Title
Modeling cognitive control: Using cross-classified IRT and structural equation modeling to understand brain and behavior predictors of academic achievement

Author
Torgrimson, Sarah Jo

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Modeling cognitive control: Using cross-classified IRT and structural equation modeling to understand brain and behavior predictors of academic achievement

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Education

by

Sarah Jo Torgrimson

2022

ABSTRACT OF THE DISSERTATION

Modeling cognitive control: Using cross-classified IRT and structural equation modeling to understand brain and behavior predictors of academic achievement

by

Sarah Jo Torgrimson

Doctor of Philosophy in Education

University of California, Los Angeles, 2022

Professor Jennie Katherine Grammer, Chair

Cognitive control skills are foundational to goal-directed action and related to children's academic achievement. These skills are often quantified using observed measures of children's brain (e.g., ERN, CRN, Pe) and behavior (e.g., accuracy, reaction time) on computerized tasks. While computerized tasks afford precision in measurement, observed measures of behavior do not allow for examination of differences in performance between trials that may result from varying task features (e.g., trial difficulty, trial randomization) or children's interactions with the task (e.g., previous trial performance impacting next trial performance). In this investigation, I leverage item response theory (IRT) models to identify task features that impact children's behavioral performance as well as accommodate for inter-trial variation in Go/No-Go performance among children. Results indicate that task features, such as trial difficulty, impact children's likelihood of making an error and also impact reaction time performance. Further, studies utilizing tasks like the Go/No-Go typically report linear regressions between an academic outcome and a single score of behavioral performance. However, relations between individual scores and real-world skills, including academics, are relatively weak and inconsistent across studies. In this project, I compare three approaches for modeling relations between cognitive control and academic skills – a traditional regression approach, an observed measure SEM model, and an adjusted (predicted) measure SEM model. Nine behavioral and four ERP measures of Go/No-Go task performance are assessed as indicator variables for a cognitive control construct. Findings show that SEM models that incorporate multiple measures of cognitive control are superior for predicting variance in academic outcome measures as compared to regression approaches. Overall, this study emphasizes that multiple measures from one computerized assessment of cognitive control can be leveraged to predict moderate variance in academic skills for children in early elementary school.

The dissertation of Sarah Jo Torgrimson is approved.

Elif Isbell

Minjeong Jeon

Catherine Lord Morrison

Jennie Katherine Grammer, Committee Chair

University of California, Los Angeles

2022

DEDICATION

This work is dedicated to my grandmother, Ruth Tarullo, whose support and encouragement has meant so much for my development both as a person and a scholar.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

There are many folks without whom my success in graduate school would not be possible. I would like to first thank my undergraduate thesis mentor, Jennifer Martin McDermott, for providing me with the foundational building blocks and emotional support that have made me the researcher I am today. To my doctoral advisor, Jennie Grammer – your mentorship has helped mold and shape my conceptual thinking, improve my skills, and guide my career path. I greatly appreciate how much you have advocated for me over the past six years. Thank you for being a trusted advisor and friend. I would also like to thank my dissertation committee members - Elif Isbell, Minjeong Jeon, and Cathy Lord - who offered diverse perspectives and expertise which greatly improved the quality of the finished dissertation. Thank you also to Andy Lin who was instrumental in helping me navigate conversion issues and providing me with standardized language around reporting model building decisions. To the students, schools, and families that participated in the original project from which this data originates, I am grateful. Further, I want to thank the researchers that conducted the original study for providing me the opportunity to complete a secondary project of this scope.

To my larger academic community, thank you for all your friendship, comradery, mentorship, and sage wisdom. Some faculty and staff I'd like to particularly thank include Amy Gershon, Elaine Dolalas, Bonnie Goff, Jason Dorio, Anne Blackstock-Bernstein, Rashmita Mistry, and Sandra Graham. Members of the BEAR lab and beyond – thank you – particularly Amira Ibrahim, Caty Solone, Remi Torres, Briana Rodriguez, and Sonia Jamani. And to Keye Xu and Sal Vazquez, thank you for the writing groups, group chats, teaching and research help, and vital peer feedback! You two have been such an important support system for my time here at UCLA. Thanks to Chaia Flegenheimer and Adaeze Egwuatu for maintaining our friendship across coasts and for all the virtual writing groups, support sessions, and job market help! And thanks to my students and mentees at UCLA and UMass who made my experiences in academia exponentially better.

There are several other friends and colleagues I'd like to acknowledge; those that have helped LA feel like home and those who supported me remotely from various far-off places. Jenna Wolter and Megan Lebre, I'm so grateful UCLA connected us. Thanks also to Ginny Sklar, Amanda Johnson, Lauren Baczewski, Alejandra Martin, Billy Cho, Freddy Novoa, Lupe Lopez, Taylor Hazelbaker, Lauren Chiang, Andy Schlink, Ramon Flores, and Natalie Aguilar for the years of HDP friendship. To my HEOC folks – the best of us! – Elaine Tamargo, Katherine Cho, and especially Christina Zavala – thank you for providing perspective and serving as pretty superb peer role models. Many thanks to Sarah Rotella, Laura Fisk, Oren Rodriguez, and Emily Mankin for being members of my faith community in LA. Thank you to Lauren Kinnard and Rebecca Smith who supported me through the last wind sprints of the dissertation process! To long-time friends that provided long-distance support and patiently listened to me ramble about the ups and downs of graduate school – thank you! – Anna Stone, Adrienne Viarengo, Lexi Abbott, Lisa Pechilis, Justin Petraitis, Brian Strange, Chris DiTullio, Mariah Lancaster, Dan Shimkus, and Kayla Storey. To my extended family, both born and chosen, thank you for your words of encouragement over the years, particular thanks to Beth Allaire and Darvin and Marilyn Torgrimson. And to the three people who always want to hear from me good or bad – Mum, Dad, Becca – I love you. Thank you for listening to all things relevant and irrelevant, for nodding your head even when things didn't make sense, and for reminding me of my purpose.

VITA

EDUCATION

M.A. in Education | Human Development & Psychology 2018
University of California Los Angeles

B.S. in Psychology | Neuroscience Track 2014
University of Massachusetts Amherst, *cum laude*

SELECTED GRANTS, AWARDS, & FELLOWSHIPS

Incorporating brain and behavior: 2018-19
Developing measures of math persistence in middle schoolers
Graduate Research Mentorship | University of California, Los Angeles

Incorporating brain and behavior: 2018
Exploring factors related to student success
Graduate Research Mentorship | University of California, Los Angeles

Examining persistence, engagement, attention, and 2016-18
Regulation in elementary school students
CONNECT Research Fellow | University of California, Los Angeles

SELECTED PUBLICATIONS

Xu, K., Torgrimson, S.J., Torres, R., Lenartowicz, A., & Grammer, J.K. (2022). EEG data quality in real-world settings: Examining neural correlates of attention in school-aged children. *Mind, Brain, & Education.* doi:10.1111/mbe.12314

Torgrimson, S.J., Tan, P.Z., & Grammer, J.K. (2021). Associations among response inhibition, motivational beliefs, and task persistence in early elementary school. *Journal of Experimental Child Psychology, 208.* doi:10.1016/j.jecp.2021.105141

SELECTED PRESENTATIONS

Torgrimson, S.J., Jeon, M., Isbell, E., & Grammer, J.K. (2022). IRT Modelling of Go/No-Go Task Performance: Considerations for Computerized Cognitive Games. Poster presented to the biannual meeting of the *International Mind Brain and Education Society, Montreal*, Quebec, Canada.

Torgrimson, S.J., Tan, P., & Grammer, J.K. (2019). *Gender Differences in Cognitive Control and Student Persistence: A Neurophysiological Investigation.* In S.J. Torgrimson & P. Tan (Chairs). Gender Differences in Psychophysiological Measures of Emotion Reactivity and Self-Regulated Behaviors. Paper presented at the biannual meeting of the Society for Research on Child Development. Baltimore, MD.

Torgrimson, S.J. & Grammer, J.K. (2018). *Task persistence: Self-regulatory and motivational factors in early elementary school.* In F. Morrison, (Chair). Executive Function Development in Young Children: Bridging Psychological, Neurological, and Educational Perspectives. Paper presented at the annual meeting of the International Mind, Brain, and Education Society Conference. Los Angeles, CA.

Torgrimson, S.J., Trane, F., & Grammer, J.K. (2018). None-the-less she persisted: Gender differences in executive function and persistence in early elementary school. Poster presented at the *Association for Psychological Science Conference*, San Francisco, CA.

Torgrimson, S.J., Trane, F., & Grammer, J.K. (2018). Gender & Cognition: Task persistence in Elementary-Aged Students. Poster presented at the *UCLA Research and Inquiry Conference*, Los Angeles, CA.

Torgrimson, S.J., Grammer, J.K, Kim, M., Morrison, F.J., & Gehring, W.J. (2017). Executive functions and academic success: Neural indices of response inhibition relate to children's literacy and math achievement. Poster presented at the *Society for Research in Child Development Conference*, Austin, TX.

Torgrimson, S.J., Lattes, E., & McDermott, J.M. (2017). Stereotype threat and executive function in girls: Gender priming impedes performance on the Tower of Hanoi. Poster presented at the *Society for Research in Child Development Conference*, Austin, TX.

SERVICE

Masters Capstone Thesis Mentor 2019-21
Graduate Student Mentor | Department of Education, UCLA

Research in Psychology Mentor 2016-21
Graduate Student Mentor | Department of Psychology, UCLA

Student Co-Chair 2019-20
Department of Education, UCLA

Committee Student Representative 2019-20
Education Executive, Teaching & Learning | Department of Education, UCLA

Student Representative 2018-19
Human Development & Psychology Division, UCLA

Modeling cognitive control: Using cross-classified IRT and structural equation modeling to understand brain and behavior predictors of academic achievement

Cognitive control skills of attention, inhibition, and response monitoring help children set, plan, and achieve task-related goals and appropriately regulate their behavior to meet environmental demands (Friedman & Miyake, 2017). These skills have been linked to important outcomes for children and youth, including academic achievement (Bierman et al., 2008; Fishbein et al., 2016; Kim et al., 2016) and diagnosis of neurodevelopmental disorders such as attention-deficit hyperactivity disorder (ADHD; Huang-Pollock & Karalunas, 2010; Musser & Raiker, 2019; Senderecka et al., 2012), autism (ASD; Delorme et al., 2007), and anxiety (Wauthia & Rossignol, 2016).

Computerized cognitive control tasks are widely used to assess children's cognitive functioning, affording researchers tight experimental control and the ability to narrow in on specific subprocesses of cognition. From these tasks, researchers can extract behavioral and neural measures of performance, perceived as more "objective" indices of cognitive control. However, relations between cognitive control and academic or clinical outcomes are mixed in the literature, in part due to inconsistencies in measurement. Researchers often estimate performance on computerized cognitive control tasks as a single score of children's cognitive functioning, though there is no consensus on which measure(s) of behavior (e.g., accuracy, reaction time) are most useful in determining predictors of academic achievement. Additionally, there is evidence to suggest that behavioral performance is impacted by task features (e.g., trial type) and child demographics (e.g., age), yet few have investigated how these factors impact children's within-task behavior or the observed average performance measures that are traditionally used to index cognitive control. Moreover, relations among concurrent brain and behavior measures derived from the same task are not fully understood. Further, there is debate over whether neural measures provide novel information, above and beyond the behavioral measures, that is helpful for predicting academic and psychological outcomes. Leveraging advanced modeling techniques to examine behavioral and neural data from one standardized cognitive control task, the goals of this current investigation are to: 1) identify task features that impact behavioral performance by investigating child, trial, and child-by-trial factors of accuracy and reaction time 2) assess how measures of task behavior relate to children's reading and math skills using structural equation modeling to estimate cognitive control as a latent construct, and 3) determine the value-added of incorporating neural measures into estimates of cognitive control when examining relations with academic skills.

Literature Review

Cognitive Control

Cognitive control includes the ability to inhibit inappropriate responses, hold information in working memory, and control attention - allowing children to adjust their behavior to environmental conditions, engage with novel information, and shift between different tasks (Miyake & Friedman, 2012). There is a large overlap between conceptualizations of cognitive control and executive function (EF; (McClelland et al., 2015); and, in practice, the two constructs are often assessed simultaneously in studies of children's cognition using a combination of EF measures (e.g., caregiver reports, observations) and computerized tasks tapping cognitive control skills. Studies of cognitive control and EF indicate that basic cognitive processes (e.g., attention orienting) emerge early in life, followed by the rapid development of higher order functioning (e.g., inhibition) that occurs between the ages of 3 and 7 years, alongside maturation in the prefrontal and anterior cingulate cortices (Crone & Steinbeis, 2017). Indeed, age-related differences in cognitive ability have been identified in both cross-sectional and longitudinal studies (Best et al., 2011; Grammer et al., 2014).

Cognitive control skills are thought to be particularly important in elementary school, when children are acquiring foundational academic skills (de Haan, 2014). By preschool-age (3-5-years-old), children's cognitive control capacity is already predictive of children's later academic success and social and emotional well-being (Blair & Raver, 2015; Bull et al., 2008; Downes et al., 2017; Espy et al., 2004; St Clair-Thompson & Gathercole, 2006). Accordingly, cognitive control skills are considered key in successful classroom adjustment helping children focus and sustain their attention, monitor their behavior and understanding, and remember directives and instructions from their teachers (Blair & Razza, 2007; McClelland et al., 2007).

Response inhibition, in particular, appears to be an important skill during this period of development and is most associated with children's academic performance in elementary school (Grammer et al., 2014; McClelland et al., 2015; Montroy et al., 2016; Roebers et al., 2011). Notably, there is some evidence of gender differences in response inhibition in elementary school, such that boys, on average, demonstrate lower response inhibition skills and greater variance in performance than girls (Matthews et al., 2009). However, inhibition skills appear to be related to academics for both boys and girls at this age.

Defined as the ability to inhibit a pre-potent response in favor of a desired, but not automatic or instructed response, response inhibition involves the coordination of many attention and response monitoring processes (Garon et al., 2008; Verbruggen & Logan, 2008). Response monitoring serves to allow the brain to piece out which incoming information is relevant to the task at hand and which information can be disregarded. These reflective processes are considered to be a form of active working memory involved in the encoding and updating of information

(Miyake et al., 2000). This skill has been linked to the ability to identify success or failure during task performance and is highly valuable during novel task learning as it allows individuals to modify their own behavior in order to improve their performance (Friedman & Miyake, 2017). While overt processing can be observed in the presence of compliant behavior (e.g., correctly responding to a prompt), response monitoring is considered a covert process, and cannot be directly observed (e.g., error processing, withholding a response). Therefore, researchers utilize methods that allow for non-invasive measurement of neural activity, quantifying event-related potentials (ERPs) that reflect covert cognitive processing such as stimulus encoding, inhibition, and response monitoring. Assessing both behavioral and neural measures of cognitive control is useful for understanding skill capacities in children and is encouraged by the National Institute of Mental Health as a best practice in research, as outlined in RDoc framework (P. Anderson, 2002; Bekker et al., 2004; Cuthbert & Insel, 2013).

Computerized Assessments of Cognitive Control

The study of cognitive control skills in relation to academic achievement dates back as early as the 1960s when public funding for educational services increased in response to Lyndon B. Johnson's 1964 "War on Poverty" (Zhai et al., 2011). In early studies, these skills were categorized as related to student self-control and often sampled children participating in compensatory education programs (Flynn, 1975). Educational studies of children's cognitive and regulatory skills continued throughout the 80s and 90s as motivational theorists incorporated contextual factors into the study of children's attention and regulatory abilities (Grolnick & Ryan, 1989; Ryan & Deci, 2000). Computerized assessments of cognitive control became popular in the 1990s, or "decade of the brain" due to the ability to concurrently measure behavioral performance indicators of cognitive control and associated neural activity (Blair, 2002).

Computerized tasks of cognitive control were designed to tap into the functions of separable, but related, componential skills including attention shifting, inhibition, and working memory (Friedman & Miyake, 2017; Miyake et al., 2000; Miyake & Friedman, 2012). These tasks were initially designed based on adult's neurological functioning. Although they have begun to be used widely with developmental populations, the relation between children's behavioral performance and neural activity during these tasks is yet to be fully understood (Hunter & Sparrow, 2012; Miller et al., 2013; Wiebe et al., 2011).

Tasks can differ on several parameters, such as stimulus complexity (e.g. image, word, sound), response required (e.g. left/right, press/don't press), and number of trials. Each of these factors impact performance. In addition, trial type and preceding context have been shown to influence task accuracy and reaction time for adult populations (Schulz et al., 2009; Wild-Wall et al., 2009). For example, in a Go/No-Go paradigm, participants must regulate their response based on trial type, pressing a button on 'Go' trials and inhibiting the prepotent response to press a button on the 'No-Go' trials. In a Go/No-Go task, preceding context of the trial can differ by the number of Go trials, increasing the difficulty of demonstrating response inhibition on a No-Go trial after consecutive Go responses. Greater number of Go trials preceding a No-Go trial has been linked to lower accuracy (Schulz et al., 2009). Additionally, the presentation of trials is typically randomized. This is done to reduce the likelihood that participants will become attenuated to the task and to increase the likelihood of error, in turn, allowing for measurement of children's performance monitoring and behavioral accommodations in response to erring. As a result, even though basic task parameters may be the same (e.g., the number of trials, percentage of No-Go trials), the order with which trials are presented varies for each participant. Thus, some individuals may receive a string of more difficult trials early in the task, which may affect their later performance due to changes in participant arousal or cognitive fatigue (Kato et al., 2009; Schulz et al., 2009). However, how task features impact children's performance remains unknown.

