August 2023

Pinpointing the Cognitive Structure of the CPT-3: A Case for Distribution Appropriate Statistical Methods

Chandler J. Zolliecoffer
University of Wisconsin-Milwaukee

PINPOINTING THE COGNITIVE STRUCTURE OF THE CPT-3: A CASE FOR DISTRIBUTION APPROPRIATE STATISTICAL METHODS

by

Chandler J. Zolliecoffer

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Psychology

at

The University of Wisconsin-Milwaukee

August 2023

ABSTRACT

PINPOINTING THE COGNITIVE STRUCTURE OF THE CPT-3: A CASE FOR DISTRIBUTION APPROPRIATE STATISTICAL METHODS

by

Chandler J. Zolliecoffer

The University of Wisconsin-Milwaukee, 2023
Under the Supervision of Professor David C. Osmon

The present study used the CPT-3 as a model to illustrate the complementary use of distribution-appropriate statistical methods to analyze non-normally distributed empirical datasets. Study results reaffirmed that while group-level analysis (e.g., via traditional parametric group-level analysis or distribution-appropriate group-level analysis procedures) offers insights into performance of the group in aggregate, it is oftentimes inappropriate to presume that patterns reflected by the group are, necessarily, applicable to a smaller subset of respondents. Thus, understanding how subgroups within the population navigate and approach a given task can have direct implications for more personalized/individualized assessment and treatment, especially in clinical and research contexts. Although results of the investigation are largely supportive of study hypotheses, the study is not without limitation; perhaps most notable is the smaller sample size. Thus, replicating the investigation's significant findings with a larger dataset would prove prudent—especially as the field of neuropsychology moves toward more consistent use of non-normally-distributed computerized reaction time tasks.

To my family:

Jordan, Loretta, James, and Chris

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My arrival at this juncture would not have been possible without the grace of God or the unwavering support of my family, friends, and mentors. To my parents, Loretta and James, thank you for your unconditional love and belief in my ability to realize this dream. To my sister, Jordan, you are my heartbeat and my best friend—thank you for your encouragement through this process and all others. To my love, Chris, I have never felt more cared for than I have by you; thank you for all that you have done and I that you continue to do. To my friends, thank you for showing me balance and holding me accountable in caring for myself; I would not be here without you. To my community, especially Black women: I have completed this endeavor for you, for us. To my mentor: Dr. David Osmon, it is an honor to be your final student. I hold this title with great pride; I hope to continue to make you proud.

Continuous performance tests (CPTs) have long been regarded as important neuropsychological instruments that are highly sensitive to brain dysfunction, but their construct validity has been elusive. First developed by Rosvold and colleagues (1956), different versions of CPTs have been designed and used profitably across varying populations (Riccio et al., 2002). Notably, prior investigations have implicated CPTs to be useful in assessing attention and executive function with involvement of various cortical, subcortical, and functional systems especially involving the right hemisphere (Riccio et al., 2002). Notably, there have been myriad attempts since its development to characterize cognitive aspects of CPTs, including constructs such as: sustained, focused, selective, and shifting attention, in addition to inhibition, hyperactivity/impulsivity, processing speed, and vigilance, among others. The present investigation narrows the scope of this vast literature using the Conners Continuous Performance test, third edition (CPT-3; Conners, 2014) as its instrument of interest.

Pinpointing the construct validity of the CPT-3 remains complex; this is due largely to the ambiguity surrounding the conceptualization of attention and its component parts (e.g., Neumann, 1996). Yet, it is crucial to better understand this complexity as CPT tasks are among the most widely used neuropsychological assessment measures and the second most frequently used instruments in the areas of attention, concentration, and working memory (Rabin et al., 2016). Moreover, CPTs have been demonstrated to differentiate clinical groups which can be helpful for differential diagnosis of attention-based conditions. Thus, clarifying the cognitive structure of the CPT-3 has significant clinical implications.

Still, although identification of the cognitive constructs deemed important for interpretation of CPT-3 results is critical, equally expository is the ability to precisely measure the extent to which such constructs contribute to an individual's performance on the task; and further, which constructs are most important in accounting for performance at both the group- and individual-level of explanation. The present investigation, thus, presents a case for the inclusion of parameterized elementary cognitive tasks (ECTs) given their ability to precisely measure cognitive speed and executive functions—the most fundamental aspects of CPT performance. Importantly, *reaction time* (RT) has a positive relationship with task complexity and therefore, by consequence, *bits of information*—a concept borrowed from the information theory literature. In this way, the amount of time needed to make a choice is both measurable and predictable (Schmitz & Wilhelm, 2016). And while ECTs represent many of the simple iterations of bit-tasks (e.g., simple and 2-choice RT tasks) extant within the literature, other, more complex versions, such as the executively-mediated RT tasks proposed by Miyake et al. (2000), exist too. Accordingly, the construct validity of the CPT-3 can be assessed relative to precisely parameterized simple and complex bit-tasks. It then becomes possible to more definitively and systematically draw conclusions about which cognitive constructs, such as mental speed or executive functioning (EF), operate on any given score or factor at the group-level of analysis, and by any given respondent at the individual-level of analysis. These ideas will be further explicated in the sections below.

Employing task-appropriate analysis methods are an important next consideration. Moreover, to best elucidate the cognitive constructs that underlie the CPT-3, the methodological processes employed must be sound. At its root, the CPT is a mental speed task—one that hinges on RT. As almost all mental speed tasks have marked non-normal (i.e., right-skewed) distributions (Luce, 1986), analyses using these data ought take such psychometric properties into account. As an initial step, the empirical distribution of the CPT-3 dependent variables must be properly fit so as to select statistical tests appropriate for the empirical distribution's shape.

Using appropriate statistical techniques to address these methodological issues ultimately strengthens the certainty with which one can draw conclusions from the data. In addition to using distribution-appropriate statistics, the present investigation acknowledges that using group-level effects to better understand the construct validity of the CPT-3 on the whole is only a first step to data-appropriate analysis methodology. A second, equally important, step is to examine individual-difference effects that operate at the respondent-level of performance. These effects are important to elucidate as CPT tasks are complex and are, therefore, subject to strategy-based performance. That is, while a given respondent might use one strategy in performing a task; another might use a different strategy. For example, on a complex measure like the Stop Signal test, one individual may have a high response threshold, preferring to delay a response until certain their answer is correct. In contrast, another person may emphasize responding in a speedy fashion. In the case of the first individual, inhibition may prove the contributing factor that explains performance while mental speed might be more important in explaining the second person's performance. As a result of these differences in task approach, the cognitive constructs underlying task performance may differ across persons, making group-level analysis of the CPT-3 construct validity applicable for the group in aggregate, but inapplicable for any given individual. Consequently, a complete picture of CPT-3 construct validity only reveals itself when both group-level and individual-level results are taken into account (Leclaire et al., 2020).

Taken together, the present investigation has three specific aims: (1) to identify the cognitive constructs underlying the CPT-3; (2) to measure the relative contribution of the identified cognitive constructs on each factor of the CPT-3 at the group-level of analysis; and (3) to examine the construct validity of the CPT-3 using distribution-appropriate statistical methods to analyze RT data at the individual-level of analysis. Accordingly, the approach for the present study follows two broad themes, namely: first, the cognitive structure of the CPT-3 can be profitably investigated using two general cognitive rubrics, executive functioning (EF) and mental speed; and second, mental speed tasks have non-normal right-skewed distributions, ultimately constraining which statistical techniques ought to be used to analyze RT data. Thus, using predominately late-adolescent and adult ADHD neuropsychology literature, the present study examines the cognitive constructs that underlie the CPT-3 and demonstrates the importance of using distribution-appropriate statistical methods to do so.

Aim 1: Cognitive Constructs of the CPT

CPT-3 Factors

Information about the cognitive abilities underlying the CPT-3 has been best summarized using factor analytic methods (Conners, 2014). Across the literature, four to five factors emerge most consistently including: focused attention, hyperactivity/impulsivity, sustained attention, vigilance, and change in control (see Conners, 2014; Egeland & Kovlik-Gran, 2008; Halperin et al., 1991). These factors are based on 13 outcome scores—many of which prove redundant; accordingly, interpretation of task performance is often challenging. Perhaps most striking, however, is the dearth of construct validity research within the field. To this end, although extant factor-analytic studies have identified which CPT-3 outcome scores hang together to form distinct factors, exactly *what* these factors represent remains yet to be fully realized. Still, it is instructive to review that which has been promulgated about the CPT-3 and its factors. Accordingly, delineated below is a brief summary of each factor as captured by the current state of the literature.

Focused Attention. The *focus* factor capture's one's ability to screen out task-irrelevant stimuli while directing attentional resources toward a given task (Mirsky et al., 1999). This factor is often identified as necessary for engagement in other types of attention (e.g., sustained attention, vigilance) as respondents must first successfully focus attention in order for it to lapse (Egeland & Kovalik-Gran, 2008; Van der Meere, 2002). With regard to CPT-3 test performance, those low on *focus*, and, thus, high on *inattention*, will demonstrate reduced hit rate accuracy, marked variability in hit reaction time (HRT), and increased omissions and perseverations (Egeland & Kovalik-Gran, 2008). Only when inattentive performance is consistently demonstrated across the entirety of the CPT task would a pure deficit in focused attention be considered. In such cases, inattention is correlated with time-on-task and inter-stimulus interval (ISI; Egeland & Kovalik-Gran, 2008).

Hyperactivity/Impulsivity. Egeland and Kovalik-Gran (2008) demonstrate that the *hyperactivity-impulsivity* factor encompasses commission errors, decreased (i.e., fast) RT, and indiscriminate response style. Moreover, although perseverations load onto this factor, they load more strongly onto the *focus* factor highlighting *sensitivity* vs. *specificity* considerations. That is, while *sensitive* to potential impairment, perseverations cannot be used as a *specific* index of performance despite being identified as a measure of impulsivity in the Conners CPT-II manual.

Sustained Attention. Mirsky and colleagues (1999) posit that *sustained attention* reflects one's ability to attend to task-relevant stimuli for an "appreciable interval" while responding to identified targets and inhibiting responses to non-targets. Perhaps most important, however, is that the *sustained attention* factor captures disruptions to one's initial ability to focus attention (Egeland & Kovalik-Gran, 2008). Thus, block change, block change standard error (SE; scored in the CPT-II [Conners, 2002] but dropped in the CPT-3), and change in commission error scores are most commonly associated with this factor.

Vigilance. Although Mirsky and colleagues (1999) conceptualized *vigilance* and *sustained attention* to be one all-encompassing construct, Egeland and Kovalik-Gran (2008) distinguish the two. The latter two authors maintain that *vigilance* reflects one's ability to keep attentional processes engaged even when there is a decrease in target stimuli. This process is associated with HRT ISI, and HRT ISI SE (also dropped in the CPT-3) scores.

Change in Control. Finally, *change in control* evaluates process changes related to impulsivity and changes in mental control where commission errors that change as a function of time on task, and changes in RT over time both load onto this factor (Egeland & Kovalik-Gran, 2008). In this way, respondents who are low on *change in control* gain control over time whereas those who are high on the factor demonstrate increased impulsivity over time (Egeland & Kovalik-Gran, 2008).

