-retest reliability of the Stroop Color and Word Test across two intervals. *Archives of Clinical Neuropsychology*, *2*(3), 265–272. doi:10.1016/0887-6177(87)90014X

Friston, K. J., & Frith, C. D. (1995). Schizophrenia: A disconnection syndrome? *Clinical Neuroscience*, *3*(2), 89–97.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2008). *Cognitive neuroscience: The biology of the mind* (3rd Ed.). New York, NY: W. W. Norton & Company.

Geurts, H. M., Grasman, R. P. P. P., Verté, S., Oosterlaan, J., Roeyers, H., van Kammen, S. M., & Sergeant, J. A. (2008). Intra-individual variability in ADHD, autism spectrum

disorders and Tourette's syndrome. *Neuropsychologia*, *46*(13), 3030–3041. doi:16/j.neuropsychologia.2008.06.013

Gläscher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M.,Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences*, *109*(36), 14681–14686. doi:10.1073/pnas.1206608109

Goldberg, T. E., Egan, M. F., Gscheidle, T., Coppola, R., Weickert, T., Kolachana, B. S., Weinberger, D. R. (2003). Executive subprocesses in working memory: Relationship to catechol-O-methyltransferase Val158Met genotype and schizophrenia. *Archives of General Psychiatry*, *60*(9), 889–896. doi:10.1001/archpsyc.60.9.889

Golden, C. (1975). A group form of the Stroop Color and Word Test. *Journal of Personality Assessment*, *39*, 386–388.

Golden, C., & Freshwater, S. (2002). *A manual for the adult Stroop Color and Word Test*. Chicago, IL: Stoelting.

Goldman- rakic, P. S. (1995). Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences*, *769*(1), 71–84. doi:10.1111/j.17496632.1995.tb38132.x

Greve, K. W., Stickle, 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. doi:10.1016/j.acn.2004.09.004

Hale, S., Myerson, J., Smith, G. A., & Poon, L. W. (1988). Age, variability, and speed: Between-subjects diversity. *Psychology and Aging*, *3*(4), 407–410. doi:10.1037/0882-7974.3.4.407

Hawking, S. (1998). *A brief history of time* (10th anniversary Ed.). New York, NY: Bantam.

Heathcote, A. (n.d.). RTSYS 1.0: A DOS application for the analysis of reaction time data [reference manual]. Retrieved from http://newcl.org/software/rtsys-manual.pdf

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating lognormal, wald, and weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, *36*, 277–290. doi:10.3758/BF03195574

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, *9*, 394–401. doi:10.3758/BF03196299

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the stroop task. *Psychological Bulletin, 109*(2), 340–347. doi:10.1037/0033-2909.109.2.340

Heaton, R. K. (1981). *Wisconsin Card Sorting Test manual*. Odessa, FL: Psychological Assessment Resources.

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1992). Intraindividual change in text recall of the elderly. *Brain and Language, 42*(3), 248–269. doi:10.1016/0093-934X(92)90100-S

Hervey, A. S., Epstein, J. N., Curry, J. F., Tonev, S., Eugene Arnold, L., Conners, K. C., Hechtman, L. (2006). Reaction time distribution analysis of neuropsychological performance in an ADHD sample. *Child Neuropsychology*, *12*, 125–140. doi:10.1080/09297040500499081

Holtzer, R., Verghese, J., Wang, C., Hall, C. B., & Lipton, R. B. (2008). Within-person across-neuropsychological test variability and incident dementia. *JAMA: The Journal of the American Medical Association, 300*(7), 823 –830. doi:10.1001/jama.300.7.823

Howell, D. (2012). *Statistical methods for psychology*. Boston, MA: Cengage Learning.

Hultsch, D. F., & MacDonald, S. (2004). Intraindividual variability in performance as a theoretical window onto cognitive aging. In R. A. Dixon, L. Bäckman, & L.-G. Nilsoon (Eds.), *New frontiers in cognitive aging* (pp. 65–88). New York, NY: Oxford University Press.

Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *57*(2), P101–115. doi:10.1093/geronb/57.2.P101

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology*, *14*(4), 588–598. doi:10.1037//0894-4105.14.4.588

Hultsch, D. F., Strauss, E., Hunter, M. A., & MacDonald, S. W. S. (2008). Intraindividual variability, cognition, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 491–556). New York, NY: Psychology Press.