Behavioral Measures of Cognitive Control

There are several behavioral performance measures that can be extracted from a Go/No-Go task that are thought to index similar but distinguishable cognitive skills (Wright et al., 2014). Typically, accuracy and reaction time measures are quantified as averages in performance across task. Most commonly used measures include errors of omission (EOM), errors of commission (ECOM), and Go trial reaction time (RTC). Errors of omission is calculated as the percentage of incorrect Go-trials (no response) and is considered a measure of sustained attention and attention shifting. Errors of commission is considered a measure of response inhibition and calculated as percent incorrect No-Go trials (response). Reaction time is reported as an average of response times to Go trials. Go trials occur more frequently than No-Go trials in Go/No-Go paradigms. Thus, reaction time to correct Go trials is thought to index processing speed and vigilance encompassing the time needed for stimulus encoding, decision making of whether to respond or not, and execution of a motor response (Dutilh et al., 2019). Previous work indicates that Go/No-Go accuracy increases, while reaction times decreases across childhood and into early adulthood - indicating improved inhibition and attention skills and faster processing speeds (Motes et al., 2018). Reaction time can also be measured for errors of commission (incorrectly responding to a No-Go stimulus; RTI), though less focused on in the literature. A critique of reaction time measures is that reaction time may vary between trials, with children responding sometimes faster to Go trials than to No-Go trials and sometimes responding slower to Go trials than to No-Go trials. This is likely dependent on the context of the previous trial (Ratcliff & Tuerlinckx, 2002). As such, variable measures such as standard deviation in reaction time (SDRT) can be helpful to include in models of cognitive control to accommodate for trial-to-trial variability in reaction time that may be masked in traditional mean reaction time measures (Weigard et al., 2020).

Notably, in adult populations, a speed-accuracy tradeoff in performance is often present, wherein, participants prioritize accuracy on No-Go trials over speed. Speed-accuracy tradeoffs can be observed by better No-Go trial accuracy, slower average reaction times, and lower Go trial accuracy. Participants may decrease their reaction time across trials to avoid committing an error of commission and, consequently, demonstrate lower hit rates for Go trials that require quick response (Votruba & Langenecker, 2013). Because of potential speed-accuracy tradeoffs, it is suggested that reaction time measures not be used in isolation (Draheim et al., 2019). Yet, many studies still use singular reaction time measures to index cognitive control performance. Speed-accuracy tradeoffs have also been shown to have a developmental pattern, such that younger children (7-8 year-olds) have displayed significantly higher rates of trading slower reaction times for better No-Go trial accuracy than young adults (18-25 year-olds; (Motes et al., 2018). However, it is unclear when during development this transition occurs. In Motes' (et al., 2018) cross-sectional study, children ages 10-15 years-old did not demonstrate significant difference in speed-accuracy tradeoffs with either younger children or adults. Additionally, arguments have been made that speed-accuracy tradeoffs rely on the false assumption that all participants are engaging in similar task strategies (Draheim et al., 2019).

As an alternative, multiple measures of accuracy and reaction time can be integrated to better understand performance and task strategy. For example, a discriminability measure (d-prime) incorporating accuracy rates on Go and No-Go trials can be calculated in order to account for response biases in performance. D-prime is calculated by subtracting a child's z-score on inaccurate No-Go trials from their z-score on accurate Go trials (Tottenham et al., 2011). Additionally, post-error adjustments in reaction time and accuracy can be calculated to gain a better understanding of participant task strategy. Post-error slowing (PES) and post-error accuracy (PEA) are calculated as differences in average performance on Go trials following error response (No-Go trials) and average performance on Go trials following correct response (Go trials). Post-error slowing is an average decrease in reaction time following error responses as opposed to correct responses. There are competing theories on PES suggesting different causes for this phenomenon. Participants may actively engage cognitive control skills to slow their reaction time and increase accuracy in response to an error (Gehring & Fencsik, 2001; MacDonald et al., 2000). Alternatively, participants may orient too much attention to a committed error, and thus demonstrate slower response times and an increased likelihood of committing another error (Notebaert et al., 2009; Steinborn et al., 2012). It is also speculated that an error of commission results in greater engagement of motor inhibition skills, which results in slower reaction times (King et al., 2010; Ridderinkhof, 2002). Though less commonly reported, post-error changes in accuracy can also be assessed following the commission of an error (Danielmeier & Ullsperger, 2011; Jentzsch & Dudschig, 2009; Maier et al., 2011). Post-error changes in reaction time are not always accompanied by post-error changes in accuracy, further suggesting that individuals may be engaging in different task strategies during performance.

Therefore, integrating multiple measures of accuracy and reaction time may help us better understand dynamics in Go/No-Go task performance.

Although parsimonious, observed measures described above do not allow for examination of differences in performance between trials that may result from varying task features (e.g., trial difficulty, trial randomization) or children's interactions with the task (e.g., previous trial performance impacting next trial performance). Latent models, like item response theory (IRT), can be used to accommodate for different sources of inter-trial variation. Yet, little research has investigated cognitive control performance using these methods. However, studies using factor analyses and structural equation modeling (SEM) have demonstrated that estimating latent constructs of cognitive control using multiple measures of performance is important.

Latent Models of Cognitive Control. As previously outlined, cognitive control tasks are often designed to target a specific subskill (e.g., attention shifting, inhibition). However, task engagement requires the coordination of several skills such as attention, inhibition, and working memory. Indeed, latent modeling of cognitive control has demonstrated that multiple skills support task performance. For example, Howard and colleagues (2014) found that latent factors of attention and inhibition supported performance across a battery of inhibition tasks in samples of children (ages 7-12) and adults (ages 17-51). Nine inhibition tasks were completed – an antisaccade task, a Stroop, a stop-signal, the Hayling task, a directed forgetting task, a proactive interference task, a flanker task, a negative priming task, and a retrieval-induced forgetting task. However, only two tasks loaded onto the inhibition construct in children, compared to seven in adults. This may in part be due to how performance was measured. Performance on each task was quantified as only one score. While some task measures may seem more related to inhibition (e.g., reaction time for inhibiting a response on the stop-signal), others seem more related to working memory (e.g., proportion of total recall from proactive interference task) though task performance may involve inhibitory skills. These measures of performance may better capture inhibition in adults, whereas other measures of performance may be more strongly related to inhibition in children. Investigating multiple measures from one task may better uncover variance in inhibition skills among children. Notably, the stop-signal task, most related to a Go/No-Go, loaded onto both constructs of attention and inhibition in both the child and adult samples.

SEM methodology has also been used to demonstrate relations between latent factors of cognitive control and academic skills in school-age children. In a longitudinal investigation, Albert and colleagues (2020) found that latent cognitive control skills mediated the relation between socioeconomic status (SES) and academic achievement in a racially and socioeconomically diverse sample in the Southeastern United States. Children's performance on a battery of computerized cognitive control and EF tasks at age 10 was related to their performance on a state exam assessing reading and math skills at age 13. Children completed tasks assessing verbal and spatial working memory, response inhibition (Stroop), and strategic planning (Tower of London). Notably, only one factor of cognitive control was used in the model and results indicated that it was most related to verbal working memory (46.19% variation explained) and least to response inhibition (5.76%). This corroborates earlier work indicating that two factors of cognitive control should be modeled when considering performance across a battery of cognitive control and EF tasks (De Franchis et al., 2017; Howard et al., 2014).

De Franchis and colleagues (2017) found a two-factor structure for cognitive control representing latent constructs of working memory/attention shifting and inhibition among 5-year-old Italian children. Cognitive control skills of inhibition (Circles Drawing Task, Tower of

London), working memory (Backward Digit Span, Dual Request Selective Task), and attention shifting (Semantic Fluency, Dimensional Card Sort) were assessed. The working memory/attention-shifting factor predicted reading comprehension, but not spelling skills, in grade 3. Inhibition was not uniquely related to academic achievement, though, it was strongly correlated with the working memory/attention shifting factor. Further, observed scores were constrained to load onto one-factor, though performance on each task may be supported by multiple constructs of cognitive control.

Similarly, Gashaj and colleagues (Gashaj et al., 2019) found a two-factor model of cognitive control among Swiss children. Inhibition (flanker task), attention shifting (fish flanker), and working memory (backward color span task) were assessed in this study, using one measure from each computerized task to represent performance. Working memory and inhibition/attention shifting skills at age 5 years were related to math achievement in grade 2. Notably, the attention shifting and inhibition tasks were both flanker tasks that used reaction time as the sole outcome measure. Whereas, behavioral performance on the working memory task was assessed with a task accuracy measure. The two-factor structure of their model may therefore, represent measurement differences rather than construct differences.

While this work highlights the importance of latent modeling during investigations of response inhibition, the examples outlined above examined cognitive control as contributing to average performance across a battery of tasks. Potential limitations to this analytical strategy include inability to 1) examine random effects of nuances in task presentation influencing children's performance (i.e., trial randomization within subject) and 2) investigate how factors of trial difficulty within task differ as a function of child latent abilities. Alternatively, multilevel latent modeling can be applied to investigate differences in performance related to child, trial, and child-by-trial characteristics. Estimating adjusted performance scores using modeling techniques like IRT can help minimize measurement error due to extraneous factors (e.g., attention) and account for trial randomization across different participants. Adjusted performance scores of Go/No-Go accuracy and reaction time may better capture individual measures of cognitive control. Additionally, only one measure from each task in the cognitive control batteries was used to represent behavioral performance, when a variety of accuracy and reaction time measures could have been extrapolated. Modeling performance using multiple scores from the same task would allow for the investigation of relations between behavioral measures for trials with varying cognitive demands.

Behavioral measures have been used more frequently in the literature predicting academic skills, with less work linking ERP measures of cognitive control to academics. Neural measures capture aspects of covert cognitive processing that are not easily observable with behavioral measures (e.g., response monitoring). The ERN, a neural correlate of response monitoring has been related to children's behavioral performance, linking cognitive processing in the brain to observed behaviors (Grammer et al., 2014; Torpey et al., 2012; Wiersema et al., 2007). However, behavioral and neural measures are typically used separately to demonstrate differences in various desired outcomes such as academic performance, trait anxiety, ADHD, or ASD (Albrecht et al., 2008; Hanna et al., 2020; Kim et al., 2016; Musser & Raiker, 2019; Wauthia & Rossignol, 2016; Zhang et al., 2009). Yet, little research has integrated measures of both brain and behavior into models predicting academic performance.

Neural Measures of Cognitive Control

Neurons, or cells, in our brain communicate through electrochemical signals that produce a measurable electrical current. Continuous electroencephalogram (EEG) activity is collected, similar to the way in which the electrical activity of the heart is measured (electrocardiograms; EKG) by placing electrodes, or sensors, on the scalp, often through usage of an EEG cap. A unique benefit of this methodology is that it is temporally sensitive and can measure immediate brain responses to events, on a millisecond time scale. In order to get a neural measure of a child's response monitoring ability, experimenters average neural reactivity following a child's response producing an average neural waveform (e.g., average neural activity following error responses) (Pontifex et al., 2010). ERP components of interest are then quantified as an average amplitude of neural activity within a specific time window. Time windows are meant to capture peaks in the data, seen as negative or positive deflections in the average waveform, and represent brain activity in response to an event onset (e.g., correct or error response). This method requires that neural signals are phase-locked in order for component "peaks" to be detected, as waveform amplitudes reflect the summation of all neural frequencies within a given time window (Cavanagh & Frank, 2014; Isbell & Grammer, 2021).

Referred to as response-locked event-related potential (ERP) components, the error-related negativity (ERN), correct-related negativity (CRN), and error positivity (Pe) are the most commonly reported measures of response monitoring in the literature (for a review see Gehring et al., 2011). These neural components have been documented in children as young as 3-years-of-age (Grammer et al., 2014) and linked to differences in task performance (Torpey et al., 2012; Wiersema et al., 2007), academic success (Hillman et al., 2012; Kim et al., 2016), and symptoms of depression, anxiety, and ADHD (Albrecht et al., 2008; Hanna et al., 2020; Musser & Raiker, 2019; Wauthia & Rossignol, 2016; Zhang et al., 2009).

The ERN. The ERN is a fast-occurring negative deflection, typically maximal at frontocentral sites and occurring around 50 milliseconds following an error commission (for a review see Gehring et al., 2011). The ERN is associated with a child's ability to detect that they have made an error (Falkenstein et al., 1991; Nieuwenhuis et al., 2001; Yeung et al., 2004). Greater (more negative) ERN amplitudes have consistently been observed when the correct answer or response was readily known or when the significance of making an error was perceived to be greater (e.g., Di Gregorio et al., 2016; Maier & Steinhauser, 2016). Thus, the ERN is thought to reflect a self-corrective mechanism in which the brain identifies a mismatch between the error response made and the known correct response (Yeung et al., 2004). Conflict monitoring theory (Yeung et al., 2004) states that greater neural response to errors in relation to correct responses is an index of better task monitoring skills and therefore a better understanding of performance.

Though few studies have investigated the development of the ERN in children under the age of 8 (Lo, 2018), the ERN has been observed in children as young as 3-years-old and develops logarithmically from mid-childhood to early adulthood (Grammer et al., 2014; Hogan et al., 2005; Santesso & Segalowitz, 2008; van Meel et al., 2012). However, reports of age-related differences in young children are mixed, with some studies reporting significant (Richardson et al., 2011) and others null (Grammer et al., 2014; Torpey et al., 2012) age-related differences in the ERN. For example, some studies suggest that ERN amplitude may be related more closely to variance in task performance than to chronological age (Downes et al., 2017; Hogan et al., 2005; Richardson et al., 2011).

The CRN. The CRN is a smaller, negative peak often observed within the first 100 ms after a correct response has been made, and generally follows a similar time course and topography as the ERN (Gruetzmann et al., 2014). The CRN has been identified as a distinct component, separate from the ERN, and is thought to reflect an attention cue to sustain engagement and avoid future errors (Luu et al., 2000). Though, the CRN has not historically been the main focus of studies investigating indices of response monitoring. Much more research has been published on developmental differences in the ERN as well as associations between the ERN and outcomes of interest. However, evidence does suggest that the CRN is enhanced (more negative) when children are younger and attenuates into adulthood as a function of development (Clawson et al., 2017).

The Pe. Accompanying the ERN/CRN complex is a more latent and broader positive component, the error positivity (Pe). The Pe is often maximal at centroparietal sites and can be observed approximately 200 to 500 milliseconds after an error response (for a review see (Overbeek et al., 2005). The Pe is thought to reflect conscious error detection, although not always observed for every error trial (Wiersema et al., 2007). Evidence indicates that larger (more positive) Pe amplitudes are associated with stronger salience of error (Nieuwenhuis et al., 2001) suggesting that the Pe reflects attention allocation to error commissions. The Pe has an earlier developmental maturation than the ERN, with adult-like patterns present in mid-childhood samples (Davies et al., 2004; Downes et al., 2017; Hogan et al., 2005; Richardson et al., 2011). However, in early elementary school, age-related differences in the Pe are still measurable (Grammer et al., 2014).

Associations between Brain and Behavior

While neural correlates of response monitoring have been under investigation for over three decades, their functional significance remains a point of debate. Some even question the purpose of assessing brain measures of cognitive control because behavioral performance measures have been shown to relate to academic achievement. However, combining neural measures with behavioral performance measures allows for investigation of both overt processing, like the commission of a response, and covert processing, like inhibiting a response or identifying whether one has made an error. Complicating our understanding, relations between ERP measures of error monitoring and task behavior, including reaction time and task accuracy, have been inconsistent (Eppinger et al., 2009). This, in part, may be due to differences among which neural (e.g., ERN, Pe) and behavioral (e.g., RT, accuracy) measures were chosen for analyses, which makes it difficult to identify the functional significance of error-related neural components. Additionally, evidence indicates that measurement error may be inflated in traditional quantifications of ERP average scores.

Neural measures of psychological processes assume that measurement is mainly driven by differences in trait constructs of interest or stimuli response (depending on the research question). This assumption inflates measurement error that may be due to individual state factors (e.g., mood, sleepiness), alternative trait characteristics (e.g., attention, maturation level), variance caused by filtering and averaging neural data, and other sources of measurement error (Gavin & Davies, 2007). There is preliminary evidence that raw averages of neural activity may not accurately portray the relation between trait levels of cognitive control, neural activity, and behavioral assessments of other outcomes of interest. Taylor and colleagues (Taylor et al., 2018) found no relation between brain and behavior in an SEM model relating constructs of neural and behavioral performance using indicator variables from a Go/No-Go task – two average E-wave (ERP) components and reaction times. Nor were brain and behavior related to a latent trait of attention, indicated by measures of attentional control and sustained attention. However, considerations such as differences in data collection procedures were left out of the model. It may, therefore, be important to adjust for child demographics such as age and data quality to help reduce measurement error in neural activity.

While some researchers have theorized that individual measures of task performance reflect similar but distinct cognitive processes, there is limited empirical evidence to support this claim and it remains and area of exploration. Researchers commonly use individual models investigating relations between outcomes of interest and either brain functioning or behavioral performance. Individual correlations limit our ability to investigate the complexity of cognitive engagement, which involves the coordination of inhibition, conflict monitoring, and response monitoring processes (Barkley, 1997; Senderecka et al., 2012). Alternatively, it may be beneficial to estimate children's cognitive control ability using multiple measures of brain and behavior from one task to capture the coordination of skills required for task performance. Indeed, a recent publication demonstrated that adult behavioral performance on a Go/No-Go task was related to activity in regions of the brain associated with response monitoring and the ERN, such as the anterior cingulate gyrus and inferior frontal gyrus (Weigard et al., 2020). Further, a meta-analysis of functional magnetic resonance imaging (fMRI) studies indicated that fMRI activity to No-Go trials was driven by the engagement of trait skills in both working memory and attention supporting that multiple cognitive subprocesses may be involved in the processing of Go/No-Go stimuli (Criaud & Boulinguez, 2013).