Aim 2: Relative Contribution of Cognitive Constructs on CPT-3 Factors

In their investigation, Egeland & Kovlik-Gran (2008) factor-analyzed CPT-2 protocols from a mixed clinical sample to determine if any of the task parameters hang together in a meaningful way; that is, in a way that evidences support for multidimensional cognitive constructs on the CPT-2. Results of the principal components analysis revealed the above-mentioned five viable factors. Factor 1, *Focus*, explained approximately 23% of the variance where omissions, HRT standard error, variability, and perseverations were the salient factor loadings (see Method section for explanation of the CPT-2/CPT-3 scores). Factor 2, *Hyperactivity-Impulsivity*, explained approximately 15% of the variance where commissions, RT, and response style were the most salient factor loadings. Factor 3, *Sustained Attention*, explained approximately 14% of the variance where increased RT, and increased HRT variability as a function of time, in addition to increased omissions, proved the most salient loadings. Factor 4, *Vigilance*, accounts for approximately 13% of the variance where the two ISI change measures were the most salient loadings. Notably, as described above, the authors draw a distinction between *Sustained Attention* and *Vigilance*. Where, traditionally the two have been thought to be tantamount to one another, Egeland & Kovlik-Gran (2008) hypothesize that differences in arousal and activation better characterize differences between vigilance and sustained attention, respectively. Where arousal is thought to be influenced by exposure to both target and non-target stimuli, activation is thought to be influenced by the performance of a motor act (Van de Meere, 2002).

Finally, Factor 5, *Change in Control*, accounts for approximately 9% of the variance where change in commissions over time and, to a lesser degree, change in RT across blocks were the most salient loadings. Although the *Sustained Attention* factor and *Change in Control* factor measure process changes on the task, respondents high on the *Sustained Attention* factor make more inattentive errors where those high on *Change in Control* make more impulsive errors. In this way, those low on the *Change in Control* factor gain mental control over time.

Synthesis of Aim 1 and Aim 2: CPT Factors as Mental Speed and EF Components

Mental Speed

Given the speeded nature of the task, the CPT-3 cognitive structure must be logically examined from the standpoint of cognitive speed as measured by RT (Carroll, 1993). To accurately assess the complexity of cognitive speed, it is important to understand the construct across various difficulty levels within the broader context of neuropsychological assessment. The parameterization of speeded tasks makes this possible by providing a systematic framework through which the reader may understand the contributions of cognitive speed (RT) vs. cognitive power (EF) on task performance. This concept is further delimited below.

Parameterization. RT tasks have been demonstrated to be multifactorial in nature; see Carroll's (1993) two broad speed factors and O'Connor and Burns' (2003) four-factor model for a more detailed review. Thus, given the complexity of the construct, organizing RT tasks by difficulty proves increasingly important. The parameterization of RT tasks offers a formulaic approach to organizing tasks by difficulty. To better understand this process, we look to information theory (IT) and the idea of a *bit* of information. Within this literature, a *bit* represents the amount of information needed to reduce uncertainty by half (Shannon, 1949). It is through this unit of measurement that ECTs are evaluated (e.g., 0-bit, 1-bit, and 1-bit executive tasks, etc., as defined in the Method section). Hick (1952) proposed that RT ought to be considered a "function of the number of bits (log2 n) implied in the number of alternatives" (Carroll, 1993, p. 479). In other words, Hick's law contends that the more choices one has, the longer it will take the individual to reach a decision in a linearly dependent fashion. Thus, the difficulty of a RT task can be parameterized according to RT = $a + b \log_2$ (Roberts & Pallier, 2001). For example, in a simple RT task, there are no choices to be made as the subject simply responds as soon as the stimulus appears. However, in a 2-choice RT task, the subject must make a choice between pressing one of two buttons, depending upon whether the stimulus appears in one of two positions. As a result, a 2-choice RT is slightly longer than a simple RT. Research has shown that each additional choice a subject must make (i.e., 2, 3, 4 choices, etc.) adds approximately 25 ms to the RT above and beyond the simple RT task. As such, RT difficulty can be precisely determined, which becomes important in more complex RT tasks—especially those adding executive processing. For example, the Stroop task, Stop Signal, Navon's Local/Global, and other tasks can be compared to simple mental speed tasks like the simple and 2-choice RT tasks in terms of their level of difficulty.

Executive Functioning

The speeded nature of the CPT-3 is more easily understood while the executive nature of the task has been less readily investigated. Like attention, EF has been thought to be a complex and difficult-to-define construct. The ill-defined nature of EF has resulted in a lack consensus about the true essence of the construct and the components that comprise it (Suchy, 2015). Models of EF have traditionally followed one of two orientations: unitary or multidimensional approaches (Suchy, 2015). Notably, *most* studies that directly compare the two theoretical models find greater support for a multidimensional conceptualization of EF (e.g., Cona et al., 2013; Friedman et al., 2006; Lehto, 1996; Lerner & Lonigan, 2014; M.R. Miller et al., 2012; Tsuchida & Fellows, 2013; and Whitney et al., 2001) than a unitary approach, particularly among adult populations. The current study adopts Miyake's multidimensional approach to EF as it has: (1) evolved out of a systematic program of work, and (2) details a manageable number of EF constructs that can be assessed within a set of mental speed tasks that lend themselves well to parameterization.

Miyake and colleagues (2000; 2012). The shifting of mental sets, updating and monitoring of working memory-related processes and inhibition of dominant, automatic responses have long been hypothesized to be components of EF (Miyake et al., 2000). Therefore, it is important that the reader have a working definition of such terms and also a general understanding of how these executive functions are recruited within the CPT-3.

EF Components and the CPT-3. Operational definitions of these executive functions are as follows: *shifting*, or "attention/task switching" encompasses shifting back and forth between mental sets, tasks, or operations; it requires that the respondent overcome interference presented by a task-irrelevant set, in favor of shifting to engagement with a task-relevant set (Miyake et al., 2000). Notably, Posner and Raichle (1994) argue that neural circuit involvement differs based on the shifting of visual attention (e.g., moving attention serially from number to number on the Trail Making Test- Part A) compared to more executively-mediated shifts (e.g., shifting mental set in go/no-go instructions, as in shifting from a *go* mental set on targets to a *no-go* mental set on non-targets on the CPT-3).

Updating refers to the updating and monitoring of working memory representations often assumed to be instantiated in the prefrontal cortex brain regions. Specifically, updating and monitoring tasks require the respondent to actively store *and* manipulate task-relevant information in mind. In this way, items held in working memory are revised and replaced once the information is no longer relevant (Morris & Jones, 1990). Such processes may be implicated in the CPT-3 where varying ISIs between trials might require refreshing task set instructions on the longer intervals (i.e., ISI Block Change score on the CPT-3). Active working memory representations may also be implicated in the CPT task where *no-go* stimuli (letter ‘X’) occur rarely among frequent *go* targets, necessitating that the mental set (i.e., push the spacebar on all letters *except 'X')* be refreshed in order to inhibit the more frequently occurring *go* response on infrequent *no-go* trials.

Finally, *inhibition* encompasses the respondent’s ability to suppress dominant, automatic responses as is seen in the case of the color-word trials of the Stroop task where the more automatic ‘word’ response must be inhibited in favor of the less automatic ‘color’ response (e.g., saying the ink color RED upon seeing the ‘BLUE’ color-word target). Inhibition is likely required on the CPT-3 where infrequent non-targets (i.e., the letter ‘X’ stimuli) require the ability to suppress the strong *go* response tendency instantiated by the frequent target (i.e., all letters other than ‘X’) proportion. As a result, the task encourages impulsive responding to non-targets as reflected in the CPT-3 commission error score. Note that it is not inconsistent to attribute both executive processes of working memory (as noted in the prior paragraph) and inhibition to the letter 'X' stimuli, as both processes can be operative in a response to such stimuli. In fact, this is exactly how individual differences can occur, making both the group- and individual-level analyses necessary.

Aim 3: Distribution-Appropriate Statistics at the Group- and Individual-Level

As the CPT-3 is, at its core, a mental speed task—one that is predicated on RT and engages higher order executive processes—evaluating its output in a clinically meaningful way proves of great import. Because it is RT-dependent, the task is presumed to have a positively-skewed, non-normal distribution; this is true for most all RT tasks (Luce, 1986). Moreover, as the right tail of empirical RT distributions is reasonably long, Luce (1986) argues that this has a direct impact on the density function of the distribution, accounting for its skewness. Given the right-skewed nature of the empirical RT distribution, some posit that there is an exponential component to the variable's distribution (e.g., McGill, 1963), highlighting the need to select analytic procedures that make sense with the data.

Parametric statistics, which assume a normal distribution, are still, however, commonly employed in the field of psychometric assessment when using RT tasks—despite traditional Gaussian (i.e., normal) analytic approaches, in many cases, rendering a poor fit of the empirical RT data (Balota & Yap, 2011). In the case of the CPT-3, outcome scores are artificially standardized into T-scores. In performing this contrived standardization, the data are made to adopt Gaussian properties despite their non-normal attributes. In the present study, we aim to highlight why this traditional method of analysis is unfavorable and instead offer an alternative approach—the use of raw RT data.

Notably, RT distributions provide information about an individual's unique approach to completing the task, offering a richness to the data that is not well captured by traditional group-level parametric analysis. That is, in traditional group-level analysis, data are averaged across the sample, effectively eliminating any opportunity to observe individual differences in the data and, ultimately, providing an incomplete picture of respondent performance. Use of distribution-appropriate statistical methods are, therefore, paramount for not only accurate interpretation of task performance, but also for discerning individual differences across respondent performance. Most often, RT distributions can be well characterized as an ex-Gaussian distribution defined by three parameters. The first parameter is *mu*—the average RT for the normal component of the ex-Gaussian distribution. It represents the 'average' time for an 'efficient' response. The second parameter is *sigma*; it is analogous to standard deviation in a Gaussian distribution and represents response variability of an efficient response. The third parameter is *tau* which represents the skewed proportion of the ex-Gaussian distribution that is made up of 'inefficient', very long RTs that result from grossly inattentive behavior and 'loss of mental set' on any given trial. See Appendix A for further explication of parameters.

Alternative methods to the traditional parametric analysis have been more readily integrated into the fields of experimental and cognitive psychology where they have been described in detail (e.g., Luce, 1986; Ratcliff, 1979; Ratcliff & Murdock, 1976). Moreover, support for the integration of such methods into clinical practice continues to mount (Osmon et al., 2018; Osmon et al., 2020). In order to adequately pinpoint the cognitive structure of the CPT-3, the statistical methods selected for data analysis must retain their integrity in order to effectively interrogate the complexities of the task and identify which constructs are required for task completion (i.e., recruitment of mental speed and/or higher order executive processes).

Thus, the present study seeks to demonstrate the utility of distribution-appropriate methodologies at the group- and individual-level of analysis for non-normal distributions, as typified in RT data across the ADHD literature. To follow is a brief introduction to distribution-appropriate group-and individual-level analysis by means as exemplified by Generalized Regression and Recursive Partitioning procedures, respectively.

Generalized Regression

Traditional parametric statistics assume normality. As such, alternative methods are needed to effectively model group-level data that are not normally-distributed. Generalized regression (GR) is one such technique. This procedure has two advantages for the present study. First, several different non-normal response distributions can be applied to make a more predictive model when the criterion variable is non-normal, as is the case for CPT-3 variables. Second, by shrinking model coefficients toward zero, GR techniques, such as Elastic Net and Lasso, can be used when the number of predictors is high relative to the number of observations (SAS, 2021). Elastic Net and Lasso techniques are also useful when variables are collinear.

Recursive Partitioning and Decision Tree Analysis

In contrast to GR, Recursive Partitioning procedures allow for the analysis of non-normally distributed data at the individual, rather than group, level. To capture individual differences in performance, Recursive Partitioning procedure splits or *partitions* data on the basis of relationships between predictors and response values generating what is termed a *decision tree* (as in Figure 2; SAS, 2021).