Iverson, G. L., Lange, R. T., Green, P., & Franzen, M. D. (2002). Detecting exaggeration and malingering with the Trail Making Test. *The Clinical Neuropsychologist*, *16*(3), 398–406. doi:10.1076/clin.16.3.398.13861

Jackson, J. D., Balota, D. A., Duchek, J. M., & Head, D. (2012). White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer disease. *Neuropsychologia*, *50*(3), 357–366. doi:10.1016/j.neuropsychologia.2011.11.024

Kaiser, S., Roth, A., Rentrop, M., Friederich, H.-C., Bender, S., & Weisbrod, M. (2008). Intra-individual reaction time variability in schizophrenia, depression and borderline personality disorder. *Brain and Cognition, 66*(1), 73–82. doi:10.1016/j.bandc.2007.05.007

Kelly, A. M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2008). Competition between functional brain networks mediates behavioral variability. *NeuroImage, 39*(1), 527–537. doi:10.1016/j.neuroimage.2007.08.008

Kelly, T. P. (2000). The clinical neuropsychology of attention in school-aged children. *Child neuropsychology: A journal on normal and abnormal development in childhood and adolescence, 6*(1), 24–36. doi:10.1076/0929-7049(200003)6:1;1-B;FT024

Kieffaber, P. D., Kappenman, E. S., Bodkins, M., Shekhar, A., O'Donnell, B. F., & Hetrick, W. P. (2006). Switch and maintenance of task set in schizophrenia. *Schizophrenia Research, 84*(2-3), 345–358. doi:10.1016/j.schres.2006.01.022

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302*(5648), 1181–1185. doi:10.1126/science.1088545

Kraepelin, E. (1919). *Dementia praecox and paraphrenia*. Chicago, IL: Chicago Medical Book

Kremen, W. S., Seidman, L. J., Faraone, S. V., Pepple, J. R., & Tsuang, M. T. (1992). Attention/information-processing factors in psychotic disorders: Replication and extension of recent neuropsychological findings. *The Journal of Nervous and Mental Disease, 180*(2), 89–93. doi:10.1097/00005053-199202000-00004

Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica, 104*(2), 167–190. doi:16/S0001-6918(00)00019-6

Lezak, M. D. (2004). *Neuropsychological assessment*. New York, NY: Oxford University Press.

Li, S.-C., Aggen, S. H., Nesselroade, J. R., & Baltes, P. B. (2001). Short-term fluctuations in elderly people's sensorimotor functioning predict text and spatial memory performance: The Macarthur successful aging studies. *Gerontology, 47*(2), 100–116. doi:10.1159/000052782

Li, S.-C., Huxhold, O., & Schmiedek, F. (2004). Aging and attenuated processing robustness. *Gerontology, 50*(1), 28–34. doi:10.1159/000074386

Li, S.-C., & Lindenberger, U. (1999). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilities in old age. In L.-G. Nilsson & H. J. Markowitsch (Eds.), *Cognitive neuroscience of memory*. Seattle, WA: Hogrefe & Huber.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15*(3), 155 –163. doi:10.1111/j.0956-7976.2004.01503003.x

Li, S. C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences, 5*(11), 479–486. doi:10.1016/S1364-6613(00)01769-1

Lie, S.-C., Specht, K., Marshall, J. C., & Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *NeuroImage, 30*(3), 1038–1049. doi:10.1016/j.neuroimage.2005.10.031

Lindenberger, U., & von Oertzen, T. (2006). Variability in cognitive aging: From taxonomy to theory. In F. I. M. Craik & E. Bialystok (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 297–314) New York, NY: Oxford University Press.

Lövdén, M., Li, S.-C., Shing, Y. L., & Lindenberger, U. (2007). Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: Longitudinal data from the Berlin Aging Study. *Neuropsychologia*, *45*(12), 2827–2838. doi:16/j.neuropsychologia.2007.05.005

Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

MacDonald, S. W. S., Cervenka, S., Farde, L., Nyberg, L., & Bäckman, L. (2009). Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. *Neuropsychologia*, *47*(11), 2299–2304. doi:10.1016/j.neuropsychologia.2009.01.016

MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2003a). Performance variability is related to change in cognition: Evidence From the Victoria Longitudinal Study. *Psychology and Aging*, *18*(3), 510–523. doi:10.1037/0882-7974.18.3.510

MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2008). Predicting impending death: Inconsistency in speed is a selective and early marker. *Psychology and Aging*, *23*(3), 595–607. doi:10.1037/0882-7974.23.3.595

MacDonald, S. W. S., Li, S.-C., & Bäckman, L. (2009). Neural underpinnings of within-person variability in cognitive functioning. *Psychology and Aging*, *24*(4), 792–808. doi:37/a0017798

MacDonald, S. W. S., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behavior: Links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, *29*(8), 474–480. doi:16/j.tins.2006.06.011

MacDonald, S. W. S., Nyberg, L., Sandblom, J., Fischer, H., & Bäckman, L. (2008). Increased response-time variability is associated with reduced inferior parietal activation during episodic recognition in aging. *Journal of Cognitive Neuroscience*, *20*(5), 779–786. doi:10.1162/jocn.2008.20502

Marenco, S., Coppola, R., Daniel, D. G., Zigun, J. R., & Weinberger, D. R. (1993). Regional cerebral blood flow during the Wisconsin Card Sorting Test in normal subjects studied by xenon-133 dynamic SPECT: Comparison of absolute values, percent distribution values, and covariance analysis. *Psychiatry Research*, *50*(3), 177–192. doi:10.1016/0925-4927(93)90029-H

Mattay, V. S., Goldberg, T. E., Fera, F., Hariri, A. R., Tessitore, A., Egan, M. F., … Weinberger, D. R. (2003). Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. *Proceedings of the National Academy of Sciences, 100*(10), 6186 –6191. doi:10.1073/pnas.0931309100

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, *16*, 798–817. doi:10.3758/PBR.16.5.798

Meyer-Lindenberg, A., Nichols, T., Callicott, J. H., Ding, J., Kolachana, B., Buckholtz, J.,…Weinberger, D. R. (2006). Impact of complex genetic variation in COMT on human brain function. *Molecular Psychiatry, 11*(9), 867–877. doi:10.1038/sj.mp.4001860

Meyers, L. S., Gamst, G. C., & Guarino, A. J. (2005). *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: SAGE Publications, Inc.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167–202.

Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(4), 907–912. doi:10.1080/14640749108400962

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, *9*(1), 90.

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. doi:10.1007/BF01109051

Mitrushina, M. N. (2005). *Handbook of normative data for neuropsychological assessment*. New York, NY: Oxford University Press.

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. doi:10.1006/cogp.1999.0734

Murtha, S., Cismaru, R., Waechter, R., & Chertkow, H. (2002). Increased variability accompanies frontal lobe damage in dementia. *Journal of the International Neuropsychological Society: JINS*, *8*(3), 360–372. doi:10.1017.S1355617701020173

Myerson, J., Adams, D. R., Hale, S., & Jenkins, L. (2003). Analysis of group differences in processing speed: Brinley plots, Q-Q plots, and other conspiracies. *Psychonomic Bulletin & Review*, *10*(1), 224–237.

Myerson, J., Hale, S., Wagstaff, D., Poon, L. W., & Smith, G. A. (1990). The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review*, *97*(4), 475. doi:10.1037/0033-295X.97.4.475

Nagahama, Y., Fukuyama, H., Yamauchi, H., Matsuzaki, S., Konishi, J., Shibasaki, H., & Kimura, J. (1996). Cerebral activation during performance of a card sorting test. *Brain*, *119*(5), 1667–1675. doi:10.1093/brain/119.5.1667

Nesselroade (1991). The warp and the woof of the developmental fabric. In R. Downs, L. Liben, & D. S. Palermo (Eds.), *Visions of Aesthetics, the environment, and development: The legacy of Joachin F. Wohlwill* (pp. 213 - 240). Hillsdale, NJ: Erlbaum.

Nesselroade, J. R. (2002). Elaborating the Differential in Differential Psychology. *Multivariate Behavioral Research*, *37*(4), 543–561. doi:10.1207/S15327906MBR3704_06

Nesselroade, J. R., & Ram, N. (2004). Studying intraindividual variability: What we have learned that will help us understand lives in context. *Research in Human Development*. doi:10.1080/15427609.2004.9683328

Nesselroade, J. R., & Salthouse, T. A. (2004). Methodological and theoretical implications of intraindividual variability in perceptual-motor performance. *The Journals of*

Gerontology. Series B, Psychological Sciences and Social Sciences, *59*(2), P49–55. doi:10.1093/geronb/59.2.P49

Nimon, K. (2010). Regression commonality analysis: Demonstration of an SPSS solution. *Multiple Linear Regression Viewpoints*, *36*(1), 10–17.