In addition to the complexities introduced by varying conventions of measurement, there are developmental considerations that further complicate our understanding of the associations between brain and behavior. Changes in children's cognitive skills occur in tandem with rapid development in associated brain regions throughout childhood and into adolescence. Indeed, there is some evidence that relations between brain and behavior may be dependent upon developmental maturity. Studies sampling 4 to 6 (Grammer et al., 2018), 5 to 7 (Torpey et al., 2012), and 7 to 8-year-old children (Wiersema et al., 2007), found null associations with behavioral performance for the ERN, CRN, and Pe. Yet, Grammer and colleagues (2018) found that larger ERN amplitudes were associated with slower reaction times in a six month follow up assessment with young children. Similarly, among 7 to 9-year-old children, Richardson et al. (2011) found that larger ERN amplitudes were related to more consistency in reaction time, but not task accuracy. However, Thurm and colleagues (Thurm et al., 2020) found that greater ERN amplitudes were associated with higher task accuracy in a slightly older sample of 9 to 11-year-old children. The CRN has been less reported in literature sampling children, however there is some evidence that greater (more negative) CRN amplitudes are associated with slower reaction times, though perhaps not when accounting for participant age (Torpey et al., 2012). Greater Pe amplitudes following error responses have been found to relate to increased age, better accuracy, and faster reaction times in Go/No-Go tasks (Thurm et al., 2020; Torpey et al., 2012). Additionally, larger Pe amplitudes have shown bivariate correlations with greater post-error slowing (Torpey et al., 2012) and post-error accuracy (Schroder et al., 2017), suggesting that children ages 5 to 8 already show a brain-behavior relation in the ability to consciously detect an error and adjust behavioral performance in response.

Combined, these findings suggest that there may be age-related differences in how error-related brain mechanisms are associated with behavior. Moreover, it remains a point of debate whether changes in the brain precede changes in children's behavior. Therefore, there may also be developmental differences in relations between academic skills and cognitive control measures of brain and behavior. Further study into relations between brain and behavior in elementary school is needed, with particular attention given to whether measures of neural processing help us predict academic outcomes above and beyond behavioral performance measures. By combining behavioral performance measures (e.g., reaction time, accuracy) with the collection of “real-time” neural data (i.e., ERPs) we can construct a comprehensive view of children’s cognitive functioning by extrapolating measures of underlying cognitive processes (e.g., performance monitoring, inhibition) that support successful task engagement (e.g., responses). Additionally, SEM modeling will allow us to better understand the dynamics between children’s brain and behavior and how functioning at each of these levels relates to outcomes of interest, such as academic performance. To my knowledge, little research has explored the cumulative variance that multiple ERP and behavioral measures may account for in academic outcomes.

The Current Study

Aims and Hypotheses

The current study aims to improve our understanding of how brain and behavior measures from a computerized cognitive control task relate to latent construct(s) of cognitive control and academic achievement in a sample of early elementary-school children. Applying advanced modeling to examine behavioral and EEG data from a child-friendly Go/No-Go task conducted with children aged 4-8, the study has three main aims:

Aim 1: Characterize Cognitive Performance

The first aim is to characterize children’s behavioral performance on a computerized Go/No-Go task tapping sustained attention and inhibition skills, using cross-classified IRT models. Task features that impact children’s performance are identified by investigating child, trial, and child-by-trial factors of accuracy and reaction time. Four independent models are used to estimate Go and No-Go accuracy and reaction time, adjusting for trial-by-trial variability in performance and other sources of measurement error.

It is hypothesized that child factors of age and gender, as well as trial difficulty will significantly impact a child's speed and likelihood of responding correctly to trial stimuli. Specifically, children that are older are expected to be more accurate and have faster reaction times. Boys are expected to demonstrate more trial-by-trial variability in performance. I also hypothesize that lower numbers of preceding Go trials will be related to greater accuracy on No-Go trials. Greater number of preceding Go trials are expected to be associated with faster reaction times for both Go and No-Go trials. No relation is expected between preceding context and Go trial accuracy.

Aim 2: Assess relations among behavioral performance measures and academics

The second aim is to leverage the various behavioral measures that can be extracted from the Go/No-Go to assess the cumulative variance that behavioral performance measures explain in children's reading and math skills. SEM models are used to assess relations between latent factors of academic skills (math, reading) and cognitive control. Academic factors of math and reading are estimated using indicator variables from standardized assessments of math and reading. Cognitive control is estimated as a latent construct(s), indexed by manifest or indicator variables from children's Go/No-Go task performance. Go/No-Go measures of accuracy (errors of omission, errors of commission, d-prime), reaction time (RTC, RTI, SDRTC, SDRTI), and post-error adjustment (PES, PEA) are used. It is hypothesized that a two-construct model of cognitive control – accommodating for differences in attention and inhibition - will fit the data best, based on previous literature modeling latent factors of cognitive control in school-age children (V. Anderson, 2002; De Franchis et al., 2017; Downing, 2016; Miyake & Friedman, 2012) and performance on Go/No-Go tasks (Winter & Sheridan, 2014; Wöstmann et al., 2013).

I also aim to assess the utility of estimating adjusted performance measures on the Go/No-Go, that accommodate for child, trial, and child-by-trial factors. Therefore, I compare two SEM models in predictions of academic skills - an Observed Model and an Adjusted Score Model. The Observed Model, utilizes observed measures (e.g., errors of omission) in Go/No-Go task performance as indicator variables for the latent cognitive control factor. The second model utilizes adjusted performance measures estimated in Aim 1 as indicator variables for the latent cognitive control factor (Adjusted Score Model).

I hypothesize that the Adjusted Score SEM model using indicator variables that account for differences in child demographics, task features, and task randomization in estimating the latent factor(s) of cognitive control will best predict differences in reading and math skills, as indicated by smaller standard errors in effect sizes. This hypothesis is supported by work indicating that Go/No-Go performance is related to trial difficulty in adult populations and to differences in age among children, adolescents, and adults. Findings from Aim 2 will help clarify best practices for quantifying behavioral measures of Go/No-Go performance as related to academic achievement.

Aim 3: Evaluate contributions of neural measures for understanding academic outcomes

The third and final aim is to determine the contribution of neural measures of cognitive control (i.e., the ERN, CRN, and Pe) in predicting unique variance in academic skills that are not able to be observed by behavior alone. It is predicted that incorporating neural measures of response monitoring as additional indicator variables for the latent construct(s) of cognitive control will increase the variance accounted for in both reading and math. These neural measures are thought to index unique covert cognitive processes that are similar but distinguishable from overt processes indexed by behavioral measures (Luu et al., 2000; Schroder et al., 2017; Torpey et al., 2012; Yeung et al., 2004). Further, I explore if reducing measurement error in neural activity is beneficial in constructing measures of cognitive control. Similar to Aim 2, I compare two models – an Observed Measures Model and an Adjusted Measures Model. Observed ERP measures are added as indicator variables for the cognitive control factor(s) in the Observed Measures Model. In the Adjusted Measures Model, adjusted measures of neural activity accommodating for child age and data quality will be used as indicator variables for the cognitive control factor(s). It is hypothesized that the Adjusted Measure SEM will better fit the data, as academic measures are normed for child ability and trial difficulty.

Methods

Data Source and Sample

The current project utilizes data from a larger longitudinal study that aimed to investigate the effects of schooling on cognitive functioning, as indexed using EEG and behavioral measures, in elementary school children. A total of 367 children were recruited from seven public elementary schools (26 classrooms) in the Midwestern United States. The seven schools in the study included populations of students from diverse SES and racial/ethnic backgrounds, with 60% of children attending a school where at least half of the students qualified for free or reduced-priced lunch. Forty-two percent of children attending the sampled schools were White, 33 % were Black, 17 % Asian, 2 % native American/Alaskan native, and 6 % more than one race. Percentage of students eligible for free or reduced-price lunch at the school level were 2%, 33%, 39%, 42%, 61%, 69% and 72%. Children's age at testing ranged from 4 to 8-years-old. Because of the longitudinal nature of the study, most children completed their first assessment in kindergarten and continued with annual follow up assessments through second grade. Some participants were added as new cohort participants in subsequent years. For this current project, I use data from children's first assessment, regardless of cohort year or grade of entry to the study. Two-hundred and ninety-seven children had behavioral data available from their first completion of the Go/No-Go task. However, five children were excluded due to low accuracy on Go trials (<25%) and thus, the final sample for Aim 1 was 292 children (M_{age}=6.15 years, N_{Male}=148). Descriptive information for the sample can be found in Table 1.

Table 1

Sample Descriptives

			School						
		N	1	2	3	4	5	6	7
	Full Sample	292	76	59	54	54	16	18	14
Grade	K	202	62	51	32	32	8	7	9
	1st	60	10	7	14	14	5	7	3
	2nd	30	4	1	8	8	3	4	2
%FRPL	Min	2.00	2.00	63.00	69.00	71.00	33.00	42.00	40.00
	Mean	47.23	2.67	67.32	70.70	72.37	33.00	42.00	40.00
	Max	83.00	5.00	71.00	76.00	83.00	33.00	42.00	40.00
	SD	29.03	1.26	3.28	2.21	2.64	0.00	0.00	0.00
Age	Min	4.96	4.96	5.08	5.03	5.16	5.25	5.24	5.09
	Mean	6.15	5.80	5.90	6.30	6.54	6.45	6.41	6.29
	Max	8.59	7.85	7.94	8.04	8.59	8.36	7.88	8.05
	SD	0.79	0.62	0.54	0.82	0.89	0.90	0.80	0.79

Measures

Children participating in the study completed individual assessments of cognitive control, EF, and academics as well as group EF assessments (see Ahmed et al. 2021). Additionally, children completed an EEG assessment. Study procedures were approved by the University of Michigan Institutional Review Board (Ahmed et al., 2021; Isbell & Grammer, 2021). The current project utilizes both behavioral and ERP measures of cognitive control as assessed during a child-friendly Go/No-Go task (Grammer et al., 2014) to estimate latent construct(s) of cognitive control. Latent measures of reading and mathematics will be estimated using observed scores on the letter word identification, reading comprehension, and applied problems subscales of the Woodcock Johnson III Tests of Achievement (WJIII; Mather, 2001) as indicator variables.

The Go/No-Go Task

Children in the study completed a child-friendly Go/No-Go task called the “Zoo Game” (Grammer et al., 2014). At the start of the task, children were given directions to the task embedded in a storyline. Children were told that they were playing a game in which their goal was to help a zookeeper put animals back in their cages. They were then informed that three of the animals (orangutans) were helping the zookeeper. Children were asked to press a button as quickly as they could every time they saw a picture of an animal (Go trials) but to inhibit their response each time they saw a helper (No-Go trials).

The task began with a brief practice block consisting of 12 trials (9 with Go images and 3 with No-Go images). Children then completed eight blocks of the task, each with 40 trials (each including 10 images of the helpers and 30 novel pictures), for a total of 320 trials. Each trial started with the presentation of a fixation cross displayed for a randomized interval ranging from 200-300 ms. Then, an image of an animal (the stimulus) was presented for 750 ms, followed by a blank black screen for 500 ms. Responses were registered during stimulus presentation as well as during the blank screen (response window=1250ms). All images were of the same size and were selected carefully so that the stimuli were easily identifiable from the background but were not particularly salient for other reasons. This was done to prevent children from being drawn to a particular stimulus because of the image background or other peripheral features. Children made responses on a standard game controller (Logitech Dual Action Gamepad USB; Logitech, Newark, CA, USA). Both speed and accuracy were emphasized; children were instructed to press the button as fast as possible, with regular reminders given not to press the button for No-

Go trials. No-Go trials are described as more difficult than Go trials because they require response inhibition skills and appear less frequently in this paradigm. To sustain enthusiasm and task engagement, children were provided with short breaks, as necessary.

Behavioral Outcome Measures. Reaction time (RT) was recorded as time (ms) from stimuli presentation to onset of response. Accuracy was also recorded within the response window (1250 ms). A correct response was indicated as a button-press for Go trials and the omission of a response for No-Go trials. RT was recorded for both correct Go-trials and incorrect No-Go trials. Trial performance (accuracy, reaction time) was extracted from E-prime data files whenever possible. Due to equipment error, trial performance for a subset of participants had to be calculated from marker files in the electrophysiological recording data (N=26). As such, data source was added a covariate in predictors of trial performance outcomes for IRT analyses (N=6494 trials).

For replication of traditional observed average scores, accuracy was analyzed as percent errors of omission (Go trials in which participants did not respond) and percent errors of commission (No-Go trials in which participants responded). Additionally, a d-prime score incorporating accuracy on both Go and No-Go trials was calculated (Z(Correct/Hit) – Z(Incorrect/False Alarm). An average RT for correct Go trials (RTC) and incorrect No-Go trials (RTI) was used. Additionally, a standard deviation of RT was used to assess variance in reaction times for Go (SDRT-C) and No-Go (SDRT-I) trials. Post-error slowing and post-error accuracy measures were calculated (average performance following error responses - average performance following correct).

Individual performance at the trial-level was utilized for estimating adjusted performance measures in Aim 1. Specifically, cross-classified multilevel IRT models were used to account for the unique presentation of stimuli that each child experienced and to estimate percent likelihood of responding correctly to Go and No-Go stimuli and average RTs. Accuracy probabilities were inverse coded to match traditional presentation of measures in terms of likelihood of erring. Separate models for Go and No-Go stimuli were used. Single values for accuracy (EOM, ECOM) and reaction time (RTC, RTI) were extracted for later SEM analyses (Adjusted Measure Model). Following model results, standard deviation in adjusted RT (SDRTC, SDRTI) and an adjusted d-prime score were also calculated for each participant. Additionally, in order to calculate PES and PES, an indicator variable of prior trial accuracy was utilized to allow for extraction of predicted performance following correct and erroneous responses.

ERP Outcome Measures. Electroencephalogram (EEG) recordings were made from 32 Ag/AgCl electrodes using a Biosemi Active 2 system and processed offline using customized EEGLAB and ERPLAB scripts (for more information see Isbell & Grammer, under review). The ERN and CRN were measured at site FCz, as an average voltage (uV) between -50 and 50 ms of response onset for error and correct trials, respectively. Pe was measured as the average voltage between 200 to 500 ms at site Pz, consistent with previous research. The morphology and amplitude of the correct Go and incorrect No-Go trials were different between frequently occurring early (0-750 ms) and infrequently occurring late (750-1250 ms) responses. Previous research indicates that there may be different cognitive processes underlying quick versus slow responses (Gehring et al., 2011). As such, late responses were excluded from analyses, meaning that all response-locked data comes from trials in which participants responded within the first 750 ms of stimulus onset (see Isbell & Grammer, 2021).

Due to differences in the quality of EEG data, the amount of “clean” data can vary substantially between participants. Particularly, movement and electrical noise create artifacts within EEG data that is difficult to extract. As such, average activity (uV) across trials at predefined time windows was utilized for investigation of response-locked components. Specifically, in the traditional comparison analyses, the ERN, CRN, and Pe components were utilized as raw averages (uV) to assess neural correlates of response monitoring. In the Adjusted Measures Model, ERP measures were estimated using multiple regression models that minimized age, data quality and other sources of measurement error in ERP values as suggested by Taylor et al. (2018).

Woodcock Johnson Tests of Achievement III

Math and reading skills were measured using subtests of the Woodcock Johnson III Tests of Achievement which is age and grade-normed for use in populations of 2-90+ years-of-age (Mather, 2001). Specifically, children completed the letter word identification, passage comprehension, and applied problems subtests (for a complete review of materials see Bradley-Johnson et al., 2004).

The Letter-Word Identification subtest includes 76 items and assesses children's early reading ability. The first 14 items assess the ability to name letters. Subsequent items include phonetically regular and irregular words, where phonetically irregular words cannot be decoded using phonics. Passage comprehension is a 47-item subtest in which most items ask participants to supply missing words in a passage, to assess children's understanding of what they read. The Applied Problems subtest assesses early math skill using word problems, pictures, and numbers and consists of 63 items. Participants were asked to listen to the item, determine the procedure to solve the problem, and successfully complete the computations. Participants are given a pencil and paper to help solve the items after they reach a certain level of difficulty in the task. For all subtests, items grow increasingly difficult as participants progress.

Due to lack of item-level data, internal consistencies of these measures were not calculated (see Ahmed et al., 2021). However, external validation efforts have demonstrated excellent reliability (Cronbach's $\alpha = .93$) for the academic achievement battery among elementary students and very high test-retest reliability ($r = .95$; Mather, 2001). The current project uses *W* scores of subtest performance, which accommodate for differences in item difficulty in estimations of child ability (Jaffe, 2009).

Missingness and Inclusion Criteria

Children with missing data for behavioral, ERP, or academic measures were deleted list-wise, given a lack of literature on which to base predefined expected associations amongst these variables. To be included in the sample, children were required to have completed at least half of the Go/No-Go task, with higher than 25% of accuracy on Go trials. It is suspected that children who do not reach this criterion do not fully participate in the task (see Isbell & Grammer, 2021). Additionally, children needed to have at least 6 epochs of usable ERP data for both correct Go and incorrect No-Go responses, in compliance with standards for ERP analyses in children (Pontifex et al., 2010).