Decision trees are comprised of *nodes* and *branches*, where nodes represent a choice resulting in two mutually exclusive subsets and branches represent outcomes originating from a node (Song & Lu, 2015). The node is chosen based upon brute force determination of which predictor variable and its corresponding value best separates the sample into two groups maximally different in value on the criterion variable. A decision tree grows in size with the aim of increasing the predictiveness of the model at each branch (Yildirim, 2020). Branching ends when the Akaike's Information Criterion reaches a minimum, signaling that further branching is no longer parsimonious.

Predictive Modeling Limitations and Corrections

Predictive modeling techniques offer many advantages, particularly with regard to non-normally distributed data. However, as with any statistical method, they are not without limitation. Perhaps most recognizable: among predictive modeling techniques, hypotheses are not *a priori,* but rather *a posteriori.* As such, these modeling techniques can overcapitalize on chance variance in the sample, posing a threat to replicability of the results. To account for this, the present study uses a k-fold cross-validation resampling procedure. This procedure is performed to evaluate model fit such that the data are divided into *k* subsets where each subset produces an R^2 value for comparison to the original R^2 value.

Study Aims and Hypotheses

In accordance with the above review of the literature, **we designed a study to (1) replicate the factors underlying the CPT-3.** Namely, we predicted that exploratory factor analysis (EFA) of the CPT-3 would replicate the Focus, Hyperactivity/Impulsivity, and Vigilance factors identified by Egeland and Kovalik-Gran (2008) in their factor analysis of the Conners CPT-II. As some ancillary scores that load onto their latter two factors were removed from the CPT-3, we did not expect them to fall out of the EFA.

A second aim of the study endeavored to **(2) measure the relative contribution of the identified cognitive constructs on each CPT-3 factor.** As no precedent literature existed, to our knowledge, to inform specific hypotheses, this aim was necessarily exploratory. Still, some general expectations were able to be specified. First, as the Egeland and Kovalik-Gran (2008) Focus factor centers around RT variability, it is plausible that ECT RT variability would prove important in predicting the Focus factor of the CPT-3. Second, as HRT and Commission errors load onto the authors' Hyperactivity/Impulsivity factor, it is possible that scores reflecting impulsivity or poor inhibition could predict this factor; such scores could include the Stop-Signal SSD, the Modified Stroop task (i.e.., Stroop Negative Priming), and the 2-bit (i.e., 1-bit executive) task. Third, as the CPT-3 task includes both mental speed and executive function abilities, it was expected that both ECT and executively-loaded Miyake tasks from the current predictor variables relate to the CPT-3.

Finally, the third aim of the study **(3) sought to examine the construct validity of the CPT-3 using distribution-appropriate statistical methods (i.e., recursive partitioning with specified decision tree parameters) to analyze RT data at the individual-level of analysis.** Therefore, we predicted that recursive partitioning with specified decision tree parameters would explain more variance than the GR results as it is likely that subgroups of individuals would perform differently than the group as a whole (Osmon et al., 2020). Thus, individual performance differences were expected to be better captured under a protocol sensitive to such differences (Osmon et al., 2020). Specifically, we hypothesized that the first branch of the decision tree would replicate the best GR predictor on each of the CPT-3 factors. Although we expected other variables to elicit branching, to our knowledge, no precedent literature existed at the time to inform *how* these variables would branch, rendering specific predictions moot.

Method

Participants

Participants were a mixed sample of n = 73 undergraduate students (ages 18-63; M_{age} = 25.52, SD = 9.58; 39 female) from a large public university in Milwaukee, WI. Participants were recruited from a larger pool of self-referred individuals seeking evaluation for learning disorder (LD) and ADHD at a University learning disability clinic. The sample comprised participants with the following diagnoses: ADHD (n = 16), LD (n = 28), psychiatric (n = 15), missing (n = 9) and no warranted diagnosis, (n = 5). Participants consented to inclusion in research activities at the time of the clinical evaluation. Approval for collection of these clinical data to be included in future research studies was granted by the University's Institutional Review Board. Participants selected for inclusion in the study passed all effort subtests of the Medical Symptom Validity Test (Green, 2004) and completed a large battery of intellectual, achievement, neuropsychological, personality, and experimental tasks.

Instruments

Elementary Cognitive Tasks (ECTs)

Elementary cognitive tasks comprise four 120-trial RT tasks parametrically-scaled in difficulty. The *bit* methodology is comprehensively described in Jensen (2006, p. 27), but, in effect, consists of scaling RT tasks on the basis of how many *bits* of information are required to complete the task. As earlier referenced, a *bit*, in information theory, represents the amount of information needed to reduce uncertainty by half (Shannon, 1948).

0-bit. As is true of all simple RT tasks, the one used in this study (i.e., 0-bit) requires zero bits of information to make a response, as there is no uncertainty. That is, upon immediate presentation of the stimulus (i.e., large black dot) on the screen, the participant is to push the designated response key (i.e., space bar). In this way, there is only one clear response. We may use this same scaling framework to understand *choice* RT tasks.

1-bit. In a 2-choice RT task (i.e., 1-bit), a single bit of information is required to arrive at the appropriate response. In this task, a large black dot appears on either the right or left side of the screen and the participant must decide which of two response keys to press (i.e., right or left key), effectively requiring a single decision to reduce uncertainty by half. Notably, all four ECT tasks in the proposed study have the same stimulus array where a large black dot appears on either the right or left side of the computer screen. In this way, variability between tasks is limited to differences in task instructions.

1-bit Executive (i.e., 2-bit). The final two parametrically-scaled RT tasks, complex choice (i.e.,1-bit executive) and cognitive control (i.e., 2-bit executive), have a different quality about them. These tasks are much harder in the sense that participants must maintain an "internal rule" to arrive at the correct response. Given their relative difficulty and the cognitive efforts needed to complete them, these tasks are thought to be executively mediated resulting in significantly longer RTs than the aforementioned 0- and 1-bit tasks. Specifically, on the 1-bit executive task, the appropriate response is opposite that of where the stimulus is presented on the computer screen. For example, if the stimulus appears on *left* side of the screen, the participant should push the *right* arrow key, following task instruction parameters. Notably, one *bit* of information is required to complete this task such that, upon presentation of the stimulus on the right side of computer screen, uncertainty about the appropriate response is reduced by half. Importantly, however, the cognitive operation requires an *opposite side* mental set response implicating inhibition of the prepotent *same side* response stereotype. It is this cognitive manipulation that highlights the executive component of the task. Moreover, this task is analogous to the anti-saccade task identified by Miyake & Friedman (2012) as an inhibition factor task. Notably, in the present study, the 1-bit Executive task is labeled "2-bit" for clarity.

2-bit Executive (i.e., 3-bit). The cognitive control task (i.e., 2-bit executive) is harder yet as it requires two *bits* of information to solve. Upon arriving at a response decision (i.e., same or opposite side key response) as in the 1-bit executive task, participants must alternate each subsequent response between a "same-side/ opposite-side" rule for the 2-bit executive task. In this way if the participant provided a *same-side* response on trial *n,* then an *opposite-side* response is indicated on trial *n+1*. One bit of information is required to determine which is the same vs. opposite side of the stimulus, and another bit is needed to ensure the participant maintains response set in order to respond accurately (same or opposite side) on the subsequent trial. Again, in the present study, the 2-bit Executive task is labeled "3-bit" for clarity.

As outlined above, the parameterization of the bit tasks included in the proposed study allow for the precise measurement of each cognitive construct. Notably, the non-executive simple and choice RT tasks scale linearly in mental speed (approximately 25 ms/bit, depending upon the task and population [Jensen, 2006, p. 53- see slope in the equation of Figure 3.3]), while the harder, executively-mediated RT tasks scale nonlinearly. This nonlinear increase in speed is presumably related to the executive nature of the more difficult tasks.

Miyake Factor Tasks

Miyake tasks have been well-documented in the literature to measure various aspects of EF including: inhibition, updating/monitoring, and shifting. The current study includes measures that assess each factor including: modified Stroop (i.e., Stroop Negative-Priming) and Stop-It task (inhibition), Keep Track and *N*-back tasks (updating/monitoring), and Navon's Local-Global and Flanker tasks (shifting). Each task is further described below.

Inhibition

Modified Stroop. The modified Stroop task (i.e., Stroop Negative-Priming) is a computerized version of the classic Stroop paradigm (Golden, 1978; Stroop, 1935)). In this version of the task, a stimulus is presented on the screen, one at a time, and remains until the participant makes a response. On the word-reading block of the task, color-words (i.e., red, blue, or green) appear in black ink across 50 trials. Participants are to select the response key (i.e.., red = left arrow, blue = down arrow, green = right arrow) that corresponds with the appropriate color-word. As in the Golden Stroop task, the color block follows the word-reading block. In this 50-trial block, asterisks are presented to the participants in red, blue, or green font colors. Participants are to then select the response key that corresponds with the appropriate stimulus color.

The color-word condition incorporates negative, positive, and neutral prime trials. Negatively primed trials occur when the inhibited response for trial *n* becomes the correct response for trial $n + 1$. For example, if in trial *n* the word 'blue' is displayed on the screen in red ink (blue is inhibited; red is the response), for negative priming to take place, the word (e.g., green) on trial $n + 1$ must be displayed in blue ink (green is inhibited; blue is the response). Because the response of the $n + 1$ trial was previously inhibited, RT for negatively primed trials are considered to be more executively-mediated and are expected to result in slower RTs. By contrast, positively primed trials occur when the same correct response is indicated across two consecutive trials. For example, if on trial *n* the word 'blue' is displayed on the screen in red ink (blue is inhibited; red is the response), in order for positive priming to take place, the word on trial $n + 1$ (e.g., green) must be displayed in red ink, too. As this response was just recently activated on trial *n,* it is expected that trial $n + 1$ would result in faster RTs. Across the color-word condition, there are 18 negatively primed trials, 18 positively primed trials, and 36 neutral trials. Neutral trials occur before priming trials in order to ready the participant for the next priming trial and are the only trials to display color-words in red, blue or green font color. All priming trials displayed the word 'yellow' in an ink color consistent with the desired priming color for that trial. In using the modified priming Stroop, detection of inhibition sensitivity is better assessed.

Stop-It. The Stop-It task is an independent computerized measure of the stop-signal paradigm that evaluates response inhibition and impulse control (Verbruggen et al., 2008). Specifically, the task assesses the type of inhibition required to stop an already activated stereotyped motor action that typically occurs during a *go* task. This type of inhibition is different from the type of inhibition observed on the traditional paper-and-pencil Stroop paradigm where a more preferred stereotype (i.e., reading a word) must be inhibited in favor of a less preferred response (i.e., naming the ink). The Stop-It task is comprised of two 64-trial blocks. Across each trial, participants are presented with a stimulus in the center of the computer screen (i.e., square or circle) at which point participants must select the key that corresponds to the presented stimulus (i.e., 'z' for square or '/' for circle) as quickly and as accurately as possible. The stimulus remains on the screen until a response is made *or* until 1250ms have elapsed. Stop-signal trials occur 25% of the time. During such trials, a 75ms 750 Hz tone will sound following stimulus onset signaling that participants should inhibit their motor response. Presentation of the tone on the first stop-signal trial will occur 250ms following display of the stimulus. Subsequent stop-signal delays (SSDs) will adjust based on performance; if successful inhibition of the motor response occurs on trial n, then SSD will *increase* by 50ms on stop-signal trial $n + 1$. In this way, successful inhibition of the response results in the stop-signal tone playing further away from the initial presentation of the stimulus making it more difficult to inhibit a motor response. By contrast, if inhibition of the motor response on trial *n* is unsuccessful, then SSD will *decrease* by 50ms on the *n* +1 stop-signal trial resulting in the stop-signal tone playing closer to the initial presentation of the stimulus making it easier to inhibit a motor response (Verbruggen et al., 2008). In modulating the SSD, the SSD RT is identified at which the participants will inhibit their response 50% of the time. The stop signal RT is obtained by subtracting the SSD from the average RT on no-signal trials. This metric provides an estimate of how long it takes the participant to inhibit a response.