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*, *71*(3), 437–451. doi:10.1016/j.bandc.2009.03.005

Ogg, R. J., Zou, P., Allen, D. N., Hutchins, S. B., Dutkiewicz, R. M., & Mulhern, R. K. (2008). Neural correlates of a clinical continuous performance test. *Magnetic Resonance Imaging*, *26*(4), 504–512. doi:10.1016/j.mri.2007.09.004

Orzack, M. H., & Kornetsky, C. (1966). Attention dysfunction in chronic schizophrenia. *Archives of General Psychiatry*, *14*(3), 323–326. doi:10.1001/archpsyc.1966.01730090099015

Owen, A. M. (1997). The functional organization of working memory processes within human lateral frontal cortex: The contribution of functional neuroimaging. *European Journal of Neuroscience*, *9*(7), 1329–1339. doi:10.1111/j.1460-9568.1997.tb01487.x

Owen, A. M. (2000). The role of the lateral frontal cortex in mnemonic processing: The contribution of functional neuroimaging. *Experimental Brain Research*, *133*(1), 33–43. doi:10.1007/s002210000398

Pandya, D. N., & Yeterian, E. H. (1996). Morphological correlations of human and monkey frontal lobe. In A. R. Damasio, H. Damasio, & Y. Christen (Eds.), *Neurobiology of Decision-Making* (pp. 13–46). Berlin: Springer.

Panza, F., D'Introno, A., Colacicco, A. M., Capurso, C., Del Parigi, A., Caselli, R. J.,…Solfrizzi, V. (2005). Current epidemiology of mild cognitive impairment and other predementia syndromes. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *13*(8), 633–644. doi:10.1176/appi.ajgp.13.8.633

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, *2*(6). doi:10.1038/35077500

Petrides, M. (1995). Functional organization of the human frontal cortex for mnemonic processing: Evidence from neuroimaging studies. *Annals of the New York Academy of Sciences*, *769*, 85–96. doi:10.1111/j.1749-6632.1995.tb38133.x

Posner, M. J., & Raichle, M. E. (1997). *Images of mind* (1st Ed.). New York, NY: W. H. Freeman.

Prado, J., Carp, J., & Weissman, D. H. (2011). Variations of response time in a selective attention task are linked to variations of functional connectivity in the attentional network. *NeuroImage*, *54*(1), 541–549. doi:10.1016/j.neuroimage.2010.08.022

Rabbitt, P., Osman, P., Moore, B., & Stollery, B. (2001). There are stable individual differences in performance variability, both from moment to moment and from day to day. *The Quarterly Journal Of Experimental Psychology. A, Human Experimental Psychology*, *54*(4), 981–1003. doi:10.1080/02724980042000534

Ragland, J. D., Glahn, D. C., Gur, R. C., Censits, D. M., Smith, R. J., Mozley, P. D., & Gur, R. E. (1997). PET regional cerebral blood flow change during working and declarative demory: Relationship with task performance. *Neuropsychology*, *11*(2), 222–231.

Ram, N., & Gerstorf, D. (2009). Time-structured and net intraindividual variability: Tools for examining the development of dynamic characteristics and processes. *Psychology and Aging*, *24*(4), 778–791. doi:10.1037/a0017915

Ram, N., Rabbitt, P., Stollery, B., & Nesselroade, J. R. (2005). Cognitive performance inconsistency: Intraindividual change and variability. *Psychology and Aging*, *20*(4), 623–633. doi:37/0882-7974.20.4.623

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. doi:10.1037/0033-2909.114.3.510

Ratcliff, R., Mckoon, D. G., & Hockley, B. (1978). A theory of memory retrieval. *Psychological Review*, 59–108. doi:10.1037//0033-295X.85.2.59

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*(3), 190–214. doi:10.1037/0033-295X.83.3.190

Ravnkilde, B., Videbech, P., Rosenberg, R., Gjedde, A., & Gade, A. (2002). Putative tests of frontal lobe function: A PET-study of brain activation during Stroop's Test and verbal fluency. *Journal of Clinical and Experimental Neuropsychology*, *24*(4), 534–547. doi:10.1076/jcen.24.4.534.1033