In Aim 1, cross-classified IRT models were based on 240 Go trials and 80 No-Go trials. However, due to task design and inclusion criterion, children had variable number of trials for reaction time and accuracy variables. Additionally, trials in which children responded in under 50 ms of Go/No-Go stimuli were excluded from behavioral analyses, consistent with methods used in the original and subsequent implementations of the task (Torgrimson et al., 2021).

Statistical Approach

The analytic approach for the project was hierarchical, such that main analyses for each aim was dependent on the results from the preceding aim. Specifically, analyses for Aim 1 produced adjusted performance measures of behavioral performance on the Go/No-Go task that were used to estimate latent construct(s) of cognitive control in Aim 2. Aim 3 then built off of the final model from Aim 2 in order to examine how neural measures load onto behavioral models of cognitive control.

Aim 1: Characterize Cognitive Performance

Cross-classified multilevel IRT modeling was used to estimate accuracy and reaction time measures of Go/No-Go performance. A cross-classified multilevel IRT model is best suited to minimize measurement error and estimate adjusted performance scores that account for randomization of trials across participants. These adjusted performance scores theoretically reveal “true” measures of trait constructs (e.g., inhibition), that are not conflated by extraneous factors or measurement error. IRT models theoretically separate measurement of the participant from performance on the item, or in this case, trial. Using a probabilistic approach, logistic IRT models produce probability scores for accuracy on an item (trial), given the child’s ability and the item’s difficulty (Kamata, 2001). A cross-classified multilevel structure further allows for separation of factor effects on responses that can be classified across an item (e.g., trial 1 of task) and individual participant. In this project, I refer to this as a child-by-trial factor (Murphy & Beretvas, 2015).

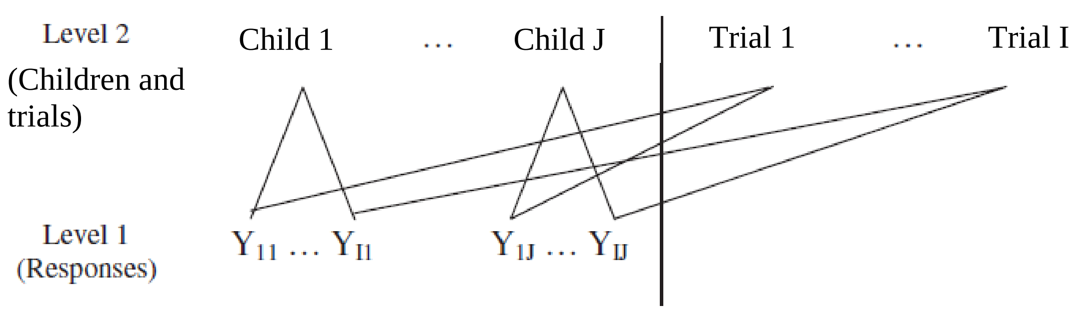
Stimuli in the Zoo Game were randomly presented within each block of the task. Therefore, participants completed a similar task in terms of overall percent of Go/No-Go stimuli, but at the level of individual stimulus presentation, each child completed a unique set of trials. For example, trial 1 could have been presented as either a Go or No-Go stimulus for each child.

Additionally, the preceding context of each trial was also unique across participants. Thus, child A might have been presented with a No-Go stimulus as the 8th trial, preceded by 3 Go stimuli. Whereas, child B may also have been presented with a No-Go stimulus at the 8th trial but preceded by a No-Go stimulus. Differences in preceding trial stimuli may alter the likelihood of responding appropriately and the reaction time of response for each trial. Because trials were randomized within each participant, preceding context was considered a child-by-trial factor.

Figure 1 depicts the cross-classified structure of factors that attribute to behavioral performance (i.e., response). As shown, response performance (Level 1) can be attributed to Level 2 factors that are related to trial (*i*) variance (e.g., trial stimuli), child (*j*) variance (e.g., age, gender), and child-by-trial (*ij*) variance (e.g., the randomization of trials across children).

Figure 1

IRT Model of Behavioral Performance Modelled from Van den Noortgate, De Boeck, and Meulders (2003).



Logistic models were used for estimating accuracy variables (EOM, ECOM) and linear models were used to estimate reaction time for correct Go (RTC) and incorrect No-Go (RTI) trials. Scores do not need to be continuous or on an interval scale to meet IRT model assumptions. Therefore, a child's accuracy estimate is based on the likelihood of responding correctly given the child's ability, trial difficulty, and unique order of trials that the child received. Similarly, a reaction time estimate accounts for child ability, trial difficulty, and randomization of trials. Model equations for logistic and continuous models read as follows:

$$\text{Logistic: } Logit(Y_{ij}) = \beta_0 + \beta_1 C_{1j} + \beta_2 C_{2j} + \beta_3 C_{3j} + \beta_4 T_{1ij} + u_{1j} + u_{2i}$$

$$\text{Continuous: } Y_{ij} = \beta_0 + \beta_1 C_{1j} + \beta_2 C_{2j} + \beta_3 C_{3j} + \beta_4 T_{1ij} + u_{1j} + u_{2i}$$

$$u_{ij} \sim N(0, \sigma_{u1}^2)$$

$$u_{2i} \sim N(0, \sigma_{u2}^2)$$

The logistic models demonstrate the likelihood that child j will accurately respond to trial *i*. The continuous models predict the average response time to trial *i* for child *j*. Child level covariates are denoted with *C* and child-by-trial covariates are denoted with *T*. There are three child covariates of age (C_{1j}; continuous), data source (C_{2j}; dichotomous), and gender (C_{3j}; dichotomous). The effect of these covariates on task performance are denoted by β_1, β_2, and β_3 respectively. There is one child-by-trial covariate, trial difficulty, operationalized as the number of preceding Go trials (T_{1ij}; continuous). The effect of trial difficulty on response is denoted as β_4. Random effects for child and trial are denoted by u_{1j} and u_{2i}, respectively. Variance parameters for child and trial effects are represented as σ_{u1}^2 and σ_{u2}^2, respectively.

A series of models using restricted and unrestricted parameters were assessed to find best model fit using the R software (R Core Team, 2017) package lme4 (Bates et al., 2015). Full information maximum likelihood (FIML) was used to calculate information criteria for model comparisons. In line with previous literature, a sample-corrected Bayesian information criterion (BIC; Schwarz, 1978) and finite Akaike’s information criterion (AIC; Hurvich & Tsai, 1989) were used as information criteria (Beretvas & Murphy, 2013). A lower information criteria value (>2) indicated better fit.

Adjusted accuracy and reaction time measures calculated in Aim 1 were used for modeling cognitive control skills in aim 2. Specifically, adjusted RTC and RTI were estimated as a child's average of predicted values for trial reaction time to Go and No-Go stimuli, respectively. Adjusted measures of EOM and ECOM were calculated as an average in a child's probability for error response on Go and No-Go trials, respectively. Other adjusted measures of accuracy (d-prime), reaction time (SDRT-C, SDRT-I), and post-error adjustments (PES, PEA) were calculated post hoc.

Aim 2: Assess relations among behavioral performance measures and academics

The second aim of my dissertation was to assess relations among behavioral performance measures and children's reading and math skills. I also sought to evaluate the utility of estimating adjusted scores that accommodate for child-by-trial variance in behavioral measures of Go/No-Go task performance. A series of SEM models were used to compare the amount of variance accounted for in academic skill by cognitive control when estimated using the Observed indicator model or the Adjusted Score indicator model. The R software (R Core Team, 2017) package lavaan (Rosseel, 2012) was used to estimate models. SEM creates weighted aggregates of latent factors and estimates both factor and error variance in the model. This is superior to composite scores that average performance across variables because it allows for indicator variables to account for varying degrees of variance in the factor of interest. Similarly, traditional regression analyses assume variables are measured without error, so modeling error variance adjusts provides a more accurate estimation of between factor relations (Weston & Gore, 2006).

The recommended two-step approach was used to estimate the SEM models (Kline, 2015). In step 1, confirmatory factor analyses (CFA) models were used to estimate factors of cognitive control and academic skills. For academic skills, a one-factor CFA was used entering subtests of the Woodcock Johnson as indicator, or manifest, variables. Measures were allowed to covary. For cognitive control, both a one-factor and two-factor model was tested using behavioral measures from the Go/No-Go task. This was done in two iterations, 1) using observed scores from the Go/No-Go (Observed Model) and 2) using adjusted performance estimates calculated in aim 1 (Adjusted Score Model). Measures were allowed to covary and load onto both measures of attention and inhibition.

In step 2, I compared full SEM models testing hypothesized interrelations between factors. Both 1-level and 2-level SEM models with children's performance nested within schools were assessed. Final CFA models of cognitive control estimated from Observed and Adjusted Measure indicator variables were used as predictors of academic skills. Strength of relations between cognitive control and academic skills were used to assess whether Observed or Adjusted Measure models of cognitive control were better at predicting variability in academic skills. For all models, FIML was used to calculate the chi-square difference test, to assess the difference in model deviance statistics, which is assumed to be chi-square distributed (Hox, 2002; Raudenbush & Bryk, 2002). Chi-square difference tests are appropriate when data is in a nested structure, versus the cross-classified structure in Aim 1 methods (Beretvas & Murphy, 2013).

Aim 3 Evaluate contributions of neural correlates of cognitive control

My third and final aim was to examine how neural measures of response monitoring load onto behavioral factors of cognitive control and whether these additional indicators change the strength of predictions between cognitive control and academic skills. I first used multiple regression models to estimate Adjusted ERP measures that minimized confounds associated with age, data quality (Z-score of artifact-free ERP epochs), and other sources of measurement error. Equation for regression models was as follows:

$$\gamma_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

The continuous models estimated average ERP (CRN, ERN, or Pe) amplitudes for child i. There are two child covariates of age (X_{1i}; continuous) and data quality (X_{2i}; continuous). Developmental differences in ERP measurement were indexed by β_1, the fixed effect parameter of age. Data quality was calculated as a Z-score of artifact-free ERP epochs. Larger numbers of artifact-free ERP epochs indicate more usable brain data. By calculating a standardized score, I was able to examine the effect of having above or below average data quality on estimates of ERP measures, as represented by the fixed effect parameter β_2. Therefore, in this model, β_0 represents the average expected ERP amplitude for children aged 0 years with an average number of artifact-free epochs. Measurement error was estimated by the random effect parameter ε_i, with variance σ^2.

Using the best-fit model from Aim 2, I then tested two additional SEM models, incorporating either 1) Observed Raw Average ERP measures (Observed Measures Model) or 2) Adjusted Measures of ERPs (Adjusted Measure Model). Model fit for the CFA step in each model was determined using FIML and chi-square difference tests. SEM models in step 2 were compared using the strength of relation between academic factors and cognitive control.

Power Analysis

To my knowledge, there is no calculator or standard for power analysis with a cross-classified multilevel IRT. However, when considering power for these types of models, total sample size is determined as the product of Level-1 and Level-2 sample sizes. Thus, trial count *and* sample size are both important factors for power consideration in the proposed project. The classification structure of the proposed models puts the level-2 sample size as the number of children (n=292) and the number of trials (n>40). Level-1 sample size (number of responses) is also equivalent to number of trials (n>40). Therefore, my minimum total sample size is 11,680 (40*292). Bevertas and Murphy (2013) compared sample sizes in a 2-level cross-classified model where minimum sample size across was 625, representing 25, level-1 students cross-classified across 25, level-2 middle and high schools. Results indicated that the sample size of level-1 was most related to model convergence. As their study used level-1 sample sizes of 25 and 50, a minimum of 40 in the current study should be sufficient. Additionally, they found that more balanced distributions of and higher degrees of cross-classification also improved convergence rates. Because every child completed at least 40 trials, I will also have a high degree of cross-classification and more balanced distributions in my proposed models.

For the SEM models outlined in Aims 2 and 3, power analyses were conducted to determine the minimum sample size needed to detect medium and large effect sizes with an alpha of = 0.05 and power = 0.95 using Soper's online calculator (2021). For an SEM with four latent variables and 14 indicator variables, the results indicated a recommended sample size of 207 participants. Specifically, 207 participants to detect an effect and 138 for the model structure. Notably, the recommended minimum sample size decreased when adjusting for more conservative probability levels (.025, .0125). The sample size required to detect an effect also significantly decreased when using a three-factor model (if one factor of cognitive control is determined to be the best fit for behavioral data). For Aim 2, the available sample size was 287 children (N_{Male}=144, M_{age}=6.14) with usable behavioral and academic data, indicating sufficient power. Similarly, in Aim 3, 246 children (N_{Male}=122, M_{age}=6.16) had complete behavioral, neural, and academic data. One child did not have information about which school they attended, and so they were removed from the 2-level SEM analyses.

Results

For clarity, results are organized by aim and analytic approach. Traditional observed behavioral measures include average (correct RT, incorrect RT, errors of omission, errors of commission, PES, PEA), variance (SDRT correct, SDRT incorrect), and standardized measures (d-prime). Alternative analytic approaches include IRT models of behavioral performance and regression models of neural measures. Adjusted performance scores are operationalized as predicted performance scores as estimated from the IRT and regression models.

Aim 1: Quantifying Behavioral Performance

Traditional Observed Behavioral Measures

Summary statistics for behavioral measures are presented in Table 2. On average, children performed well on the Go/No-Go task with an average of 8.93% errors of omission (EOM) on Go trials and 39.35% errors of commission (ECOM) on No-Go trials. However, there was a range of performance with EOM ranging from 0.00% to 61.92% (*sd*=10.00%) and ECOM ranging from 5.00% to 96.92% (*sd*=18.63%). A variance ratio test indicated that boys had greater variance in RT Correct than girls ($F(147)=1.40$, $p=0.43$). Additionally, boys had greater variance between trial reaction times on correct Go trials ($F(30922)=1.07$, $p<.001$) but less variance than girls in reaction time for incorrect No-Go trials (*F*(4863)=0.88, *p*<0.001). No gender differences in variance were found for accuracy measures.

Table 2

Behavioral Data: Descriptive Statistics

Variable	Min	Mean	Max	SD
EOM	0.00	8.93	61.92	10.00
ECOM	5.00	39.35	96.92	18.63
RT Correct (ms)	362.35	555.99	769.76	71.35
RT Error (ms)	309.67	470.01	667.54	71.57
SDRTC (ms)	58.76	158.83	295.01	42.89
SDRTI (ms)	29.77	153.75	352.87	61.64
PEA	-70.72	-9.57	7.58	13.31
PES (ms)	-131.29	17.45	178.91	51.51
d-prime	-0.42	1.86	3.74	0.83

On average children responded faster to No-Go trials, *t*(291)=27.05, *p*<.001 and had more variability in trial reaction times for Go (as indexed by SDRTC) than No-Go trials (SDRTI), *t*(291)=2.14, *p*=0.033. Variability within subjects on reaction time measures (SDRTC, SDRTI) was also descriptively larger than variability between subjects (sample standard deviation of RT Correct and RT Error; see Table 2). On average, children showed post-error slowing (*mean*=17.45 ms) but not post-error improvements in accuracy (*mean*=-9.57%), as mean values were negative. Yet, there was a subset of children that demonstrated post-error improvements in accuracy (N=71, 24.32%). More than half of the children demonstrated post-error slowing (N=183, 62.67%). Three children had negative d-prime values, indicating that their false alarm (ECOM) rates were higher than their accuracy on Go trials. There were outliers present in all behavioral variables. As such, all measures were winsorized at a 10% level, or to 2 standard deviations of the mean, allowing for maximal sample retention (Dixon, 1960).

Table 3

Correlations between Behavioral Measures

	EOM	ECOM	RTC	RTI	SDRTC	SDRTI	d-prime	PEA	PES	Age
EOM		0.052	0.208***	0.265***	0.456***	0.441***	-0.693***	-0.752***	0.149*	-0.302***
ECOM	0.378		-0.457***	-0.111+	0.425***	0.307***	-0.704***	0.127*	0.054	-0.009
RTC	<.001	<.001		0.727***	0.320***	0.364***	0.112+	-0.239***	0.096	-0.267***
RTI	<.001	0.057	<.001		0.302***	0.504***	-0.139*	-0.241***	0.113+	-0.268***
SDRTC	<.001	<.001	<.001	<.001		0.766***	-0.633***	-0.273***	0.190**	-0.230***
SDRTI	<.001	<.001	<.001	<.001	<.001		-0.540***	-0.290***	0.172**	-0.233***
d- prime	<.001	<.001	0.057	0.018	<.001	<.001		0.425***	-0.155**	0.261***
PEA	<.001	0.031	<.001	<.001	<.001	<.001	<.001		-0.181**	0.219***
PES	0.011	0.359	0.101	0.055	0.001	0.003	0.008	0.002		0.066
Age	<.001	0.881	<.001	<.001	<.001	<.001	<.001	<.001	0.263	

Note: N=292 for all variables. Above the diagonal line are *r*-values and below the diagonal, are associated *p*-values.