Updating/Monitoring

Keep Track. The Keep Track task used in the current study is adapted from Yntema (1963) and measures one's ability to update working memory representations. This is achieved by attending to task-relevant information (monitoring) while also disregarding task-irrelevant information once deemed no longer appropriate (updating). Participants are first presented with six possible categories at the bottom of a computer screen including: animals, colors, countries, distances, metals, and relatives. Fifteen words comprising 2-3 exemplars from each category are then presented in random order at a rate of 1500ms/stimulus while the target categories remain at the bottom of the screen. Participants are tasked with recalling the *last* word presented from each of the *target* categories ($n = 4$) for any given trial. Participants must then record their responses at the end of the trial where percentage correct serves as the index of performance. For instance, if color were one of the target categories and the participants first saw "green", and later "blue" with no other color words presented later still, then the participant ought record "blue" at the end of the trial. To ensure comprehension of the task, participants are first shown all six categories and all of the exemplars from each category. Participants then complete a single trial run with three categories for practice.

N-Back. The proposed study also incorporates an *N*-back task (Kirchner, 1958) to assess updating and monitoring abilities. On the task participants are presented with a sequence of letters one at a time. Participants must decide if the current stimulus is the same or different from the one presented *n* (i.e., 0, 1, 2, or 3) trials ago by making a button-press response that corresponds to the appropriate choice (i.e., 'N' for same as *n* trials ago or 'M' for different as *n* trials ago) where, again, percentage of correct responses serves as the index of performance.

Shifting

Local-Global. Shifting represents one's ability to effectively shift between tasks, operations, and/or mental sets (Monsell, 1996). The Local-Global task originally developed by Navon (1977) is used to measure shifting using RT and percentage correct at both the local and global levels. In this task, large 'S' and 'H' letters are composed of small 's' and 'h' letters with blocked trials of selective attention to the local and the global perceptual levels. Participants are expected to respond to either large or small letters on any given trial, ignoring the other type per instruction parameters. Trials include an equal number of congruent stimuli (e.g., a large 'S' made up of small 's' letters) and incongruent stimuli (e.g., a large 'S' made up of small 'h' letters). Participants complete three 36-trial blocks where local letters are the target stimuli on the first and third blocks and are global letters are the target stimuli on the second block.

Flanker. Originally developed by Eriksen and Eriksen (1974), the Flanker task is a measure of attentional switching. It is comprised of two 20-trial blocks wherein participants are presented with five arrows displayed in the center of the screen. Participants must respond appropriately to the middle stimulus "flanked" by either congruent or incongruent stimuli. That is, participants must select the corresponding key on the computer using the index finger of their dominant hand. RT and percentage correct serve, again, as the index of performance.

Attention Tasks

CPT-III. The CPT-3 is a commonly used computerized measure of attention that indexes performance across areas of inattention, impulsivity, sustained attention, and vigilance (Conners, 2014). On this task, stimuli (i.e., letters) are presented in the center of the computer screen, one at a time. Participants are instructed to respond (i.e., push the spacebar) to all letters *except* 'X'. How long the stimulus is displayed on the screen and the length of time between trials (i.e., inter-stimulus interval [ISI] is intentionally variable so as to evaluate aspects of inattention, impulsivity, sustained attention, and vigilance. The following output scores are obtained upon completion of the 360-trial task administration: detectability (d'), omissions, commissions, perseverations, hit reaction time (HRT), HRT standard deviation (HRT SD), variability, HRT block change, omissions by block, commissions by block, HRT inter-stimulus interval change (HRT ISI change), omissions by ISI, and commissions by ISI.

Procedure

As noted above, participants were a clinically-mixed sample of self-referred undergraduate students seeking evaluation for LD/ADHD at a university learning disability clinic. Clients were consented for participation in research activities at the time of clinical evaluation. Following the consenting process, the examiner read a standardized introduction script to the clients describing what to expect over the course of the full-day evaluation. Clients were oriented to the types of tests they would complete during the evaluation, and were encouraged to try their hardest to ensure valid, interpretable test data. The battery comprised a number of cognitive, achievement, neuropsychological, personality, and experimental tests. The examiner facilitated administration of the tests (e.g., reading test instructions, providing feedback where appropriate, etc.). All tests were completed in a fixed order to control for order effects across research subjects. One scheduled 45-minute lunch break occurred approximately halfway through the evaluation; shorter unscheduled breaks were allowed between subtests. The treating clinical neuropsychologist completed a 30-60-minute clinical interview with the client near the session's end. Upon completion of the evaluation, the examiner scored the tests and gave the results to the provider. A separate date for feedback was selected where the provider would relay the findings and recommendations to the client. After scheduling the feedback session, the client was then dismissed. In total, the clinical evaluation lasts approximately 8 hours.

Data Analysis

All data analysis was completed using JMP 15 Pro (SAS, 2019). To test the first aim of the study, the EFA used principal components analysis with orthogonal varimax rotation in order to obtain theoretically independent CPT-3 factors. The following CPT-3 variables were included in the model: Omissions, Commissions, HRT, HRT SE, Variability, Detectability (*d'*), Perseverations, HRT Block Change, and HRT ISI Change. The number of factors were preset at three. Only factors attaining eigenvalues of 1.00 or higher were retained (where Factor 3 was rounding up) and only factor loadings greater than .30 were considered significant loadings. Consistent with the Egeland and Kovalik-Gran (2008) variables included in this study, we expected to replicate the Focus, Hyperactivity/Impulsivity, and Vigilance factors.

Hypotheses for the second aim were evaluated using generalized regression (GR). GR can select various response distributions to model empirical data that have a right-skew (e.g., LogNormal, Weibull, etc.); it also adds a penalty term which forces coefficients to zero for non-contributory predictors. In the present study, a GR was run on each CPT-3 factor where the ECT and Miyake variables described in the Instruments subsection above served as predictors. We expected that each CPT-3 factor would have a unique empirical distribution that was likely to be right-skewed in shape. As such, selection of an appropriate response distribution proved prudent in order to appropriately model the identified CPT-3 factor. Response distribution models were selected after analyzing each CPT-3 factor distribution. Adaptive elastic net version GR models were used in the present investigation as they include the penalty advantages of both ridge and lasso regression techniques. Moreover, their statistical properties are superior to the elastic net option. Contributing predictors were selected based not only upon their statistical significance (i.e., Wald χ^2), but also based on their Variable Importance values which are analogous to beta coefficients in standard regression models. The success of the GR was judged based upon generalized R^2 values, which are analogous to standard R^2 values in multiple regression. The third hypothesis for aim 2 was assessed by evaluating whether both mental speed ECTs and Miyake EF scores contribute to predicting CPT-3 factors.

Finally, recursive partitioning (RP) was used to examine the two hypotheses for Aim 3. First, the R^2 values from the GR and RP analyses were compared across their respective analyses for each CPT-3 factor. We expected that RP would predict more variance than GR as participants were expected to differ in their approach to completing all factors CPT-3 task across. As a result of its complexity, we predicted several branches on both sides of the decision tree (‘good’ and ‘bad’ performance on the CPT-3 factors), indicating substantial individual differences in performance across all factors of the CPT-3. Second, we expected that the GR would yield reliable and replicable results, which would be reflected in the most predictive variable for each CPT-3 factor as the first node in the respective RP. Absent precedent literature we had no predictions for subsequent nodes on any of the RP analyses.

Results

Exploratory Factor Analysis

Table 1

Variance Explained

Factor	Eigenvalue	Variance	Percent	Cum. %
1	4.44	2.93	29.26	29.26
2	2.44	2.59	26.86	55.12
3	0.93	2.30	22.97	78.09

Note. Results of the EFA yielded three distinct factors as identified by the above eigenvalues.

Table 2

Rotated Factor Loadings

	Factor 1		Factor 2		Factor 3	
Variable	**Variability**	**Focus[a]**	**H/I**	**H/I[a]**	**Inattention**	**Sustained Attention[a]**
RT-SE	**0.85[b]**	**.843[b]**	0.07	.265	**0.43**	.145
Variability	**0.79[b]**	**.870[b]**	-0.17	.086	**0.30**	.133
RT-Block	**0.79**	-	-0.13	-	0.05	-
RT ISI	**0.69**	.097	0.17	.021	0.18	.014
Persev. Err.	**0.38[b]**	**.768[b]**	**-0.49**	.612	**0.61**	.019
Omission Err.	**0.36[b]**	**.747[b]**	**-0.37**	.255	**0.78**	.224
RT	**0.34**	.363	**0.80[c]**	**.760[c]**	0.25	.049
Resp. Style	0.21	.206	0.12	**.688[c]**	**0.92**	.089
Commission Err.	0.16	.335	**-0.90[c]**	**-.804[c]**	0.10	.048
d'	-0.06	-	**0.82**	-	-0.30	-
Block Change SE	-	.162	N/A	.047	N/A	**.842**
Block Change	-	.068	N/A	-	N/A	**.731**
Δ Omissions	-	-.137	N/A	-	N/A	**-.707**

Note. This table provides the rotated factor loadings for each factor alongside corresponding factors and factor loadings from Egeland and Kovalik-Gran (2008). Significant loadings are bolded. Variables from the present study reflect raw scores while Egeland and Kovalik-Gran reflect standardized T-scores.

[a] Egeland and Kovalik-Gran (2008) Factors and rotated factor loadings; see Egeland and Kovalik-Gran (2008) Table 3.

[b] Corresponding significant Factor 1 rotated factor loadings across studies

[c] Corresponding significant Factor 2 rotated factor loadings across studies

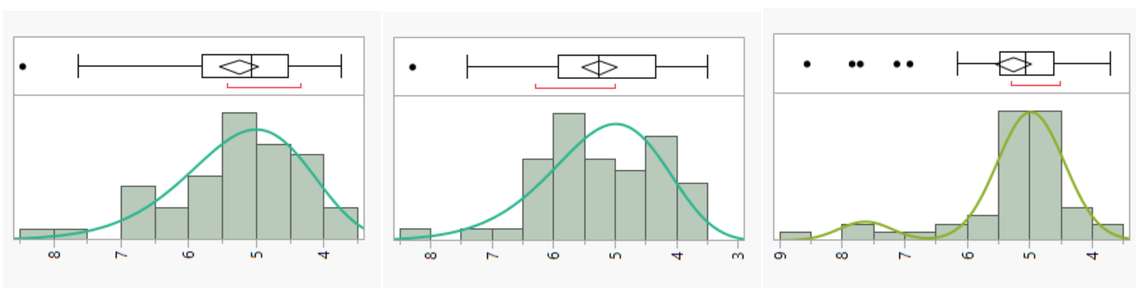
Distributional Analyses

A Lognormal distribution provided the best fit of the empirical distributions for the Variability and Hyperactivity/Impulsivity factors (AIC-c = 113.47 and 13.67, respectively). The Normal-2 Mixture distribution provided the best fit of the empirical distribution for the Inattention factor (AIC-c = 126.71). See *Figure 1*.

Figure 1

Distributions

Variability *H/I* *Inattention*



Note. Lognormal AICc = 133.47 Lognormal AICc = 136.67 Normal-2 Mixture AICc = 126.71

Generalized Regression

In order to identify the best predictors of the Variability, Hyperactivity/Impulsivity, and Inattention factors, respectively, three adaptive elastic net generalized regressions were performed.