Rentrop, M., Rodewald, K., Roth, A., Simon, J., Walther, S., Fiedler, P., & Kaiser, S. (2010). Intra-individual variability in high-functioning patients with schizophrenia. *Psychiatry Research*, *178*(1), 27–32. doi:10.1016/j.psychres.2010.04.009

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. doi:10.1016/S0887-6177(01)00111-1

Ridderinkhof, R. K., van den Wildenberg, P. M. W., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140. doi:10.1016/j.bandc.2004.09.016

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome Jr., E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, *20*(5), 343–350. doi:10.1037/h0043220

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397. doi:10.1038/nn2066

Russell, E. W. (1986). The Psychometric Foundation of Clinical Neuropsychology. In Filskov & T. J. Boll, T. J. (Eds.), *Handbook of Clinical Neuropsychology* (Vols. 1-2). New York: Wiley.

Sakagami, M., Tsutsui, K., Lauwereyns, J., Koizumi, M., Kobayashi, S., & Hikosaka, O. (2001). A code for behavioral inhibition on the basis of color, but not motion, in ventrolateral prefrontal cortex of macaque monkey. *The Journal of Neuroscience*, *21*(13), 4801–4808.

Salthouse, T. A. (1993). Attentional blocks are not responsible for age-related slowing. *Journal of Gerontology*, *48*(6), P263–270. doi:10.1093/geronj/48.6.P263

Salthouse, T. A., & Berish, D. E. (2005). Correlates of within-person (across-occasion) variability in reaction time. *Neuropsychology*, *19*(1), 77–87. doi:10.1037/0894-4105.19.1.77

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2009). On the relation of mean reaction time and intraindividual reaction time variability. *Psychology and Aging*, *24*(4), 841–857. doi:10.1037/a0017799

Schmiedek, F., Oberauer, K., Wilhelm, O., Sü[beta], H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414–429. doi:37/0096-3445.136.3.414

Schretlen, D., Pearlson, G. D., Anthony, J. C., Aylward, E. H., Augustine, A. M., Davis, A., & Barta, P. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society*, *6*(01), 52–61. doi:10.1017/S1355617700611062

Sener, E. C., Ozkan, S. B., Aribal, M. E., Sanac, A. S., & Aslan, B. (1996). Evaluation of congenital Brown's syndrome with magnetic resonance imaging. *Eye (London, England)*, *10 (Pt. 4)*, 492–496. doi:10.1038/eye.1996.108

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(1), 379–423,623–656. doi:10.1002/j.1538-7305.1948.tb00917.x

Shipley, B. A., Der, G., Taylor, M. D., & Deary, I. J. (2006). Cognition and all-cause mortality across the entire adult age range: Health and lifestyle survey. *Psychosomatic Medicine*, *68*(1), 17 –24. doi:10.1097/01.psy.0000195867.66643.0f

Simmonds, D. J., Fotedar, S. G., Suskauer, S. J., Pekar, J. J., Denckla, M. B., & Mostofsky, S. H. (2007). Functional brain correlates of response time variability in children. *Neuropsychologia*, *45*(9), 2147–2157. doi:10.1016/j.neuropsychologia.2007.01.013

Smith, G. L., Large, M. M., Kavanagh, D. J., Karayanidis, F., Barrett, N. A., Michie, P. T., & O'Sullivan, B. T. (1998). Further evidence for a deficit in switching attention in schizophrenia. *Journal of Abnormal Psychology*, *107*(3), 390–398. doi:10.1037/0021-843X.107.3.390

Somsen, R. J. M., van der Molen, M. W., Richard Jennings, J., & van Beek, B. (2000). Wisconsin Card Sorting in adolescents: Analysis of performance, response times and heart rate. *Acta Psychologica*, *104*(2), 227–257. doi:10.1016/S0001-6918(00)00030-5

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology-Human Perception and Performance*, *22*(2), 461–479. doi:10.1037/0096-1523.22.2.461

Stefanis, N. C., van Os, J., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., & Stefanis, C. N. (2005). Effect of COMT Val158Met polymorphism on the continuous performance test, identical pairs version: Tuning rather than improving performance. *American Journal of Psychiatry*, *162*(9), 1752–1754. doi:10.1176/appi.ajp.162.9.1752

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York, NY: Oxford University Press.