***p<.001,**p<.01,*p<.05,+p<.01

A correlation table, indicating relations between winsorized behavioral variables and age is presented in Table 3. Trial accuracy variables were not correlated ($r(292)=0.05$, $p=.378$), providing one piece of evidence that accuracy on Go and No-Go trials may reflect different cognitive processes. All trial reaction time measures were positively correlated, indicating similar average ($r(292)=0.73$, $p<.001$) and standard deviation ($r(292)=0.77$, $p<.001$) values between Go and No-Go trials. Additionally, RTC was correlated with both EOM and ECOM, such that more EOM were related to slower reaction times ($r(292)=0.21$, $p<.001$), whereas more ECOM were related to faster reaction times on the frequent Go trials ($r(292)=0.21$, $p<.001$). The d-prime measure was negatively correlated with trial accuracy and standard deviation in trial reaction times. This finding demonstrates that higher d-prime values, indicating better ratio of correct "hit" Go trials to incorrect "false alarm" No-Go trials, was related to lower errors of omission ($r(292)=-0.69$, $p<.001$) and commission ($r(292)=-0.70$, $p<.001$) and less variable reaction times for Go ($r(292)=-0.63$, $p<.001$) and No-Go ($r(292)=-0.54$, $p<.001$) trials. Post-error accuracy and was most correlated with EOM ($r(292)=-0.75$, $p<.001$), such that fewer EOM was related to more positive values of PEA, indicative of post-error adjustments in behavior. Age was correlated with better cognitive performance on all behavioral variables except ECOM and PES, such that older children were more likely to have fewer EOM, faster and less variable reaction times, and more positive scores on d-prime and PEA variables.

Next, gender was assessed as a between-subjects factor for winsorized behavioral measures. Table 4 shows the results of corresponding t-tests. Performance differed as a function of gender across three behavioral measures. On average, females had fewer percent ECOM ($t(290)=3.65$, $p<.001$), longer reaction times to correct Go trials ($t(290)=-2.70$, $p<.001$), and

Table 4

Gender Differences in Behavioral Data

Variables	*Mean* Male	*Mean* Female	*t*	*p*
Age	6.17	6.12	0.52	.601
%EOM	8.67	8.05	0.65	.516
%ECOM	42.47	35.18	3.65***	<.001
RT Correct	545.33	566.82	-2.70**	.007
RT Error	462.32	475.76	-1.69	.093
SDRTC	159.02	157.73	0.27	.790
SDRTI	150.09	154.74	-0.68	.497
%PEA	-8.37	-9.35	0.75	.453
PES	13.74	21.04	1.29	0.199
d-prime	1.73	2.00	-2.87**	.004

better (more positive) d-prime scores ($t(290)=-2.87$, $p=.004$). Overall, results indicated that female participants showed better response inhibition skills and had higher percentages of successful “hit” rates on Go trials as compared to incorrect “false alarm” rates on No-Go trials.

IRT Models of Behavioral Performance

Tables 5 and 6 demonstrate the between and within subject trial-level descriptive statistics, including the range of each variable included in IRT analyses. For each behavioral outcome variable (EOM, ECOM, RT Correct, RT Error), the same six models were assessed. Model 1 included predictions of the fixed intercept and the random effect of level-2 factors of trial and child ID. Model 2 built off of Model 1, adding the predictor variable of trial difficulty (number of preceding Go trials). Covariates were then incrementally added in each model iteration, such that data source, gender, and age were added as one additional covariate in Models 3, 4, and 5, respectively. Last, a dichotomous indicator variable identifying whether the trial was following a correct or erroneous response was added to Model 6.

Table 5

Between Subject Trial-Level Descriptive Information

Trial Type		N	Trial Difficulty	Post-Correct	Post-Error
All	Min	160	0	121	10
	Mean	309.23	6.82	252.29	49.13
	Max	320	22	301	146
	SD	25.92	4.48	36.93	25.53
Go Trials	Min	120	0	88	4
	Mean	231.76	6.37	188.87	37.13
	Max	240	21	228	109
	SD	19.45	4.23	27.89	19.42
No-Go Trials	Min	40	0	25	1
	Mean	77.45	5.70	63.42	12.00
	Max	80	22	78	39
	SD	6.59	4.04	9.67	6.83

Table 6

Within Subject Trial-Level Descriptive Information

Trial Type		Trial Difficulty
All	Min	0-0
	Mean	2.02-3.21
	Max	8-22
	SD	1.99-3.71
Go Trials	Min	0-0
	Mean	1.85-3.27
	Max	7-21
	SD	1.79-3.76
No-Go Trials	Min	0-0
	Mean	2.43-3.00
	Max	8-22
	SD	2.23-3.76

Note: Values indicate a range between participants. For example, some participants had a maximum of 8 preceding Go trials, while some had a maximum of 22 preceding Go trials.

Model 6 was the best fit model for all outcome measures. Table 7 shows the factor estimates for Model 6 for each of the four main outcome variables. Differences in performance between trials accounted for approximately 2.53% of the variance in RT Correct compared to differences between children accounting for approximately 12.23% of the variance in performance. For RT Error, differences between trials accounted for approximately 1.39% of variance and differences between children accounted for approximately 10.10%. Trial variance was expectedly smaller for logistic regression models, as well, compared to variance between children. Trial variance was approximately 0.08(*sd*=0.29) compared to 1.07(*sd*=1.03) for child variance in EOM. Similarly, trial variance was 0.02(*sd*=0.13) compared to 0.71(*sd*=0.84) for child variance in ECOM.

Table 7

Parameter Estimates for Predicted Value Models

	RT Correct	RT Error	EOM	ECOM
Fixed Effect (SE)				
(Intercept)	729.71(2907)***	621.63(32.14)***	0.03(0.52)	-0.24(0.42)
Trial Difficulty	-4.64(0.26)***	-1.48(0.69)*	0.01(0.01)*	0.02(0.01)***
Data Source	-111.54(12.79)***	-50.25(14.27)***	0.42(0.22)	0.29(0.18)
Female	16.54(7.23)*	11.61(7.98)	-0.06(0.13)	-0.38(0.10)***
Age	-26.54(7.23)	-25.03(5.09)***	-0.53(0.08)***	-0.02(0.07)
Post-Error	10.84(1.99)***	22.01(5.35)***	1.19(0.03)***	-0.10(0.04)*
Random Effect (SD)				
Trial	758.1(27.53)	465.8(21.58)	0.08(0.29)	0.02(0.13)
Child	3666.2(60.55)	3383.2(58.17)	1.07(1.03)	0.71(0.84)
Residual	25564.5(159.89)	29661.1(172.22)		
Fit Statistics				
AIC	784257.2	113648.5	31943.1	27139.2
BIC	784338.3	113712.0	32015.9	27203.2
Log Likelihood	-392119.6	-56815.2	-15963.6	-13561.6
Deviance	784239.2	113630.5	31927.1	27123.2
Residual (df)	6056	8602	65984	22017

Note: ***p<.001, **p<.01, *p<.05, +p<.01

Notably, trial difficulty was a significant predictor in all finalized models. Trial difficulty was negatively related to reaction time for both No-Go (β=-1.48, p=.027) and Go (β=-4.64, p<.001) trials, such that larger numbers of preceding Go trials were related to faster reaction times. Specifically, for every additional preceding Go trial, reaction time is expected to decrease (get faster) by 4.64 ms for Go trials and by 1.48 ms for No-Go trials, while holding all other predictors constant. Greater number of preceding Go trials was also related to increased probability of error on No-Go trials (β=0.02, p<.001) and Go trials (β=0.01, p=.048). In other words, it was harder for participants to inhibit a response to No-Go trials or sustain vigilance to Go trials if the context of the previous trials included a greater number of Go trials. For ECOM, each additional preceding No-Go trial is thought to relate to a 0.42%-0.50% increase in probability of error, given that all other variables remain constant. Similarly, for EOM, each additional preceding Go trial is thought to relate to a 0.03%-0.14% increase in probability of error for children of average age, given that all other variables remain constant.

The post-error indicator was also statistically significant of performance on accuracy and reaction time measures for both Go and No-Go trials. Specifically, average children are expected to demonstrate post-error slowing for both Go (β=10.84, p<.001) and No-Go trials (β=22.01, p<.001) as well as an increased likelihood of making an EOM (β=1.19, p<.001) following error, but a decreased negatively likelihood of an ECOM (β=-0.10, p=.022) following error. This finding suggests that in the current sample, children may have oriented too much to an error and slowed their reaction time, leading to less task engagement or greater motor inhibition (greater chance of EOM and less chance of ECOM). For children of average age, probability of making an error is expected to increase by 7.32-10.85% for Go trials and decrease by 2.14%-2.49% for No-Go trials for post-error versus post-correct trials if the trial was preceded by a No-Go trial (0 preceding Go) and other predictors are held constant. That means for a male child of average age with data from E-prime, if the current trial has is preceded by 0 goal trials and a correct response, their probability of making an error of omission is expected to be 3.8%, whereas their probability of making an error of commission is expected to be 41.02%. If that child had a trial with 22 preceding Go trials (the maximum observed in the current sample), we would expect the probability of EOM to be 4.7%, and the probability of ECOM to be 51.92%. Likewise, the same child's probability of exhibiting an EOM would be 11.5% and their probability of demonstrating an ECOM would be 38.63% if the trial was preceded by an erroneous response but 0 Go trials. Whereas for a trial following an error and 22 preceding Go trials, the likelihood of error would be 13.95% for EOM and 49.43% for ECOM.

Following model specification, predicted scores were calculated utilizing parameter estimates identified in Model 6, which include analytic adjustments for trial randomization, trial difficulty, age, gender, and data source as well as random effects of trial and child variance. Because parameter estimates of trial difficulty did not differ practically between models 2-5 (see Tables A1-A4 in Appendix A), all covariates regardless of statistical significance were retained as control factors for estimating predicted values, as gender, age, and data source were significant in at least one of the models. Of note, estimated scores were not calculated for the first trial in each block, as it was not directly preceded by a trial (max=8 per participant). Predicted performance measures of accuracy (EOM, ECOM, d-prime), reaction time (RTI, RTC, SDRTI, SDRTC), and post-error adjustments (PEA, PES) were then calculated for each child to index an "adjusted" performance measure. These predicted measures allowed for a comparison of factor loadings in observed versus adjusted indicators of cognitive control.

Adjusted (predicted) versus observed performance measures are presented in Table A5 (see Appendix A). Mean scores between predicted and observed measures were comparable. However, predicted values had smaller standard deviations in performance between children. Descriptive statistics for residual scores are presented in Table 8. Predicted scores were closest to the fitted model for EOM and d-prime scores. Reaction time variables of SDRTC, SDRTI, and PES had the largest residual variances. Notably, all SDRTC and SDRTI values were positive, indicating that observed standard reaction times were more variable than predicted standard deviations in reaction time. Negative residual values indicate better observed performance for EOM and ECOM measures (less errors) and slower reaction times (RTC, RTI) for observed than predicted measures. In contrast, negative values indicate less post-error adjustments in accuracy (PEA) and slowing (PES) as well as smaller ratios between hit and false alarm rates (lower d-prime performance) for observed than predicted measures.

Table 8

Residual Values: Descriptive Statistics

Variable	Min	Mean	Max	SD
%EOM	-0.79	0.15	2.70	0.53
%ECOM	-3.76	-0.07	5.92	1.67
RT Correct (ms)	-6.81	1.34	15.94	2.76
RT Error (ms)	-108.12	2.38	72.90	18.74
SDRTC (ms)	30.80	130.98	265.15	42.55
SDRTI (ms)	14.28	138.36	337.72	60.98
%PEA	-55.99	0.72	23.99	9.10
PES (ms)	-146.87	-1.07	143.12	50.91
d-prime	-0.46	0.02	0.48	0.11

Table 9

Correlations between Predicted Behavioral Measures

	EOM	ECOM	RTC	RTI	SDRTC	SDRTI	d-prime	PEA_W	PES_W	age
EOM		0.049	0.197***	0.269***	0.113+	0.199**	-0.714***	-0.974***	-0.132*	-0.309***
ECOM	0.405		-0.460***	-0.117*	0.105+	0.195**	-0.694***	-0.020	0.196**	-0.015
RTC	0.001	0.000		0.757***	0.150*	0.040	0.118*	-0.269***	-0.271***	-0.271***
RTI	0.000	0.046	0.000		0.183**	0.171**	-0.144*	-0.320***	-0.115*	-0.343***
SDRTC	0.054	0.073	0.010	0.002		0.111+	-0.152**	-0.103+	0.321***	-0.080
SDRTI	0.001	0.001	0.498	0.003	0.059		-0.255***	-0.188**	0.017	-0.110+
d-prime	0.000	0.000	0.045	0.014	0.009	0.000		0.710***	-0.031	0.280***
PEA_W	0.000	0.736	0.000	0.000	0.078	0.001	0.000		0.167**	0.352***
PES_W	0.024	0.001	0.000	0.049	0.000	0.778	0.596	0.004		0.099+
Age	0.000	0.797	0.000	0.000	0.174	0.060	0.000	0.000	0.090	

Note: N=292 for all variables. Above the diagonal line are *r*-values and below the diagonal, are associated *p*-values.

***p<.001, **p<.01, *p<.05, +p<.01

Correlations between predicted measures are presented in Table 9. Results reflect many of the same patterns of relations between observed variables outlined in Table 3. Main differences include the standard deviation of predicted values of RT (SDRTI, SDRTC), such that almost all correlations with SDRT measures decreased, perhaps due to the smaller range in these variables for adjusted versus observed measures. Notably, some correlations with post-error adjustment variables increased. PEA was highly correlated with EOM ($r(292)=-0.97$, $p<.001$) and d-prime ($r(292)=0.710$, $p<.001$) for predicted values, indicating greater improvements in accuracy following error were related to fewer EOM and better d-prime scores. For predicted values, greater PES was related to increased average ECOM ($r(292)=0.20$, $p=.001$) and faster overall reaction times for Go ($r(292)=-0.27$, $p<.001$) and No-Go trials ($r(292)=-0.12$, $p<.001$), as well as more trial variability in reaction time for Go trials ($r(292)=0.321$, $p<.001$). Factor loadings of predicted versus observed measures of task behavior onto a latent construct of cognitive control are described below in the Estimating Cognitive Control Construct section.

Aim 2: Relations between Cognition and Academics

Sample descriptives for academic subscales of the Woodcock-Johnson are presented in Table 10. Note, values are children's *W* scores on the individual subscales. *W* scores are centered at a value of 500, representing the expected performance for a child aged 10 years or at the beginning of Grade 5 (Jaffe, 2009). As indicated in Table 9, there is quite a large range in *W* scores represented in the sample, with some children in early elementary school demonstrating the expected ability of a 5^{th} grader and others performing at or below grade level. The largest range in performance was captured in the Letter Word subscale and the smallest range found for the Applied Problems subscale. *W* scores are on an equal-interval scale, such that a difference of value *x* in scores, is the same along the continuum in each subscale. In other words, the difference in skills between children with *W* scores of 300 and 304 is the same as the difference in skills between children with *W* scores of 480 and 484. To retain outliers in the sample, adjust for skewness in the sample distributions, and align the scaling of observed variances among measures, all variables were winsorized at a 10% level (Dixon, 1960), or to 2 standard deviations of the mean, and then standardized to a Z-score, prior to CFA analyses.

Table 10

Sample Descriptives for Academic Measures

Variable	Min	Mean	Max	SD
Applied Problems	372	439.63	506	18.86
Passage Comprehension	377	430.86	497	27.86
Letter Word ID	283	402.28	525	40.77

Table 11

Bivariate Correlations between Observed Behavior Measures and Academic Skills

	Letter Word ID	Passage Comprehension	Applied Problems
EOM	-0.261***	-0.213***	-0.262***
ECOM	-0.135*	-0.140*	-0.185**
RTC	-0.159**	-0.139*	-0.130*
RTI	-0.153*	-0.162**	-0.131*
SDRTC	-0.303***	-0.278***	-0.288***
SDRTI	-0.243***	-0.244***	-0.236***
d-prime	0.308***	0.285***	0.324***
PEA	0.191**	0.191**	0.180**
PES	-0.013	0.006	0.046
Age	0.519***	0.450***	0.443***
Letter Word ID		0.829***	0.693***
Passage Comprehension	0.829***		0.588***
Applied Problems	0.693***	0.588***	

Note: N=287 for all variables.

***p<.001, **p<.01, *p<.05

Correlations between winsorized and standardized scores for observed behavioral measures and academic skills are presented in Table 11. Overall, bivariate correlations between behavior and academic outcomes were small, but in expected directions. Better academic skills were related to fewer EOM and ECOM, faster reaction times for Go and No-Go trials, less variance in reaction times between trials (SDRTC, SDRTI), better d-prime scores, and greater post-error adjustments in accuracy. The d-prime measure had the highest correlation with academic measures (0.285>*r*<0.324). Post-error slowing was not related to any of the subscales of the Woodcock-Johnson. All academic subscales were also positively related with one another, and to age such that older children demonstrated higher math and literacy skills.

Correlations between observed behavioral measures and academic skills were similar in direction and magnitude to those between academics and predicted behavioral measures (see Tables 11,12). Some correlations increased slightly in strength, most notably between PEA and academic subscales. Conversely, correlations between standard deviation in reaction time measures (SDRTC, SDRTI) decreased or even lost statistical significance.

Table 12

Bivariate Correlations between Predicted Behavior Measures and Academic Skills

	Letter Word ID	Passage Comprehension	Applied Problems
EOM	-0.261***	-0.213***	-0.265***
ECOM	-0.139*	-0.143*	-0.187**
RTC	-0.160**	-0.139*	-0.133*
RTI	-0.192**	-0.193**	-0.171*
SDRTC	-0.112+	-0.134*	-0.103+
SDRTI	-0.048	-0.043	-0.034
d-prime	0.314***	0.287***	0.337***
PEA	0.295***	0.246***	0.286***
PES	0.083	0.033	0.065
Age	0.519***	0.450***	0.443***
Letter Word ID		0.829***	0.693***
Passage Comprehension	0.829***		0.588***
Applied Problems	0.693***	0.588***	

Note: N=287 for all variables.