Variability

An adaptive elastic net generalized regression with LogNormal response distribution model and AICc was performed yielding generalized $R^2 = 0.25$. One of 30 variables were significant (i.e., 2-bit error) where Wald χ^2 (1, *69*) = 29.92, p < 0.01.

Hyperactivity/Impulsivity

An adaptive elastic net generalized regression with LogNormal response distribution model and AICc was performed yielding generalized $R^2 = 0.67$. Four of 30 variables were significant (i.e., 2-bit mu, 3-bit error, 1-bit sigma, and Stroop NP-RT, in respective order). Wald $\chi^2_{\text{2-bit mu}}$ (1, 69) = 12.55, p < 0.01. Variable Importance Main Effect $_{\text{2-bit mu}}$ = 0.42. Wald $\chi^2_{\text{3-bit error}}$ (1, *69*) = 9.42, p < 0.01. Variable Importance Main Effect $_{\text{3-bit error}}$ = 0.15. Wald $\chi^2_{\text{1-bit sigma}}$ (1, *69*) = 12.27, p < 0.01. Variable Importance Main Effect $_{\text{1-bit sigma}}$ = 0.12. Wald $\chi^2_{\text{Stroop NP-RT}}$ (1, 69) = 8.02, p < 0.01. Variable Importance Main Effect $_{\text{Stroop NP-}}$= 0.09.

Inattention

An adaptive elastic net generalized regression with *t*(5) response distribution model and AIC-c was performed yielding Generalized R^2 = 0.13. One of 30 variables were significant (i.e., 1-bit tau) where Wald χ^2 (1, *69*) = 6.15, p < 0.05.

Decision Trees

An RP procedure was used to evaluate which predictor variable and its corresponding value best separated a given sample (i.e., Variability, Hyperactivity/Impulsivity, or Inattention) into two groups maximally different in value on a criterion variable. Thirty variables were used as classifying predictors including: five ex-Gaussian parameters (i.e., Mu, sigma, tau, error, and coefficient of variation [CoV]) across all bit-tasks (i.e., 0-bit, 1-bit, 2-bit, and 3-bit), in addition to RT and accuracy parameters on the identified Miyake tasks.

Variability

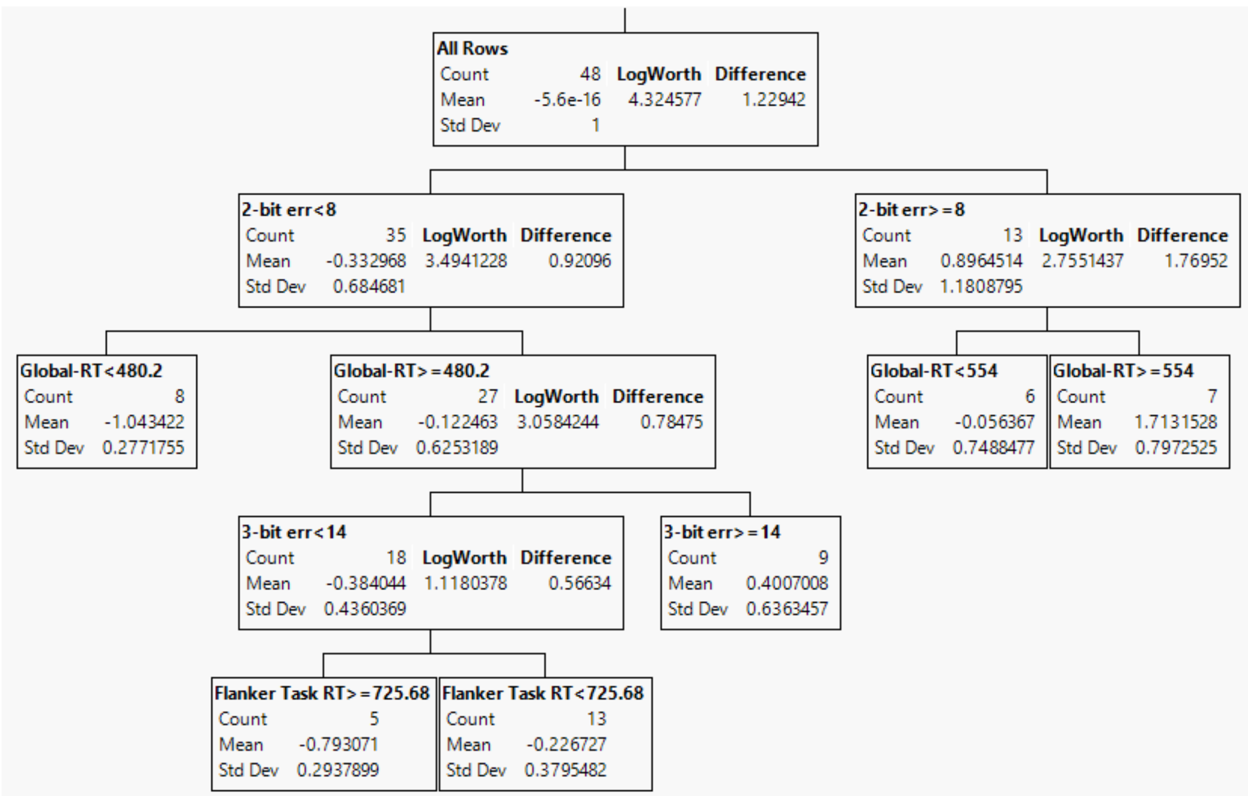
Table 3

Variability Decision Tree Attributes

R^2	Num. of Splits	AICc	k-fold R^2
0.74	**5**	**88.32**	**0.66**

Figure 2

Variability Decision Tree



Note. Decision Tree showing that Variability is first partitioned by 2-bit error, the criterion variable, where 8 errors represents the cut-off value. A *lower* 2-bit error score (<8 errors) is associated with less Variability (M = -0.33). A *higher* 2-bit error score (>8 errors) is associated with greater Variability (M = 0.90). Variability is further partitioned by Global-RT, 3-bit error, and Flanker Task RT performance.

Variability. Five splits yielded an optimum prediction value (AIC-c = 88.32). Overall, the model yielded R^2 = 0.74 and RASE = 0.51. A k-fold cross-validation analysis revealed 8%

shrinkage in the R^2 value. Four variables contributed to the classification including: 2-bit error, Global-RT, 3-bit error, and Flanker Task RT.

Variability is first partitioned by 2-bit error, where 8 errors represents the cut-off value (~79th percentile; M= 5 errors; Median = 3 errors)). Greater Variability (M = 0.90) is associated with a *higher* 2-bit error score (>8 errors). This groups is further partitioned by Global-RT where 554 ms represents the cut-off value (~51st percentile; M= 601ms; Median = 551ms). A *slower* Global-RT score (>554 ms) is associated with even greater Variability (M = 1.71). By contrast, a *faster* Global-RT score (<554) is associated with less Variability (M = -0.06).

Although Variability is most elevated when 2-bit errors are *greater*, the data reveal gradations in performance even when 2-bit errors are *fewer*. In this way, when 2-bit error score is < 8, the sample is again partitioned by Global-RT, however, now, with a cut-off value of 480.2 (~23rd percentile). A *slower* Global-RT score (>480.2 ms) is associated with greater Variability (M = -0.12) relative to a *faster* Global-RT score (<480.2 ms) which is associated with less Variability (M = -1.04).

The sample is partitioned again by 3-bit error where 14 errors represents the cut-off value (75th percentile; M = 11 errors; Median = 4 errors;). A *higher* 3-bit error score (>14 errors) is associated with greater Variability (M = 0.40) relative to a *lower* 3-bit error score (<14 errors) which is associated with less Variability.

The final split of the data is partitioned by Flanker Task RT where 725.68 ms represents the cut-off value (~81st percentile; M = 581ms; Median = 562 ms). A *faster* Flanker Task RT (<725.68) is associated with greater Variability (M = -0.23) relative to a *slower* Flanker Task RT (> 725.68) which is associated with less Variability (M = -0.79).

Hyperactivity/Impulsivity

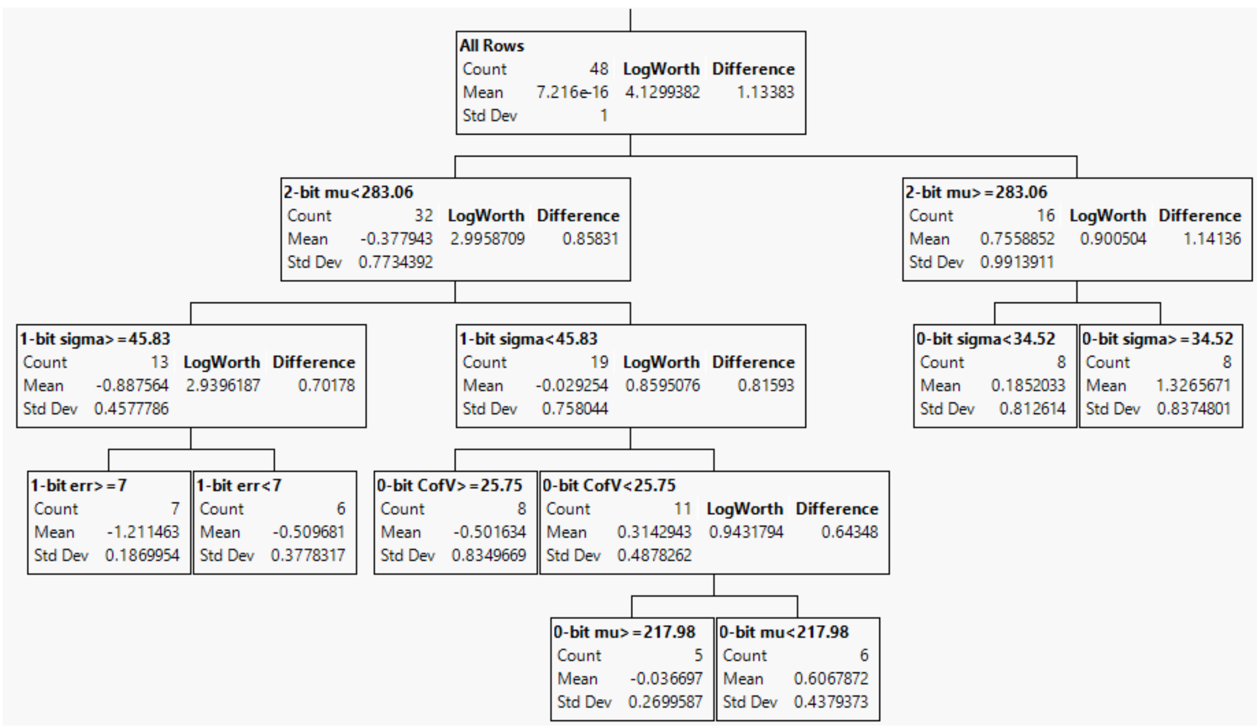
Table 4

H/I Decision Tree Attributes

R^2	Num. of Splits	AICc	k-fold R2
0.65	**6**	**104.90**	**0.56**

Figure 3

H/I Decision Tree



Note. Decision Tree showing that H/I performance is first partitioned by 2-bit mu, the criterion variable, where 283.06 errors represents the cut-off value. A *faster* 2-bit mu score (<283.06 ms) is associated with decreased H/I (M = -0.38). A *slower* 2-bit mu score (>283.06 ms) is associated with increased H/I (M = 0.76). H/I is further partitioned by 1-bit sigma, 1-bit error, 0-bit CofV, and 0-bit mu scores.