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia*, *39*(8), 771–786. doi:10.1016/S0028-3932(01)00013-6

Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society*, *17*(5), 759–765. doi:10.1017/S1355617711000695

Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 901–915. doi:10.1098/rstb.2007.2096

Stuss, D. T., & Knight, R. T. (2013). *Principles of frontal lobe function*. New York, NY: Oxford University Press.

Stuss, D. T., Murphy, K. J., Binns, M. A., & Alexander, M. P. (2003). Staying on the job: The frontal lobes control individual performance variability. *Brain*, *126*(11), 2363 –2380. doi:10.1093/brain/awg237

Stuss, D. T., Pogue, J., Buckle, L., & Bondar, J. (1994). Characterization of stability of performance in patients with traumatic brain injury: Variability and consistency on reaction time tests. *Neuropsychology*, *8*(3), 316–324. doi:10.1037/0894-4105.8.3.316

Tekin, S., & Cummings, J. L. (2002). Frontal-subcortical neuronal circuits and clinical neuropsychiatry: An update. *Journal of Psychosomatic Research*, *53*(2), 647–654. doi:10.1016/S0022-3999(02)00428-2

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *14*(2), 167–177. doi:10.1016/S0887-6177(97)00095-4

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, *24*(3), 300–315. doi:10.1037/a0018274

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review*, *7*(3), 424–465. doi:10.3758/BF03214357

Wagenmakers, E.-J., & Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological Review*, *114*(3), 830–841. doi:10.1037/0033-295X.114.3.830

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3–22. doi:10.3758/BF03194023

Wager, T. D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: A meta-analysis. *NeuroImage*, *22*(4), 1679–1693. doi:10.1016/j.neuroimage.2004.03.052

Walhovd, K. B., & Fjell, A. M. (2007). White matter volume predicts reaction time instability. *Neuropsychologia*, *45*(10), 2277–2284. doi:10.1016/j.neuropsychologia.2007.02.022

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, *46*(4), 361–413. doi:10.1016/S0010-0285(02)00520-0

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition* (WAIS–IV). San Antonio, TX: The Psychological Corporation.

West, R. (1999). Age differences in lapses of intention in the Stroop task. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 54B*, P34–P43. doi:10.1093/geronb/54B.1.P34

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. *Brain and Cognition*, *49*(3), 402–419. doi:10.1006/brcg.2001.1507

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. doi:10.1037/0033-2909.121.3.331

Williams, B. R., Hultsch, D. F., Strauss, E. H., Hunter, M. A., & Tannock, R. (2005). Inconsistency in reaction time across the life span. *Neuropsychology*, *19*(1), 88–96. doi:10.1037/0894-4105.19.1.88

Williams, L. M., Brammer, M. J., Skerrett, D., Lagopolous, J., Rennie, C., Kozek, K., Gordon, E. (2000). The neural correlates of orienting: An integration of fMRI and skin conductance orienting. *Neuroreport*, *11*(13), 3011–3015. doi:10.1097/00001756-200009110-00037

Winterer, G., Coppola, R., Goldberg, T. E., Egan, M. F., Jones, D. W., Sanchez, C. E., & Weinberger, D. R. (2004). Prefrontal broadband noise, working memory, and genetic risk for schizophrenia. *American Journal of Psychiatry*, *161*(3), 490–500. doi:10.1176/appi.ajp.161.3.490

Winterer, G., & Weinberger, D. R. (2004). Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Trends in Neurosciences*, *27*(11), 683–690. doi:10.1016/j.tins.2004.08.002

Wu, C.-T., Pontifex, M. B., Raine, L. B., Chaddock, L., Voss, M. W., Kramer, A. F., & Hillman, C. H. (2011). Aerobic fitness and response variability in preadolescent children performing a cognitive control task. *Neuropsychology*, *25*(3), 333–341. doi:10.1037/a0022167

Zakzanis, K. K., Mraz, R., & Graham, S. J. (2005). An fMRI study of the Trail Making Test. *Neuropsychologia*, *43*(13), 1878–1886. doi:10.1016/j.neuropsychologia.2005.03.013

Zheng, Y., Myerson, J., & Hale, S. (2000). Age and individual differences in visuospatial processing speed: Testing the magnification hypothesis. *Psychonomic Bulletin & Review*, *7*(1), 113–120. doi:10.3758/BF03210729