****p*<.001, ***p*<.01, **p*<.05, +*p*<.10

Estimating Cognitive Control Construct

A series of models were run to estimate one-factor and two-factor CFAs of cognitive control. Model fit criterion included chi-square (p>.05), comparative fit index (CFI, >0.900), root mean square approximation (RMSEA, <.05), and standardized room mean square residual (SRMR, <.05). One and two-factor models were compared using AIC criterion. For all models, variables and co-variances were removed/added one-at-a-time to increase goodness of fit. Variables were removed if they demonstrated large negative residual variances and lambda factor loadings -1> λ>1 or fell below +/- 0.100 for lambda factor loadings (p>.05). Covariances were then added in incrementally, using the modindices function in lavaan, while also accounting for theoretically relevant residual shared variances (e.g., reaction time variables likely had shared variance relating to motor skills that would not be explained by cognitive control construct) in order to obtain a well-fitting model. Covariances were removed if they were not statistically significant (p>.05) unless retention of covariance in the model increased goodness-of-fit.

Observed Behavioral Measures of Cognitive Control. For the one-factor CFA, Model 1 included all nine behavioral variables. Details of the one-factor and two-factor model building can be found in the Appendix B. The final one-factor model is presented below in Table 13 and Figure 2. The final model (χ^2=2.11, *df*=3, *CFI*=1, *AIC*=3483.04, *RMSEA*=.000, *SRMR*=.009) included 5 indicator variables (EOM, RTC, SDRTC, SDRTI, PEA) and two covariances (EOM~~PEA and RTC~~PEA). The standard deviation variables (SDRTC, SDRTI) best fit the cognitive control latent construct (λ_{SDRTC}=0.870, λ_{SDRTI}=0.874). The latent construct explained the variance in EOM (λ=0.508) moderately well and explained the least variance in RTC (λ=0.390) and PEA measures (λ=-0.310).

Table 13

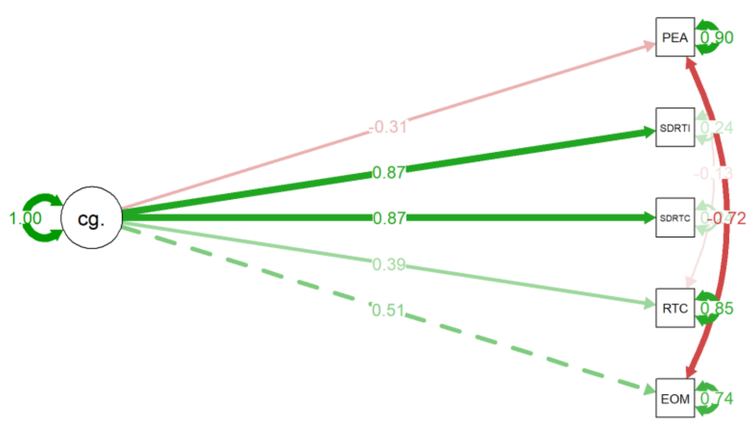
One-factor CFA for Observed Behavioral Indicator Variables

	Latent Estimate	Std. Error	Z-Value	Variance Estimate	Std. Error	Z-Value	λ
EOM	1.000			0.740	0.096	7.71***	0.508
RTC	0.768	0.130	5.91***	0.845	0.066	12.84***	0.390
SDRTC	1.713	0.197	8.71***	0.243	0.051	4.73***	0.870
SDRTI	1.720	0.203	8.45***	0.236	0.059	4.01***	0.874
PEA	-0.611	0.093	-6.58***	0.899	0.092	9.80**	-0.310
Cognitive Control				0.257	0.057	4.53***	
Covariances							
EOM~~PEA				-0.588	0.081	-7.26***	
RTC~~PEA				-0.109	0.041	-2.64***	

***p<.001, **p<.01, *p<.05, +p<.10

Figure 2

Cognitive Control CFA: One-factor Solution for Observed Indicator Variables

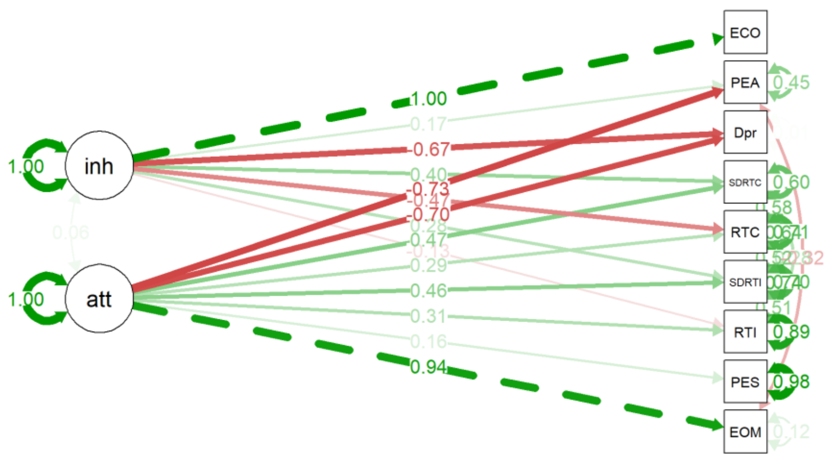


Note: The anchor variable, EOM is indicated with a dotted line. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

For the two-factor CFA, variables were assigned to “attention” and “inhibition” constructs, given that the Go/No-Go task is thought to tap these cognitive control skills. Attention was anchored on EOM and inhibition was anchored on ECOM, as these represent performance on Go and No-Go trials which theoretically signal continued attention versus response inhibition skills.

Figure 3

Cognitive Control CFA: Two-factor Solution for Observed Indicator Variables



Note: Anchor variables for each construct are indicated with dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

The final two-factor model is presented in Table 14 and Figure 3. The final model (χ^2=21.28, *df*=14, *CFI*=0.996, *AIC*=5396.45, *RMSEA*=.043, *SRMR*=.028) included all nine behavioral indicator variables such that attention significantly explained variance in eight variables (ECOM was not added as an indicator to this construct) and inhibition explained

Table 14

Cognitive Control CFA: Two-factor Solution for Observed Behavioral Indicators

	Latent Estimate Attention (*SE*)	Latent Estimate Inhibition (*SE*)	Variance Estimate	Std. Error	Z-Value	Attention λ	Inhibition λ
EOM	1.000	-	0.123	0.037	3.32***	0.936	
ECOM	-	1.000	0.000				1.000
RTI	0.330(0.08)	-0.134(0.06)	0.888	0.077	11.57***	0.309	-0.134
SDRTI	0.487(0.06)	0.280(0.06)	0.696	0.053	13.24***	0.456	0.280
RTC	0.311(0.07)	-0.474(0.05)	0.704	0.061	11.48***	0.291	0.474
SDRTC	0.502(0.06)	0.398(0.06)	0.597	0.057	10.51***	0.470	0.398
d-prime	-0.750(0.04)	-0.665(0.02)	0.012	0.018	0.63	-0.702	-0.665
PEA	-0.779(0.04)	0.172(0.04)	0.451	0.047	9.56***	-0.729	0.172
PES	0.167(0.06)	-	0.972	0.073	8.64***	0.156	
Attention			0.873	0.101	8.64***		
Inhibition			0.997	0.071	14.068**		
Covariances							
Attention~~Inhibition			0.053	0.057	0.93		
SDRTC~~SDRTI			0.415	0.049	8.51***		
RTC~~SDRTC			0.374	0.050	7.50***		
RTC~~SDRTI			0.366	0.050	7.37***		
RTI~~RTC			0.586	0.058	10.09***		
RTI~~SDRTC			0.207	0.048	4.32***		
RTI~~SDRTI			0.401	0.053	7.55***		
EOM~~PEA			-0.075	0.032	-2.34***		

***p<.001, **p<.01, *p<.05, +p<.10

variance in seven variables (PES did not load onto this construct). The two-factors were not correlated ($_{\text{standardized}}r$=0.057, p=0.352) and explained low to high variance in indicator variables.

Overall, the model best explained EOM and d-prime indicators and least explained RTC, RTI, and PES indicators, which had standardized error variances of 0.123, 0.012, 0.891, 0.706, 0.976 respectively. The attention construct best explained EOM (λ=0.936), d-prime (λ=-0.702), and PEA (λ=-0.729). Whereas, inhibition best explained variance in ECOM (λ=1.000) and d-prime (λ=-0.665) indicators. The model moderately explained SDRTC, SDRTI, and PEA; standardized error variances of 0.599, 0.699, and 0.453 respectively. Notably, in both the one- and two-factor CFAs, constructs should be interpreted as lower cognitive skills, given the directionality of the factor loadings.

A post-hoc exploratory EFA analysis was run using the factanal function in r to determine how many factors might best explain all nine behavioral variables (R Core Team, 2017). Results indicated that even five factors were not sufficient to explain variance in the behavioral measures (χ^2=3.89, p=.049). No indicator variable loaded onto all factors. Factor 1 best described variance in EOM (λ=0.935), PEA (λ=-0.755), and d-prime (λ=-0.543). Factor 2 best accounted for variance in SDRTC (λ=0.891) and SDRTI (λ=0.741). Factor 3 accounted for the most variance in ECOM (λ=0.944) and Factor 4 accounted for the most variance in RTI (λ=0.957) and RTC (λ=0.635). Factor 5 best described variance in RTC (λ=0.551). No factor explained variance in PES very well, such that PES demonstrated the most unexplained variance (λ=0.958). Results are presented in Table 15.

Table 15

Exploratory EFA Model Results: Observed Behavioral Indicators

	Uniqueness	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
EOM	0.031	0.935	0.274	0.104		
ECOM	0.005	-0.112	0.262	0.944	-0.122	
RTC	0.005	0.102	0.312	-0.425	0.635	0.551
RTI	0.005	0.130	0.244		0.957	
SDRTC	0.086	0.181	0.891	0.248		0.143
SDRTI	0.272	0.199	0.741	0.181	0.325	
PES	0.958		0.171			
PEA	0.391	-0.755	-0.149		-0.104	
d-prime	0.052	-0.543	-0.412	-0.701		-0.149
SS loadings		1.860	1.857	1.677	1.466	0.336
Proportion of Variance		0.207	0.206	0.186	0.163	0.037
Cumulative Variance		0.207	0.413	0.599	0.762	0.799

Adjusted Measures of Behavioral Performance. The final one-factor model is presented below in Table 16 and Figure 4. The final model (χ^2=4.03, *df*=4, *CFI*=1, *AIC*=4223.16, *RMSEA*=.005, *SRMR*=.020) included seven indicator variables (EOM, RTC, RTI, SDRTI, PEA, d-prime, and PES) and 11 covariances. Details about model testing can be found in Appendix B. Accuracy measures included best fit the cognitive control latent construct (λ_{EOM}=0.974, λ_{PEA}=1.00, $\lambda_{d\text{-}prime}$=-0.709). The latent construct explained relatively small amounts of variance in reaction time measures (λ_{RTC}=0.272, λ_{RTi}=0.321, λ_{SDRTI}=0.191, λ_{PES}=-0.181).

The final two-factor model is presented in Table 17 and Figure 5. The final model (χ^2=19.92, *df*=14, *CFI*=0.997, *AIC*=5011.78, *RMSEA*=.043, *SRMR*=.028) included all nine behavioral indicator variables such that attention significantly explained variance in seven variables (SDRTC did not load onto this construct; ECOM was not examined as an indicator variable of attention) and inhibition explained variance in eight variables (EOM was not examined as an indicator variable of inhibition). The two-factors were not correlated ($_{standardized}r$=0.049, p=0.386) and explained low to high variance in indicator variables.

Table 16

One-factor CFA for Predicted Behavioral Indicator Variables

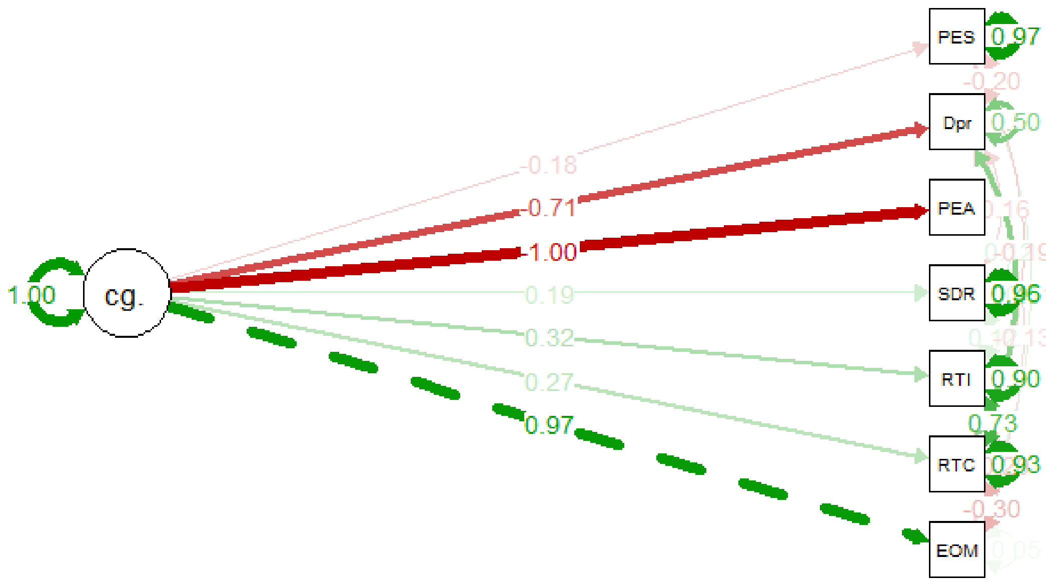
	Latent Estimate	Std. Error	Z-Value	Variance Estimate	Std. Error	Z-Value	λ
EOM	1.000			0.052	0.005	10.49***	0.974
RTC	0.277	0.064	4.33***	0.907	0.066	13.04***	0.272
RTI	0.330	0.065	5.09***	0.895	0.051	13.11***	0.321
SDRTI	0.196	0.068	2.90***	0.960	0.059	12.97***	0.191
PEA	-1.027	0.021	-49.89***	0.000			-1.00
d-prime	-0.727	0.042	-17.15***	0.493		10.73***	-0.709
PES	-0.186	0.059	-3.15***	0.964		13.93***	-0.181
Cognitive Control				0.945	0.057	9.59***	
Covariances							
RTC~~RTI				0.662	0.060	10.98***	
RTC~~d-prime				0.298	0.042	7.15***	
EOM~~RTC				-0.064	0.014	-4.49***	
EOM~~RTI				-0.042	0.014	-2.92**	
RTC~~PES				-0.175	0.038	-4.64***	
d-prime~~PES				-0.140	0.043	-3.24**	
RTI~~SDRTI				0.111	0.038	2.90**	
SDRTI~~d-prime				-0.108	0.038	-2.88**	
RTI~~d-prime				0.078	0.041	1.93+	
EOM~~d-prime				-0.021	0.010	2.24*	
EOM~~PES				0.025	0.012	2.06*	

***p<.001, **p<.01, *p<.05, +p<.10

Overall, the model best explained EOM, PEA, and d-prime indicators and least explained SDRTC, PES, SDRTI, and RTI indicators, which had standardized error variances of 0.047, 0.004, 0.031, 0.983, 0.924, 0.924, and 0.876 respectively. The attention construct best explained EOM (λ=0.976), d-prime (λ=-0.697), and PEA (λ=-0.999). Whereas, inhibition best explained variance in ECOM (λ=1.000) and d-prime (λ=-0.660) indicators.