Hyperactivity/Impulsivity. Six splits yielded an optimum prediction value (AIC-c = 104.90). Overall, the model yielded $R^2 = 0.65$ and RASE = 0.59. A k-fold cross-validation analysis revealed 10% shrinkage in the R^2 value. Six variables contributed to the classification including: 2-bit mu, 0-bit sigma, 1-bit sigma, 1-bit error, 0-bit CoV, and 0-bit mu.

Hyperactivity/Impulsivity is first partitioned by 2-bit mu, where 283.06 ms represents the cut-off value (~51st percentile; M = 287ms; Median = 281ms). Increased H/I (M = 0.76)) is associated with a *slower* 2-bit mu score (>283.06 ms). This group is further partitioned by 0-bit sigma where 34.52 ms represents the cut-off value (~63rd percentile; M = 34ms; Median = 28ms; 75th percentile = 41ms). A *slower* 0-bit sigma score (>34.52ms) is associated with even greater H/I (M = 1.33). By contrast, a *faster* 0-bit sigma score (<34.52ms) is associated with less H/I (M = -0.06).

Although H/I is most elevated when 2-bit mu performance is *slower*, the data reveal gradations in performance even when 2-bit mu performance is *faster.* When 2-bit mu score is < 283.06 ms, H/I is decreased (M = -0.89) relative to when the 2-bit mu score is >283.06ms, (M = 0.76). The sample is further partitioned by 1-bit sigma where 45.83 ms represents the cut-off value (~69th percentile; M = 38ms; Median = 34ms). A *slower* 1-bit sigma score (>45.83 ms) is associated with decreased H/I (M = -0.89) relative to a *faster* 1-bit sigma score (<45.83 ms) which is associated with increased H/I (M = -0.03).

When 1-bit sigma scores are faster (<45.83 ms), the sample is partitioned even further. It is first partitioned by 0-bit CoV, where 25.75 represents the cut-off value (~43rd percentile; M = 33ms; Median = 28ms). A *slower* 0-bit CoV score (>25.75ms) is associated with decreased H/I (M = -0.50) whereas a *faster* 0-bit CoV score (<25.75ms) is associated with increased H/I (M = 0.31). The final split of the data is partitioned by 0-bit mu where 217.98 ms represents the cut-off value (~29th percentile; M = 236ms; Median = 227ms). A *slower* 0-bit mu score (>217.98 ms) is associated with decreased H/I (M = -0.04) relative to a *faster* 0-bit mu score (> 218.98 ms) which is associated with increased H/I (M = 0.61).

Inattention

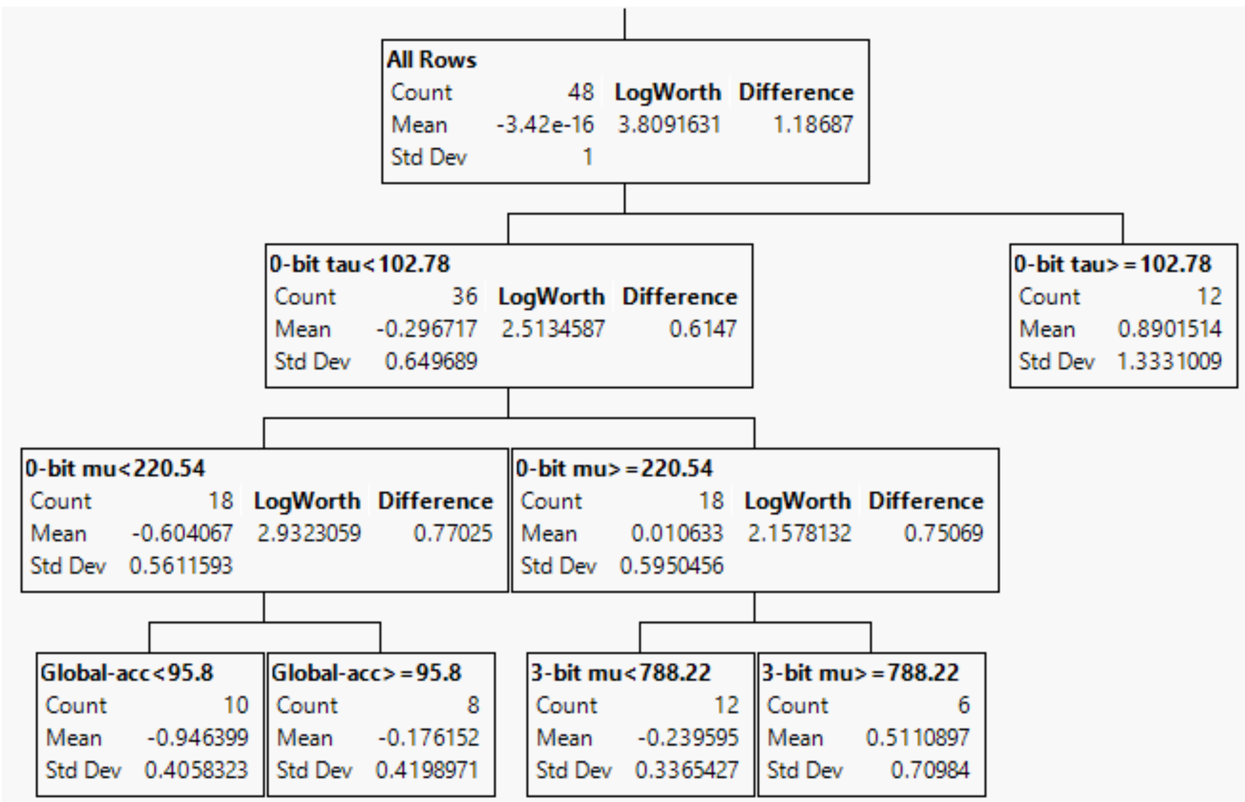
Table 5

Inattention Decision Tree Attributes

R^2	Num. of Splits	AICc	k-fold R2
0.44	**4**	**120.89**	**0.34**

Figure 4

Inattention Decision Tree



Note. Decision Tree showing that Inattention is first portioned by 0-bit tau, the criterion variable, where 102.78 ms represents the cut-off value. A *faster* 0-bit tau score (<102.78 ms) is associated with decreased Inattention (M = -0.30). A *slower* 0-bit tau score (>102.78 ms) is associated with increased Inattention (M = 0.89). Inattention is further partitioned by 0-bit Mu, Global accuracy, and 3-but Mu performance.

Inattention. Four spits yielded an optimum prediction value (AIC-c = 120.89). Overall, the model yielded R^2 = 0.44 and RASE = 0.74. A k-fold cross-validation analysis revealed 10% shrinkage in the R^2 value. Four variables contributed to the classification including: 0-bit tau, 0-bit mu, Global-acc, and 3-bit mu.

Inattention is first partitioned by 0-bit tau, where 102.78 ms represents the cut-off value (~76th percentile). Increased Inattention (M = 0.89) is associated with a *slower* 0-bit tau score (>102.78 ms). By contrast, a *faster* 0-bit tau score (<102.78 ms) is associated with decreased Inattention (M = -0.30).

Although Inattention is most elevated when 0-bit tau performance is *slower*, the data reveal gradations in performance even when 0-bit tau performance is *faster*. In this way, when 0-bit tau score is <102.78, the sample is again partitioned by 0-bit mu with a cut-off value of 220.54 (~36th percentile). A *slower* 0-bit mu score (>220.54 ms) is associated with increased Inattention (M = 0.01) relative to a *faster* 0-bit mu score (<220.54.2 ms). A final split for this portion of the sample is partitioned by Global-acc with a cut-off value of 95.8 where decreased Inattention (M = -0.95) is associated with a *faster* Global-acc score (<95.8). By contrast, a *slower* Global-acc score is associated with increased Inattention (M = -0.18).

When 0-bit mu performance is slower (>220.54 ms), Inattention is increased (M = 0.01) relative to faster 0-bit mu performance (M = -0.60). A final split for this portion of the sample is partitioned by 3-bit mu with a cut-off value of 788.22 ms (~83rd percentile) where decreased Inattention (M = -0.24) is associated with a *faster* 3-bit mu score (<788.22 ms). By contrast, a *slower* 3-bit mu score (>788.22 ms) is associated with increased Inattention (M = 0.51).

Discussion

The present study examined the cognitive constructs that underlie the CPT-3 and presents a case for using distribution-appropriate statistical methods to do so. Specifically, the investigation evidenced clear results with respect to the study's three main aims: (1) to identify the cognitive constructs underlying the CPT-3; (2) to measure the relative contribution of the identified cognitive constructs on each factor of the CPT-3 at the group-level of analysis; and (3) to examine the construct validity of the CPT-3 using distribution-appropriate statistical methods in order to analyze raw RT data at the individual-level of analysis. Overall, results of the investigation are largely supportive of the study hypotheses meriting support for the use of raw RT scores in place of transformed standardized scores when analyzing the CPT-3. In doing so, study findings revealed a more nuanced and complete picture not only of the cognitive structure of the CPT-3, but also of respondent performance at both the group- and individual-level of analysis. Where predictions are less well supported by the data, we offer up alternative explanations for observed discrepancies between hypotheses and the empirical findings. The subsequent sections provide an explication of the investigation results.

Group-Level Analysis Discussion

Factor Analysis

In order to discern the cognitive constructs underlying the CPT-3, we performed an exploratory factor analysis, addressing the first aim of the study. From this analysis, three factors emerged: Variance, Hyperactivity/Impulsivity, and Inattention (see Table 2). In order to adequately compare results between studies, it is important to note the differences in variables used on the Conners CPT-II compared to those adopted by the more current iteration of the task. As earlier described, *block change, block change SE, changes in commissions, and HRT ISI SE,* though included on the CPT-II, were dropped in the CPT-3. Thus, the factors identified by the present study were, where appropriate, redefined and labeled in consideration of their significant rotated factor loadings. That raw scores were used in the EFA represent a second important consideration—the implications of which are further delimited below.

Table 2 provides a visual representation of the rotated factor loadings common to each factor, as identified by the present study and those from Egeland and Kovalik-Gran (2008). As predicted, all four of the significant Egeland and Kovalik-Gran Factor 1 (Focus) variables loaded significantly onto the first factor (i.e., Variability) of the present study. Perhaps more interesting, however, is that these four significant variables loaded significantly onto the third factor (i.e., Inattention) of the present study, too. This finding suggests that the Egeland and Kovalik-Gran Focus factor is an inappropriate mixture of Variability and Inattention. The scores that underlie the collapsed factor reflect two related, but ultimately distinct constructs. In using raw RT scores, the present study is better able to tease apart these differences. That is, RT SE and variability scores reflect primary loadings on the Variability factor while perseverations and omission errors reflect secondary loadings. The inverse is true on the Inattention factor; that is, perseverations and omission errors reflect primary loadings while RT-SE and variability reflect secondary loadings. Thus, the evidence for use of raw RT scores to better parse nuances in the data is clear.

Of the three significant Egeland and Kovalik-Gran (2008) Factor 2 variable loadings, two emerged as significant on the second factor of the present study. In this way, Fact 2 of the present study reproduced a factor characterized by hyperactivity and impulsivity lending further support to study hypotheses. By contrast, none of the three significant Egeland and Kovalik-Gran variables loaded significantly onto the third factor of the present study. This result is not surprising, however, as these three loadings constitute variables dropped from the CPT-3 (i.e., Block Change SE, Block Change, and Change in Commissions). Thus, while the results do not replicate the Egeland and Kovalik Sustained Attention factor as predicted, this pattern of performance instead suggests the existence of a factor characterized by inattention, lending credence to the construct validity of the CPT-3 task.