Figure 4

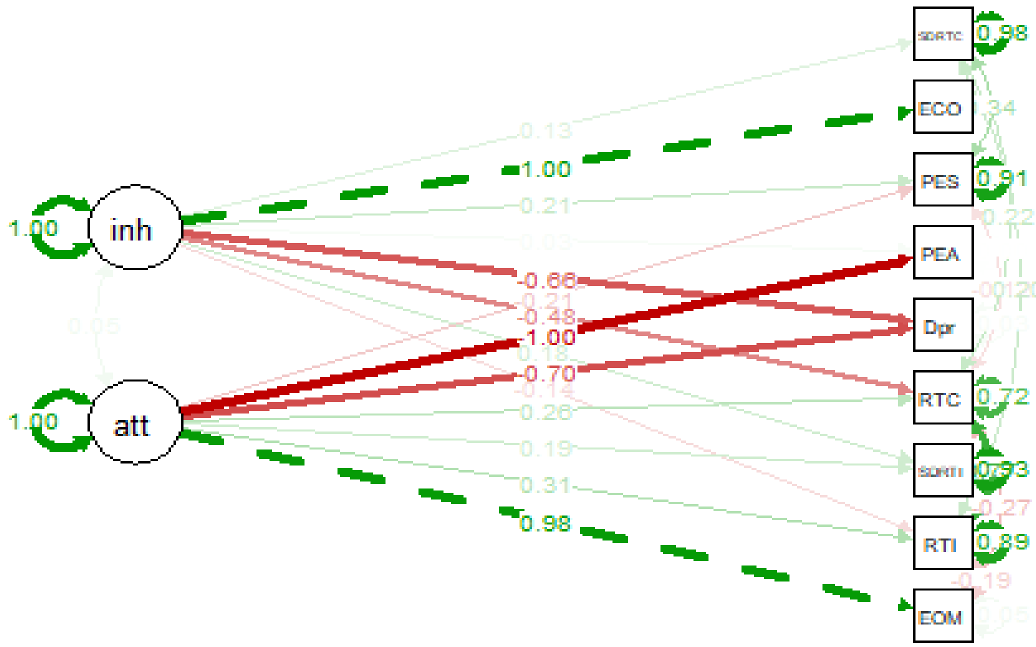
Cognitive Control CFA: One-factor Solution for Predicted Indicator Variables



Note: The anchor variable, EOM, is indicated with by dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Figure 5

Cognitive Control CFA: Two-factor Solution for Predicted Indicator Variables



Note: The anchor variables are indicated with by dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Table 17

Cognitive Control CFA: Two-factor Solution for Predicted Behavioral Indicators

	Latent Estimate Attention (*SE*)	Latent Estimate Inhibition (*SE*)	Variance Estimate	Std. Error	Z-Value	Attention λ	Inhibition λ
EOM	1.000	-	0.047	0.004	10.68***	0.976	
ECOM	-	1.000	0.000				1.000
RTI	0.313(0.07)	-0.139(0.06)	0.876	0.068	12.81***	0.308	-0.140
SDRTI	0.193(0.07)	0.184(0.06)	0.924	0.071	13.04***	0.188	0.184
RTC	0.267(0.06)	-0.474(0.05)	0.700	0.055	12.63***	0.263	0.478
d-prime	-0.715(0.02)	-0.660(0.01)	0.031	0.005	6.45***	-0.697	-0.660
PEA	-1.024(0.02)	0.033(0.01)	0.004	0.066	0.800	0.999	0.033
PES	-0.220(0.06)	0.213(0.06)	0.924	0.066	14.04***	-0.214	0.212
SDRTC		0.127(0.06)	0.983	0.077	12.71***		0.126
Attention			0.949	0.099	9.58***		
Inhibition			0.997	0.070	14.25***		
Covariances							
Attention~~Inhibition			0.049	0.057	0.867		
RTI~~RTC			0.607	0.055	11.13***		
PES~~SDRTC			0.320	0.057	5.64***		
RTC~~PES			-0.097	0.032	-3.00**		
EOM~~RTC			-0.050	0.012	-4.05***		
EOM~~RTI			-0.038	0.014	-2.69**		
RTC~~SDRTC			0.182	0.053	3.46**		
RTI~~SDRTC			0.187	0.059	3.17**		

***p<.001, **p<.01, *p<.05, +p<.10

Estimating Academic Construct

Results of the academic CFA are presented in Table 18 and Figure 6. Because there were only three indicator variables for academic skills, the model was saturated at step 1.

Table 18

WJ CFA: One-factor Solution

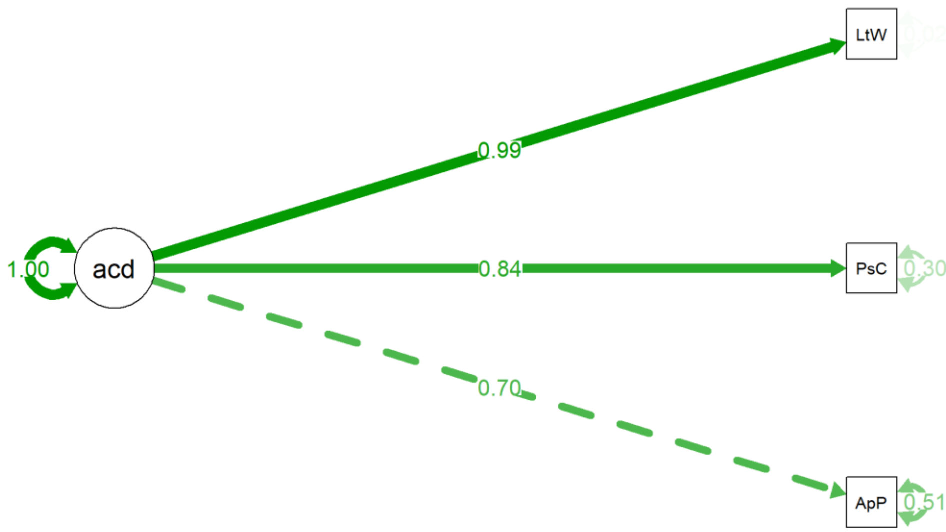
	Latent Estimate	Std. Error	Z-Value	Variance Estimate	Std. Error	Z-Value	λ
Applied Problems	1.000			0.507	0.047	10.84***	0.701
Passage Comprehension	1.196	0.086	13.88***	0.295	0.038	7.85***	0.839
Letter Word ID	1.409	0.102	13.81***	0.024	0.039	0.61	0.988
Academic Skills				0.490	0.074	6.59***	

****p*<.001, ***p*<.01, **p*<.05, +*p*<.10

Performance on the letter word ID subscale was most highly correlated with the academic skills latent factor (λ=0.988). Both passage comprehension (λ=0.839) and applied problems (λ=0.701) were moderate to highly correlated. Standardized residual variances, indicate that the latent factor explained 51.2% of the variance in applied problems, 70.4% of the variance in passage compression, and 97.6% of the variance in letter word ID. All subscales had positive factor loadings, indicating positive relations between the academic latent factor and both literacy and math observed subscales. A covariance between letter word ID and passage comprehension was tested in the full SEM model, as covariances were not able to be estimated in this saturated model. This was decided due to an expected shared residual variance among literacy subscales, separate from the math subscale.

Figure 6

Academic Skills CFA



Note: The anchor variable, applied problems, is indicated with a dotted line. Green lines are associated with positive factor loadings.

Estimating Relations between Cognition and Academics

For both the observed and predicted CFA models of cognitive control, the one-factor solution had a better model fit than the two-factor model as indicated by a lower *AIC* value. Therefore, the one-factor construct of cognitive control was regressed onto the one-factor construct of academic skills in SEM analyses. A traditional multiple regression model using observed behavioral measures as predictors of each academic construct was first run to compare SEM results with regression analyses more frequently reported in the literature. Linear regressions weight each predictor variable the same, as opposed to SEM models which allow for behavioral indicator variables to have different weights onto the cognitive control factor that is then used to predict variance in the academic factor. The academic factor in the SEM models below then adjusts the weight of variance explained in each academic indicator variable, considering the variance predicted by the cognitive control construct. Results are outlined below by analytic technique.

Traditional Multiple Regression Approach. Linear regression models predicting one of the three academic subscales was run to estimate relations between observed behavioral measures and academic skills. School was included as a covariate predictor. Results are presented in Table 19. Only one behavioral measure from the Go/No-Go was significantly related to each academic outcome. Academic outcomes were either related to more positive d-prime, indicating better performance, or shorter average Go trial reaction times. The cognitive model predicted approximately 1.90% of variance in the applied problems subscale, such that a 1 SD increase in RTC (longer reaction times) was related to a 0.130 standard deviation decrease in math skills. The d-prime measure helped predict 10.20% and 8.10% of variance in the letter word ID and passage comprehension subsets, respectively. For every 1 SD increase in d-prime, we would expect to see a 0.316 standard deviation increase in letter word ID and a 0.290 standard deviation increase in passage comprehension.

Table 19

Multiple Regression Models Predicting Academic Skills

	Applied Problems	Letter Word ID	Passage Comprehension
(Intercept)	-6.224e^$^{-03}$	-7.351 e^$^{-03}$	-4.887 e^$^{-01}$
RTC	-1.298e^$^{-01}$		
d-prime		3.155 e^$^{-01}$	2.898 e^$^{-01***}$
School	1.647e^$^{-04}$	1.945 e^$^{-04}$	1.293 e^$^{-04}$
Model Statistic	3.827*	17.23***	13.55***
R^2	0.026	0.108	0.087
Adjusted R^2	0.019	0.102	0.081

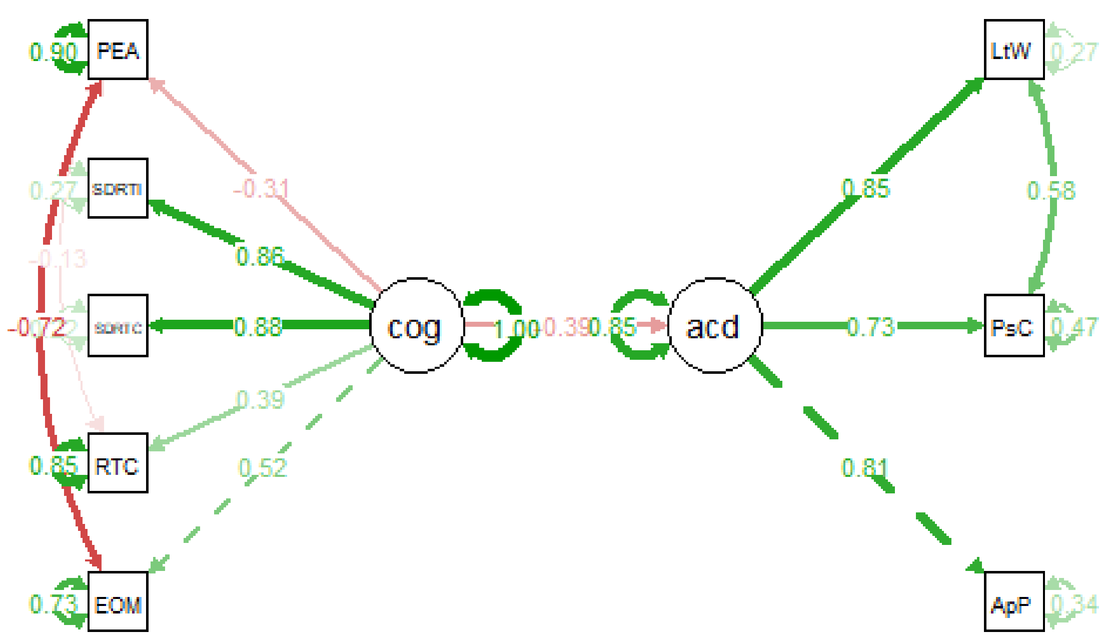
***p<.001, **p<.01, *p<.05, +p<.10

SEM model with Observed Indicators. Using the one-factor solutions for the cognitive control and academic skills constructs above, both one-level and two-level SEM models were assessed for goodness of fit. For the 2-level model, academic and cognitive variables were added as indicators for a between school factor of variance. However, no 2-level models reached convergence. Therefore, school was added as a covariate predictor for the cognitive control and academic latent factors. However, school was not a significant predictor for either the cognitive control (β=0.002, p=.938) or academic (β=0.021, p=.594) factors and were thus removed from the model estimation.

Results (χ^2=16.01, *df*=16, *CFI*=1.00, *AIC*=5385.84, *RMSEA*=.002, *SRMR*=.039) are outlined in Table 20 and Figure 7. Results indicate a moderate negative relation between cognitive control and academics (β=-0.612, p<.001). However, given the factor loadings, the cognitive control construct reflects poor cognitive control skills – such that higher factor loadings are associated with greater percentage of errors and longer and more variable reaction times. Therefore, better cognitive control skills (negative Z-scores), would be associated with higher academic skills (positive Z-scores).

Figure 7

Observed Behavioral Measures Predict Academic Skills



Note: Anchor variables for each construct are indicated with dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Table 20

Behavioral Indicators SEM: Solution for Observed Indicators

	Latent Estimate	Std.Error	Z-Score	Variance Estimate	Std. Error	Z-Value	Cognitive λ	Academic λ
EOM	1.00			0.732	0.065	11.35***	0.515	
RTC	0.753	0.138	5.44***	0.847	0.073	11.67***	0.388	
SDRTC	1.717	0.198	8.66***	0.217	0.049	4.44***	0.884	
SDRTI	1.663	0.192	8.68***	0.266	0.048	5.52***	0.856	
PEA	-0.611	0.083	-7.33***	0.897	0.075	11.91***	-0.315	
Applied Problems	1.00			0.339	0.100	3.40***		0.812
Passage Comprehension	0.897	0.145	6.18***	0.467	0.092	5.10***		0.729
Letter Word ID	1.049	0.161	6.52***	0.273	0.108	2.53*		0.852
Cognitive Control				0.264	0.060	4.40***		
Academic Skills				0.559	0.105	5.32***		
Regressions								
Academic ~ Cognitive	-0.612	0.133	-4.69***					
Covariances				Estimate	Std. Error	Z-Value		
EOM~~PEA				-0.584	0.061	-9.63***		
RTC~~PEA				-0.110	0.037	-2.99**		
Passage Comprehension~~ Letter Word ID				0.207	0.095	2.19*		

****p*<.001, ***p*<.01, **p*<.05, +*p*<.10

Table 21

Behavioral Indicators SEM: Solution for Adjusted Indicators

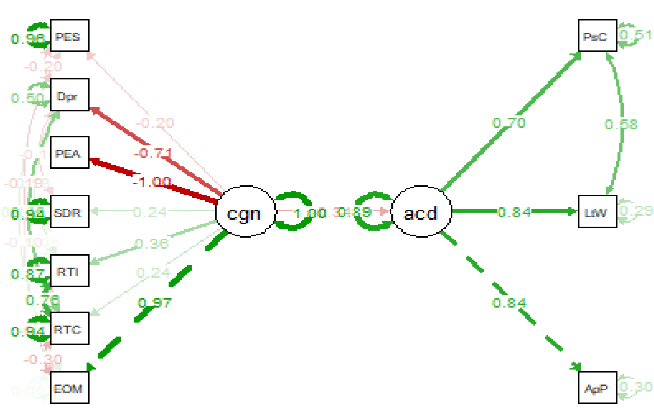
	Latent Estimate	Std.Error	Z-Score	Variance Estimate	Std. Error	Z-Value	Cognitive λ	Academic λ
EOM	1.00			0.046	0.004	11.09***	0.974	
RTC	0.255	0.066	3.87***	0.895	0.079	11.33***	0.244	
RTI	0.379	0.064	5.97***	0.830	0.075	11.11***	0.361	
SDRTI	0.257	0.066	3.87***	0.936	0.084	11.09***	0.240	
PEA	-1.037	0.015	-67.81***	0.000			-1.00	
d-prime	-0.757	0.048	-15.76***	0.488	0.044	11.16***	-0.710	
PES	-0.214	0.066	-3.24**	0.919	0.083	11.09***	-0.204	
Applied Problems	1.00			0.307	0.121	2.55***		0.836
Letter Word ID	0.963	0.166	5.82***	0.270	0.112	2.42****		0.843
Passage Comprehension	0.815	0.149	5.48***	0.499	0.095	5.24**		0.699
Cognitive Control				0.867	0.082	10.54***		
Academic Skills				0.634	0.131	4.85***		
Regressions								
Academic ~ Cognitive	-0.308	0.067	-4.62***					
Covariances								
RTC~~RTI				0.651	0.068	9.58***		
RTC~~d-prime				0.301	0.045	6.65***		
EOM~~RTC				-0.061	0.013	-4.51***		
EOM~~RTI				-0.030	0.013	-2.41*		
RTC~~PES				-0.172	0.039	-4.37***		
d-prime~~PES				-0.133	0.043	-3.13**		
RTI~~SDRTI				0.046	0.036	1.27		
SDRTI~~d-prime				-0.103	0.039	-2.66**		
RTI~~d-prime				0.085	0.040	2.13*		
EOM~~d-prime				-0.015	0.010	-1.62		
Covariances (cont.)				Variance Estimate	Std. Error	Z-Value		
EOM~~PES				0.024	0.013	1.86+		
Passage Comprehension~~ Letter Word ID				0.213	0.098	2.18*		

***p<.001, **p<.01, *p<.05, +p<.10

SEM model with Adjusted Indicators. Similar to the observed model, both one-level and two-level SEM models using one-factor solutions for cognitive control and academic factors were assessed for goodness of fit. For the 2-level model, academic and cognitive variables were added as indicators for a between school factor of variance. However, no 2-level models reached convergence. Therefore, school was added as a covariate predictor for the cognitive control and academic latent factors. Yet, again, school was not a significant predictor for either the cognitive control (β=-0.050, p=.166) or academic (β=0.004, p=.897) factors and was removed from the model estimation. Results (χ2=29.53, *df*=23, *CFI*=0.996, *AIC*=5150.51, *RMSEA*=.034, *SRMR*=.037) are outlined in Figure 8 and Table 21. Results indicate a similar but weaker negative relation between cognitive control and academics (β=-0.308, p<.001) than presented in the observed model.

Figure 8

Adjusted Behavioral Measures Predict Academic Skills



Note: Anchor variables for each construct are indicated with dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Aim 3: Examining the Contribution of Neural Measures

It is theorized that neural components reflect covert cognitive processes not fully captured by observations of behavior, such as inhibitory and self-monitoring processes that occur in the absence of behavior. In order to examine the contribution of neural measures to our ability to predict variance in academic achievement, ERP components were added to models investigated in Aim 2. Specifically, observed measures of the ERN, CRN, Pe Correct, and Pe Error were examined as additional predictors in the final traditional multiple regression models and additional indicators in the Observed CFA and SEM models identified in Aim 2. This allowed for investigation both of how neural measures load onto a factor of behavioral cognitive control and whether the addition of neural indicators changes the estimated relation between cognitive control and academic skills. Next, adjusted scores for neural measures accounting for potential measurement bias from differences in child age and data quality were estimated using linear regression models. Predicted scores of ERP components were then calculated from these regression models. These predicted ERP measures were added as indicators in the Adjusted Measures CFA and SEM Models identified in Aim 2. Results are organized by analytic technique.