Generalized Regression

After confirming the construct validity of the CPT-3, next evaluated were which distributions best fit the empirical data for each EFA-identified factor, in order to address the second aim of the investigation. As predicted, all of the data yielded non-linear, right-skewed empirical distributions with varying degrees of skewness. Specifically, the Variability and Hyperactivity/Impulsivity factors were best modeled by a Lognormal response distribution while a Normal-2 Mixture response distribution best captured the empirical distribution of the Inattention factor. Thus, as established at the outset of the investigation, GR analyses were used to effectively and appropriately model group-level CPT-3 data demonstrated to be non-normally distributed. Whereas each EFA-identified factor served as the criterion variable for each GR, ECT, ex-Gaussian parameters, and Miyake tasks served as the predictor variables. Adaptive elastic net penalties were selected to account for collinearity and for a greater number of predictor variables than observations. Notably, as the JMP 15 PRO (SAS, 2019) statistical package does not model Normal-2 Mixture response distributions, a t(5) response distribution was used to model group-level data on the Inattention GR as it represented the next best fit of the empirical data.

Results of the GR models revealed the relative contribution of mental speed and EF cognitive constructs (i.e., ECTs, ex-Gaussian parameters, and Miyake tasks) on each factor. As no precedent literature was available to inform specific hypotheses about which predictor variables would turn up significant, this aim of the investigation was exploratory. Though some general expectations were offered as plausible given our understanding of ECTs, ex-Gaussian parameters, and Miyake factor tasks, results of the analyses yielded nominal support for such expectations. Differences in factor composition for the current study compared to those observed in Egeland and Kovalik-Gran (2008) partially account for this observation.

Given its prominence in the Egeland and Kovalik-Gran (2008) Focus factor, it was plausible for RT variability (i.e., sigma) on ECT tasks to be similarly important for predicting the first factor of the present investigation (i.e., Variability), for example. Instead, results of the Variability GR analysis found 2-bit error (i.e., 1-bit executive error) to be the only significant predictor of the 30 variables entered into the model, accounting for 25% of the observed variance. That is, the number of errors made on an easy executively-mediated task predicted one-fourth of the variance observed on the Variability factor of the CPT-3, underscoring the modulating role of EF on Variability performance.

By contrast, four of the 30 variables entered into the Hyperactivity/Impulsivity GR model were significant when the Hyperactivity/Impulsivity factor was selected as the criterion variable, accounting for just over two-thirds of the variance observed. These variables include: 2-bit mu (i.e., 1-bit EF mu), 3-bit error, (i.e., 2-bit EF error,)1-bit sigma, and Stroop NP-RT reflecting efficient RT responses on an easy executive task, number of errors on a difficult executive task, variability on a speeded task, and performance on an inhibition. Consistent with general expectations, these significant variables suggest that scores characterized by impulsivity and poor inhibition significantly inform performance on the Hyperactivity/Impulsivity factor.

The final GR analysis found 1-bit tau to be the only significant predictor of the 30 variables entered into the model where Inattention was selected as the criterion variable, accounting for 13% of the observed variance. This finding suggests that inefficient RT responses due to grossly inattentive behavior (e.g., loss of mental set) on a speeded task are predictive of Inattention performance. Although neither *a priori* hypotheses nor general expectations were offered for this GR model, that an inattentive ex-Gaussian parameter is demonstrated to be important for predicting results of the Inattention GR is not unexpected.

Overall, results of the GR analyses provide insights into how the group behaves as a whole. And though group-level analyses are thought to ostensibly reflect general trends of the group in aggregate, it must be underscored that what is true for the average respondent is not unequivocally true for *every* respondent. Accordingly, recursive partitioning with decision tree predictive modeling techniques were recruited to address the third and final aim of the investigation which sought to clarify individual differences in performance at the respondent-level, and provide a more nuanced perspective of the CPT-3 construct validity.

Individual-Level Analysis Discussion

While they offer an initial lens through which to evaluate data, group-level analyses can, at times, provide a reductive view of the empirical dataset. For this reason, we argued that such results would be better contextualized alongside outcomes of RP analyses that were expected to better evaluate individual differences in the empirical data. Specifically, as it was expected that subgroups of respondents would differ in their approach to completing the executively-mediated, speeded CPT-3 measure, individual-level RP analyses were expected to capture more nuance in the dataset and were, therefore, predicted to explain more variance than accounted for by group-level GR analyses. Notably, the k-fold R-squared values were used for comparison to GR R-squared values. Furthermore, we hypothesized that the first branch of each DT would replicate the most significant GR predictor variable identified across each CPT-3 factor. Finally, while it was expected that other variables would elicit branching, as no precedent literature was available to inform subsequent partitioning beyond the initial criterion variable, no specific *a priori* predictions were made. Instead, it was expected that results of the current investigation would contribute to the literature in this way.

Hyperactive/Impulsive RP

As predicted, results of the RP analyses were, indeed, largely supportive of study hypotheses such that two of the three RP analyses (i.e., Variability and Inattention) explained greater variance than did their respective GR analysis correlates. Contrary to study predictions, however, the Hyperactivity/Impulsivity RP analysis accounted for *less* variance than its GR analogue. Nonetheless, despite this unexpected finding, this result, in fact, acts as an exemplar for the complementary use of RP alongside group-level analysis given its layered dissection of the Hyperactivity/Impulsivity factor. That is, where RP distinguishes between factor profiles highlighting variability in approach to task completion, GR appears to erroneously collapse these cognitively distinct subtypes into one unidirectional factor, consequently overestimating the variance explained by the four significant variables.

Hyperactivity/Impulsivity DT Right Side

To fully appreciate results of the Hyperactivity/Impulsivity RP analysis, it is important to consider that in its analysis, RP identifies two antagonistic Hyperactivity/Impulsivity subtypes, evidencing support for a factor that: 1) reflects bipolar directionality, and, 2) is further characterized by differing cognitive structures (see Figure 3). The factor is first partitioned by 2-bit mu (i.e., 1-bit EF mu), a relatively easy executively-mediated, speeded RT task, replicating the most significant GR variable predictor, as predicted. The right side of the DT reflects outcomes of slower 2-bit mu performance such that slower ‘efficient’ RT responses are associated with increased Hyperactivity/Impulsivity (i.e., the more hyperactive/impulsive subtype). It is not possible to definitively state, at this time, why slower 'efficient' RT responses are correlated with an increased Hyperactive/Impulsive presentation. To better elucidate this finding, further research is warranted. If called to speculate though, it is possible that respondents exhibiting a Hyperactive/Impulsive profile may respond slower than 'normal' on a relatively easy executively-mediated, speeded RT task because they are unable to recruit the required executive power to attend consistently to the task, ultimately resulting in slower RTs. Still, future investigations that are able to replicate these findings will prove important.

The right side of the DT is partitioned once more by 0-bit sigma where more variable performance on the relatively easy speeded RT task demonstrates *increased* Hyperactivity/Impulsivity while less variable performance reflects *decreased* Hyperactivity/Impulsivity. In contrast to the first branch on this side of the DT, this result reflects findings consistent with what one might intuit—increased variability associated with increased Hyperactivity/Impulsivity.

Hyperactivity/Impulsivity DT Left Side

The left side of the Hyperactivity/Impulsivity DT reflects faster 2-bit mu performance where faster 'efficient' RT responses are associated with the *decreased* Hyperactivity/Impulsivity (i.e., the less Hyperactive/Impulsive subtype). Still, while Hyperactivity/Impulsivity is *most* elevated when 2-bit mu performance is *slower* (i.e., on the right side of the DT), the data reveal gradations in performance even when 2-bit mu performance is faster, as observed on the left side of the DT. This is further delineated below.

Following the initial 2-bit mu branching, the left side of the DT is further partitioned by 1-bit sigma, 1-bit error, 0-bit CoV, and 0-bit mu where decreased variability, fewer errors, and faster ‘efficient’ RT responses on easy mental speed tasks suggest *increased* Hyperactivity/Impulsivity. This result is, is in direct opposition to the findings obtained on the right side of the DT adding further support for a cognitively distinct subtype of Hyperactivity/Impulsivity. Moreover, the inverse is true when the respondent is more variable, makes more errors, and has slower ‘efficient’ RT responses on easy mental speed tasks such that they are less Hyperactive/Impulsive, as defined by this second Hyperactivity/Impulsivity subtype. Again, further research is warranted to better contextualize these findings. What remains clear, however, is the RP analysis method sufficiently demonstrates its utility providing a method by which to systematically examine nuances in the empirical data.

Variability RP

Of the findings generated from the current investigation, perhaps most compelling is the contrast in variance explained by the Variability RP analysis compared to the amount explained by its GR correlate. The Variability GR accounts for 25% of the explained variance and identifies only one variable as important for predicting Variability performance (i.e., 2-bit error). By contrast, the analogous RP analysis accounts for a 41% *more* variance (k-fold $R^2 = 0.66$). Consistent with study predictions, the Variability RP analysis also identifies 2-bit error as the criterion variable that separates the sample into two maximally different groups. Unlike the GR, however, the RP analysis recognizes other variables that contribute to, and ultimately help explain, more variance in the model.

Variability DT Right Side

A relatively easy speeded executive task, 2 bit-error performance reflects greater Variability when errors are higher, as one might expect. Partitioned further by Global-RT, a task that measures one’s ability to effectively shift between tasks, operations, and/or mental sets

(Monsell, 1996), slower performance on the task again reflects greater Variability while faster Global-RT scores are associated with less Variability. Again, these results fit well with our understanding of attention, executive functioning, RT, and Variability.

Variability DT Left Side

In contrast to the right side of the Variability DT, 2-bit error performance reflects decreased Variability when errors are lower, as observed on the left side of the DT. Still, as earlier demonstrated on the H/I factor, RP analysis makes plain gradations in performance. That is, even on the left side of the DT where 2-bit errors are fewer and Variability is less relative to the right side, subsequent branching on the left side of the DT, under a particular set of circumstances, reflects increased Variability. For example, the DT is again partitioned by Global-RT, though this time on the left. As on the right side of the DT, slower Global-RT scores reflect greater variability while fast Global-RT scores are associated with less Variability. The DT is next partitioned by 3-bit error (i.e., 2-bit EF)—a difficult, executively-mediated speeded RT task. When 3-bit errors are greater, Variability is greater while fewer 3-bit errors reflect less Variability. The DT is partitioned by one final variable: Flanker Task RT, a measure of attentional switching. Curiously, this result suggests that one's ability to quickly switch attention between tasks, operations, or mental sets is associated with greater Variability despite another measure of attentional switching (i.e., fast Global RT) being associated with *less* Variability only two splits prior. Under certain conditions, it would therefore seem that quickly switching one's attention between tasks, operations, mental sets or the like can leave one vulnerable to increased Variability. This is, perhaps, due to the respondent learning to anticipate an upcoming shift on the task. Again, results of the RP analysis offer insights about subsets of individuals in the dataset and how their performance may differ from the group as a whole.

Inattention RP

As earlier reported, the Inattention RP analysis accounts for more variance (k-fold R^2 = 0.34) than explained by the group-level GR analysis (R^2 = 0.13), consistent with study hypotheses. However, contrary to study predictions, the Inattention RP analysis does not identify the same criterion variable as generated by the GR analysis (i.e., 1-bit tau). Instead, the RP analysis identifies 0-bit tau as the variable that separates the sample into two maximally different groups. Thus, it is clear that 'loss of mental set' (i.e., inefficient RT responses) as measured by tau, is important for understanding the Inattention factor. Less clear, however, is why the RP analysis appears to place greater importance on the easier 0-bit tau task in comparison to the 1-bit tau task. Perhaps, however, this discrepancy suggests that the Inattention factor is quite sensitive—to the extent that even very simple, elementary attentional processing skills are enough to adequately predict Inattentive performance on the CPT-3 task.

Inattention DT Right and Left Sides

As one might reasonably expect, slower 0-bit tau RT responses on the right side of the DT are associated with increased Inattention while faster 0-bit tau RT responses, as displayed on the left side of the DT, are associated with less Inattention. And while Inattention is most increased on the right side of the DT when 0-bit tau is slower, again, under the specified circumstances, the left side of the DT also reflects increased Inattention.

Following 0-bit tau on the left, the DT is further split by 0-bit mu. When 0-bit mu is slower, Inattention is greater. That is, slower 'efficient' responses on easy speeded RT tasks are associated with increased Inattention. By contrast, faster 0-bit mu is associated with decreased Inattention. The sample is then subsequently partitioned by Global-acc where a slower score is associated with increased Inattention while a faster score is associated with less Inattention—

again, falling within general expectations. Finally, following slower 0-bit mu performance, the sample is partitioned one final time by 3-bit mu where faster ‘efficient’ responses on a complex executively-mediated speeded RT task are associated with increased Inattention. By contrast, the inverse is true of slower ‘efficient’ responses on the same task such that they are associated with less Inattention. Results of this RP analysis demonstrate, one final time, that mental speed and executive functioning constitute the cognitive constructs underlying the CPT-3; moreover, more variance is accounted for in this close examination of individual subsets of performance compared to what can be explained at the group-level of analysis.

Implications, Limitations, and Future Directions

Using the CPT-3 as a model, the present study presents a case for the complementary—and arguably urgent—use of distribution-appropriate statistical methods to analyze non-normally distributed empirical datasets. As reflected in results of the exploratory factor analysis, raw RT scores offer an opportunity for a nuanced, close examination of the empirical data providing insights about patterns of performance that could not otherwise be discerned by imposing a normal distribution on non-normally distributed scores.

Moreover, results of the present study reaffirm that while group-level analysis by any method (e.g., traditional parametric group-level analysis or distribution-appropriate group-level analysis procedures) offers insights into performance of the group in aggregate, it is oftentimes inappropriate to presume that the patterns reflected by the group are, necessarily, applicable to a subset of respondents. Thus, understanding how subgroups within the population navigate and approach a given task can have direct implications for more personalized/individualized assessment and treatment, especially in clinical and research contexts. For example, although results of both the GR and RP analyses demonstrate support for mental speed and EF tasks as cognitive underpinnings of the CPT-3, RP analyses more frequently account for greater explained variance while also identifying under what set of conditions each significant variable proves salient on each factor.

Although results of the investigation are largely supportive of study hypotheses, the present study is not without limitation. Undoubtedly impacted by its smaller sample size, this limitation is perhaps most apparent. To confirm significant findings, it is warranted to invest in a larger dataset for future studies to ensure replicability. Still, while a larger sample size would certainly reify confidence in the observed results and help to clarify power-related queries, several steps to mitigate any spurious findings that could result in ambiguous conclusions have been taken. For example, the present study used an Adaptive Elastic Net generalized regression procedure given the unbalanced predictor to observation ratio. Moreover, as predictive modeling techniques can overcapitalize on chance variance, the present study used a k-fold cross-validation resampling procedure to account for this. Future studies may consider running additional cross-validation procedures such as boosted tree analysis to ensure a convergence of evidence is achieved.

What's more, although the parameterization of RT tasks provides a systematic approach to classifying *bit* tasks by difficulty, the integrated use of *bit* tasks within the neuropsychological and psychological literature, at this time, remains limited. For some, this fact may raise several questions. However, while there is little extant precedent literature upon which to reference within the general psychological literature base, *bit* theory is much more prevalent and readily integrated into the Information Theory literature base. Thus, in order to quell any uncertainty on behalf of the reader, it is important to adequately and sufficiently convey the theoretical rationale and utility of the *bit* system, as we have aimed to do in the present investigation.

Pursuant to findings of the present study, there is a clear case for the use distribution-appropriate statistical methods in the context of non-normally-distributed empirical data. Accordingly, as the field of neuropsychology continues to move toward more frequent use of computerized tasks, evaluating RT effectively will, undoubtedly, become more critical. It would, therefore, prove prudent for the field to adopt a distribution-appropriate analysis framework moving forward.

References

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, *20*(3), 160-166.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies* (No. 1). Cambridge University Press.

Cona, G., Arcara, G., Amodio, P., Schiff, S., & Bisiacchi, P. S. (2013). Does executive control really play a crucial role in explaining age-related cognitive and neural differences? *Neuropsychology*, *27*(3), 378–389.

Conners, C. K. (2002). *Conners continuous performance test (CPTII). Technical guide and software manual.* North Tonawanda, NY: Multi Health Systems.

Conners, C. K. (2014). *Conners continuous performance test 3rd edition manual*. Toronto: Multi-Health Systems.

Egeland, J., & Kovalik-Gran, I. (2008). Measuring several aspects of attention in one test: The factor structure of Conners' Continuous Performance Test. *Journal of Attention Disorders*, *13*(4), 339-346.

Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non-search task. *Perception and Psychophysics, 16*, 143-149.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179.

Golden, C. J. (1978). *Stroop color and word test*. Wood Dale, IL: Stoelting Co.

Halperin, J. M., Sharma, V., Greenblatt, E., & Schwartz, S. T. (1991). Assessment of the Continuous Performance Test: Reliability and validity in a nonreferred sample. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*(4), 603.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of experimental psychology*, *4*(1), 11-26.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences.* Oxford, UK: Elsevier.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, *55*(4), 352.

Leclaire, K.N., Osmon, D.C., Driscoll, I. (2020). A distributional and theoretical analysis of reaction time in the reversal task across adulthood. *Journal of Clinical and Experimental Neuropsychology*, *42*(2), 199-207. Doi:1080/13803395.2019.1703909

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *49*A(1), 29–50.

Lerner, M. D., & Lonigan, C. J. (2014). Executive function among preschool children: Unitary versus distinct abilities. *Journal of Psychopathology and Behavioral Assessment*, *36*(4), 626–639.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization* (No. 8). Oxford University Press on Demand.

Monsell, S. (1996). Control of mental processes. *Unsolved mysteries of the mind: Tutorial essays* *in cognition*, 93-148.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353-383. doi:10.1016/0010-0285(77)90012-3

McGill, W.J. Stochastic latency mechanisms. In R.D. Luce, R.R. Nush, & E. Galanter (eds.). *Handbook of Mathematical Psychology,* Vol. I. New York: Wiley, 1963, pp. 309-360.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, *13*(3), 395–423.

Mirsky, A. F., Pascualvaca, D. M., Duncan, C. C., & French, L. M. (1999). A model of attention and its relation to ADHD. *Mental Retardation and developmental disabilities research reviews*, *5*(3), 169-176.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current directions in psychological science*, *21*(1), 8-14.

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.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology,* **81,** 111–121.

Neumann, O. (1996). Theories of Attention. In O. Neumann & A. F. Sanders (Eds). Handbook of

Perception and Action, Vol. 3, New York: Elsevier.

O'Connor, T. A., & Burns, N. R. (2003). Inspection time and general speed of processing. *Personality and individual differences*, *35*(3), 713-724.

Osmon, D.C., Kazakov, D., Santos, O. *et al.* Non-Gaussian Distributional Analyses of Reaction Times (RT): Improvements that Increase Efficacy of RT Tasks for Describing Cognitive Processes. *Neuropsychol Rev* 28**,** 359–376 (2018). https://doi.org/10.1007/s11065-018-9382-8

Osmon, D. C., Leclaire, K. N., Driscoll, I., & Zolliecoffer, C. J. (2020). Reversal learning in young and middle-age neurotypicals: Individual difference reaction time considerations. *Journal of Clinical and Experimental Neuropsychology*, *42*(9), 902-913.

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

Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology*, *31*(3), 206-230.

Ratcliff, R. (1979). A theory of memory retrieval. *Psychological Review*, 86, 446–461. Doi: 10.1037/0033-295X.85.2.59.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. Psychological Bulletin, 86, 446–461.

Ratcliff, R., & Murdock, B.B., Jr. (1976). Retrieval processes in recognition memory.

Psychological Review, 83, 190–214.

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.

Roberts, R. D., & Pallier, G. (2001). Individual differences in performance on elementary cognitive tasks (ECTs): Lawful vs. problematic parameters. *The Journal of general psychology*, *128*(3), 279-314.

Rosvold, H., Mirsky, A., Sarason, I., Bransome, E. D. Jr., & Beck, L. H. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343–350.

SAS (2021). *Generalized regression models*. Cary NC: SAS Institute, Inc.

SAS (2021). *Overview of the bootstrap forest platform*. Cary NC: SAS Institute, Inc.

SAS (2021). *Overview of the partition platform*. Cary NC: SAS Institute, Inc.

SAS (2021). *Overview of the boosted tree platform*. Cary NC: SAS Institute, Inc.

Schmitz, F., & Wilhelm, O. (2016). Modeling mental speed: Decomposing response time distributions in elementary cognitive tasks and correlations with working memory capacity and fluid intelligence. *Journal of Intelligence*, *4*(4), 13.

Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, *37*(1), 10-21.

Song, Y. Y., & Ying, L. U. (2015). Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, *27*(2), 130.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,* 18, 643–662.

Suchy, Y. (2015). *Executive functioning: A comprehensive guide for clinical practice*. Oxford University Press.

Tsuchida, A., & Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *49*(7), 1790–1800.

Van der Meere, J.J. (2002). The role of attention. In Sandberg (Ed.), *Hyperactivity and attention disorders of childhood.* Cambridge, UK: Cambridge University Press.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in cognitive sciences*, *12*(11), 418-424.

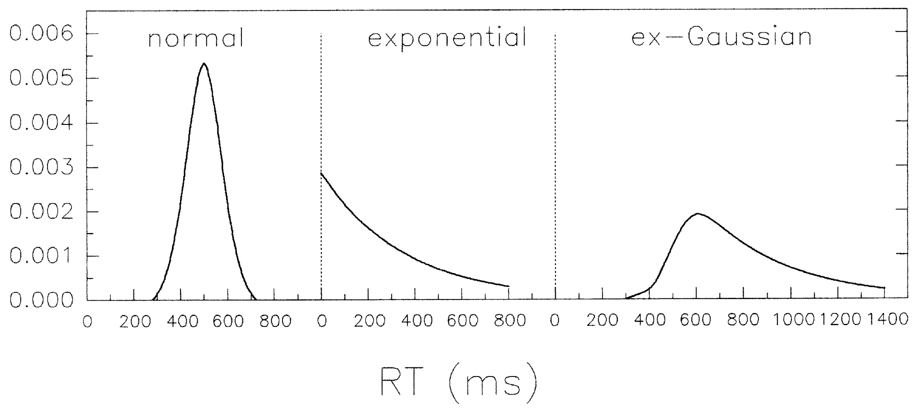
Whitney, P., Arnett, P. A., Driver, A., & Budd, D. (2001). Measuring central executive functioning: What's in a reading span? *Brain and Cognition*, *45*(1), 1–14.

Yildirim, S. (2020, April 25). *Random forests vs. gradient boosted decision trees.* Artificial intelligence. https://ai.plainenglish.io/random-forests-vs-gradient-boosted-decision-trees-dd4f0ef86554

Yntema, D. B. (1963). Keeping track of several things at once. *Human factors*, *5*(1), 7-17.

APPENDIX

Distributions Explained



Appendix. The ex-Gaussian distribution is composed of a normal distribution and an exponential component. The normal component has mu (central tendency) and sigma (deviation) parameters while the exponential has a tau (mean and deviation combined) parameter.