Descriptive statistics for observed neural measures are presented in Table 22. Value for ERP measures varied among participants, more than expected. CRN and ERN components are defined as negative deflections in EEG activity. However, some children had positive values of the ERN and CRN. A paired t-test indicated that children did demonstrate the expected relation between CRN and ERN measures, such that the ERN was more negative than the CRN ($t(245)=-20.42$, $p<.001$). Similarly, the Pe component is defined as a positive deflection in EEG activity, yet some children showed negative values for the Pe. Children did demonstrate expected relations between the Pe for error and correct trials, such that the Pe was more positive following error than correct trials ($t(245)=20.13$, $p<.001$). Discrepancies in expected values for ERP components may have been a result of the time windows used to average ERP components or other sources of variation, unable to be observed in this secondary analysis. All data met the 6-epoch inclusion criterion.

Table 22

Descriptive Statistics for Observed ERP Components

Variable	Min	Mean	Max	SD
CRN	-14.64	5.056016	28.54	7.044793
ERN	-32.6	-6.33785	35.15	8.123426
Pe Correct	-38.09	3.665976	37.76	9.517027
Pe Error	-31.28	19.72069	80.86	14.18503
Correct Epochs	35	165.50	235	45.14
Error Epochs	6	23.95	69	12.54

Note: ERP measures are reported in μV. Correct and Error Epochs are measures of data quality. Specifically, they are the count of usable segments of EEG data for Correct Go and Error No-Go trials.

To retain outliers in the sample and adjust for skewness in the sample distributions ERP measures were winsorized at a 10% level and then standardized to a Z-score. Correlations between winsorized and standardized scores for observed neural, behavioral, and academic measures are presented in Table 23. Neural components were positively related to one another, though not strongly. The CRN was most related to ECOM ($r(246)=-0.53$, $p<.001$), d-prime ($r(246)=0.47$, $p<.001$) measures. Similarly, Pe Error had the strongest correlations with d-prime ($r(246)=0.43$, $p<.001$), ECOM ($r(246)=-0.36$, $p<.001$), and SDRTI ($r(246)=-0.37$, $p<.001$) measures. Neither the ERN or Pe Correct were related to any behavioral or academic measure above 0.200. The CRN was weakly related to performance on the applied problems subscale (*r*(246)=0.20, *p*=.002). Likewise, the Pe Error was only weakly associated with literacy subscales of the letter word ID (*r*(246)=0.18, *p*=.005) and passage comprehension (*r*(246)=0.15, *p*=.018).

Table 23

Correlations among Observed Neural, Behavioral, and Academic Measures

	CRN	ERN	Pe Correct	Pe Error
Age	-0.071	0.046	-0.074	0.094
EOM	-0.156*	0.019	0.022	-0.243***
ECOM	-0.534***	0.124+	-0.066	-0.360***
RTC	0.352***	0.095	0.156*	0.055
RTI	0.198**	0.134*	0.092	-0.127*
SDRTC	-0.387***	0.032	0.043	-0.324***
SDRTI	-0.275***	-0.021	0.019	-0.365***
d-prime	0.474***	-0.110+	0.026	0.432***
PEA	-0.007	-0.084	-0.118*	-0.009
PES	-0.065	0.016	0.036	-0.077
Applied Problems	0.197**	0.033	0.026	0.119+
Letter Word ID	0.150*	0.118+	-0.015	0.177**
Passage Comprehension	0.109+	0.052	-0.038	0.151*
CRN		0.326***	0.219**	0.382***
ERN	0.326***		0.104	0.221***
Pe Correct	0.219**	0.104		0.451***
Pe Error	0.382***	0.221***	0.451***	

***p<.001, **p<.01, *p<.05, +p<.10

Traditional Multiple Regression Approach

Observed measures of the ERN, CRN, Pe Correct, and Pe Error were examined as additional predictors of academic skills in the traditional multiple regression models identified in Aim 2 (see Table 19). Table 24 describes findings from these analyses. No neural components were significantly related to passage comprehension. With the addition of the CRN predictor, cognitive control measures (RTC, CRN) explained 6.3% of the variance in applied problems which was an increase from the regression model with only behavioral predictors (1.9% explained variance). For every 1 SD increase (less negative) in CRN amplitude, applied problems is expected to increase by 0.026 standard deviations. Similarly, the ERN increased the variance explained in the letter word ID subscale from 10.2% (behavioral predictors only) to 12.2%. A 1 SD increase (less negative) ERN amplitude is expected to result in a 0.158 standard deviation increase in letter word ID skills.

Table 24

Multiple Regression Models Predicting Academic Skills

	Applied Problems	Letter Word ID
(Intercept)	-2.393e^-02	-4.724 e^-02
RTC	-1.704e^-01	
d-prime		3.181 e^-01
CRN	2.63e^-01	
ERN		1.577 e^-01
School	1.83e^-04	2.136 e^-04
Model Statistic	6.53***	12.30***
R^2	0.075	0.132
Adjusted R^2	0.063	0.122

****p*<.001, ***p*<.01, **p*<.05, +*p*<.10

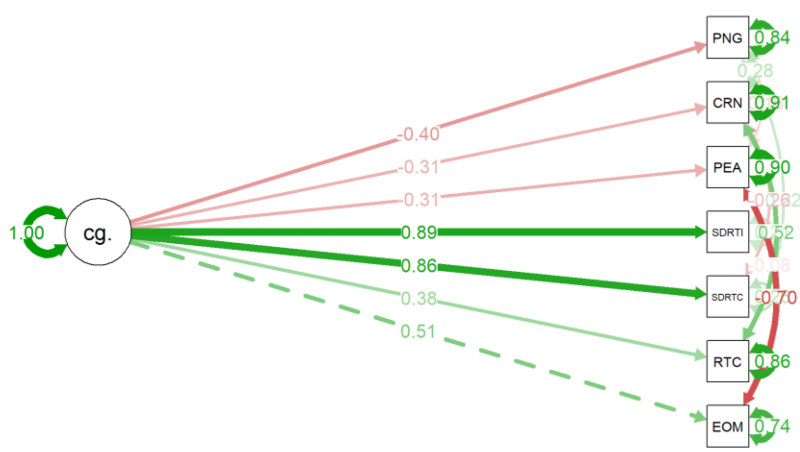
Observed Measures SEM Model

Estimating Broader One-Factor CFA. Observed ERP measures were added as additional indicators to the finalized one-factor CFA model identified in Aim 2 (see Table 13). The ERN and Pe Correct did not load well onto the factor, and so they were removed from the model. Next, covariances were examined incrementally to try and reach a better model fit. Details of the one-factor CFA model testing can be found in the Appendix B. The final one-factor model is presented below in Figure 9 and Table 25.

The final model (χ^2=8.60, *df*=7, *CFI*=0.998, *AIC*=4126.19 *RMSEA*=.031, *SRMR*=.025) included an additional two indicator variables (CRN, Pe Error) and five covariances from the behavioral only CFA. Table 26 details the differences between the behavior-only and the one-factor CFA that includes neural indicators. Results indicate that there was little difference in factor or residual variance between models. Similarly, there were only small changes in factor loadings. However, standard error variances decreased largely from the behavior only model to the model with neural indicator variables added.

Figure 9

ERP and Behavioral Indicators of Cognitive Control: CFA One-factor Solution



Note: The anchor variable, EOM is indicated with a dotted line. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Table 25

Observed Behavior and ERP CFA: One-factor Solution

	Latent Estimate	Std. Error	Z-Value	Variance Estimate	Std. Error	Z-Value	λ
EOM	1.00			0.675	0.098	6.92***	0.507
RTC	0.762	0.147	5.20***	0.805	0.069	11.66***	0.380
SDRTC	1.750	0.217	8.05***	0.254	0.049	5.20***	0.859
SDRTI	1.780	0.229	7.76***	0.198	0.051	3.90***	0.888
PEA	-0.602	0.100	-5.99***	0.776	0.094	8.29***	-0.313
CRN	-0.636	0.148	-4.30**	0.907	0.077	11.80***	-0.307
Pe Error	-0.828	0.158	-5.24***	0.820	0.077	10.62***	-0.404
Cognitive Control				0.233	0.056	4.15***	
Covariances							
EOM~~PEA				-0.505	0.081	-6.22***	
RTC~~CRN				0.442	0.058	7.59***	
PEA~~Pe Error				-0.126	0.040	-3.13**	
CRN~~Pe Error				0.238	0.057	4.16***	
RTC~~Pe Error				0.182	0.053	3.41**	
SDRTC~~CRN				-0.126	0.042	-2.99**	
RTC~~PEA				-0.067	0.033	-2.01*	

***p<.001, **p<.01, *p<.05, +p<.10

Table 26

One-Factor Cognitive Control CFA: Comparison between One-factor Solutions

	Behavior λ	Neural λ	Behavior Std. Error Variance	Neural Std. Error Variance
EOM	0.508	0.507	0.740	0.098
RTC	0.390	0.380	0.845	0.069
SDRTC	0.870	0.859	0.243	0.049
SDRTI	0.874	0.888	0.236	0.051
PEA	-0.310	-0.313	0.899	0.094
CRN		-0.307		0.077
Pe Error		-0.404		0.077
Cognitive Control	*unstandardized estimate*		0.257(0.057)***	0.233 (0.056)***
Covariances	Behavior	Neural		
EOM~~PEA	-0.588***	-0.505***		
RTC~~CRN		0.442***		
PEA~~Pe Error		-0.126***		
CRN~~Pe Error		0.238***		
RTC~~Pe Error		0.182**		
SDRTC~~CRN		-0.126**		
RTC~~PEA	-0.109***	-0.067*		

***p<.001, **p<.01, *p<.05, +p<.10

Table 27

Exploratory EFA Model Results

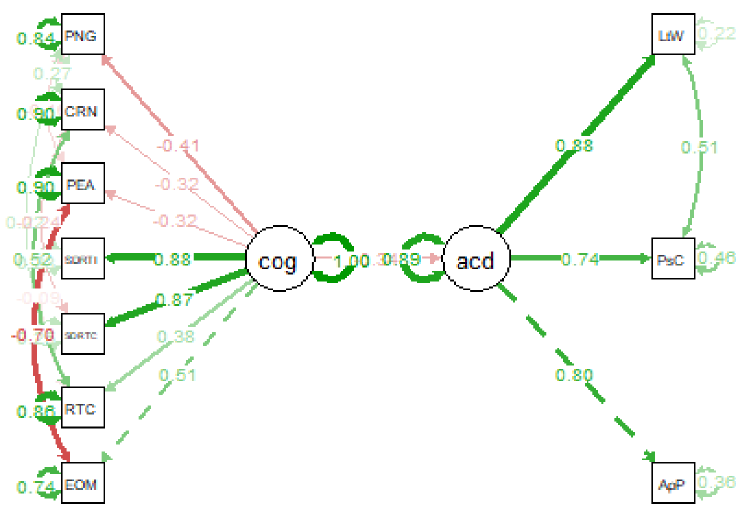
	Uniqueness	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
EOM	0.051	0.126	0.927	0.228	0.116			
ECOM	0.062	-0.921		0.228	-0.132			
RTC	0.005	0.456		0.327	0.744		0.102	0.323
RTI	0.077		0.207	0.181	0.912			
SDRTC	0.005	-0.243	0.197	0.935	0.146			
SDRTI	0.296	-0.255	0.221	0.643	0.378		-0.111	-0.137
PES	0.947		0.130	0.172				
PEA	0.385		-0.754	-0.127	-0.124		-0.110	
d-prime	0.005	0.741	-0.554	-0.328			0.133	
ERN	0.005					0.987	0.117	
CRN	0.454	0.477		-0.313	0.247	0.315	0.240	
Pe Correct	0.554						0.657	
Pe Error	0.298	0.305		-0.200	-0.111	0.165	0.722	
SS loadings		2.074	1.907	1.823	1.683	1.127	1.096	0.147
Proportion of Variance		0.160	0.147	0.140	0.129	0.087	0.084	0.011
Cumulative Variance		0.160	0.306	0.447	0.576	0.663	0.747	0.758

A post-hoc exploratory EFA analysis was run using the factanal function in r to determine how many factors might best explain all 13 measures that can be extracted from the Go/No-Go (R Core Team, 2017). Results indicated that seven factors were sufficient to explain variance in the Go/No-Go measures (χ^2=9.18, *p*=0.327). No indicator variable loaded onto all factors. Factor 1 best described variance in ECOM (λ=-0.921), and d-prime (λ=0.741) and moderately described variance in RTC (λ=0.456), CRN (λ=0.477), and Pe Error (λ=0.305). Fewer ECOM were related to slower reaction times, more attenuated (less negative) CRN values, greater (more positive) Pe Error values, and better d-prime ratios. Factor 2 best accounted for variance in EOM (λ=0.927) and PEA (λ=-0.754), with moderate variance accounted for in d-prime (λ=-0.554). Fewer EOM were related to greater post-error adjustments in accuracy and better d-prime ratios. Factor 3 accounted for the most variance in SDRTC (λ=0.935) and SDRTI (λ=0.643). Factor 4 accounted for the most variance in RTI (λ=0.912) and RTC (λ=0.744). Factor 5 best described variance in the ERN (λ=0.987). Notably, no behavioral variables loaded onto this factor. Factor 6 explained moderate variance in Pe values for correct (λ=0.657) and error (λ=0.722) trials. Only RTC (λ=0.323) and SDRTI (λ=-0.137) loaded weakly onto Factor 7. No factor explained variance in PES very well, such that PES demonstrated the most unexplained variance (λ=0.947). Results are presented in Table 27.

SEM model with Observed Measures. The one-factor solution for the observed behavior and ERP CFA (see Table 25) was regressed onto the academic skills CFA (see Table 18) to investigate whether the addition of neural indicator variables increase the variance explained in academic skills. Results of the model (χ^2=31.83, *df*=26, *CFI*=0.995, *AIC*=5762.86, *RMSEA*=.030, *SRMR*=.043) are outlined in Table 28 and Figure 10. Results indicate a moderate negative relation between cognitive control and academics (*β*=-0.557, *p*<.001). The strength of this association is slightly less than Aim 2 (β=-0.612, p<.001). While the addition of neural indicators did not increase the overall amount of variance explained in academic skills (Behavior = 44.1%, Behavior and ERP = 42.5%), results do indicate better predictions of literacy skills for passage comprehension (Behavior = 53.3%, Behavior and ERP = 55.5%) and letter word ID (Behavior = 72.7%, Behavior and ERP = 79.7%). Table 29 shows a comparison of factor loadings and standard error variances between models.

Figure 10

Observed SEM Model: Behavioral and Neural Indicators of Academic Skills



Note: Anchor variables for each construct are indicated with dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Table 28

Behavioral and ERP Indicators SEM: Solution for Observed Indicators

	Latent Estimate	Std. Error	Z-Score	Variance Estimate	Std. Error	Z-Value	Cognitive λ	Academic λ
EOM	1.00			0.668	0.063	10.53***	0.515	
RTC	0.751	0.152	4.96***	0.805	0.074	10.86***	0.380	
SDRTC	1.742	0.215	8.10***	0.239	0.046	5.18***	0.868	
SDRTI	1.729	0.213	8.11***	0.218	0.045	4.89***	0.876	
PEA	-0.601	0.090	-6.70***	0.773	0.070	11.07***	-0.318	
CRN	-0.649	0.157	-4.14**	0.899	0.082	10.99***	-0.318	
Pe Error	-0.819	0.157	-5.22***	0.819	0.076	10.82***	-0.406	
Applied Problems	1.00			0.373	0.120	3.11**		0.797
Passage Comprehension	0.902	0.175	5.14***	0.445	0.109	4.09***		0.737
Letter Word ID	1.060	0.196	5.40***	0.203	0.131	1.55		0.884
Cognitive Control				0.241	0.059	4.08***		
Academic Skills				0.575	0.125	4.61***		
Regressions								
Academic ~ Cognitive	-0.557	0.149	-3.73***					
Covariances								
EOM~~PEA				-0.500	0.057	-8.71***		
RTC~~CRN				0.444	0.061	7.34***		
PEA~~Pe Error				-0.126	0.037	-3.45***		
CRN~~Pe Error				0.235	0.055	4.29***		
RTC~~Pe Error				0.182	0.054	3.41**		
SDRTC~~CRN				-0.113	0.039	-2.87**		
RTC~~PEA				-0.067	0.031	-2.15*		
Letter Word~~Passage Comprehension				0.153	0.114	1.34		

***p<.001, **p<.01, *p<.05, +p<.10

Table 29

SEM: Comparison between Aim 2 and Aim 3 Models

	Aim 2		Aim 3		Aim 2	Aim 3
	Cognitive λ	Academic λ	Cognitive λ	Academic λ	Error Variance	Error Variance
EOM	0.515		0.515		0.732***	0.668***
RTC	0.388		0.380		0.847***	0.805***
SDRTC	0.884		0.868		0.217***	0.239***
SDRTI	0.856		0.876		0.266***	0.218***
PEA	-0.315		-0.318		0.897***	0.773***
CRN			-0.318			0.899***
Pe Error			-0.406			0.819***
Applied Problems		0.812		0.797	0.339***	0.373**
Passage Comprehension		0.729		0.737	0.467***	0.445***
Letter Word ID		0.852		0.884	0.273*	0.203
Cognitive Control					0.264***	0.241***
Academic Skills					0.559***	0.575***
Regressions	*B*	Std. Error	*β*	Std. Error		
Academic~Cognitive	-0.612***	0.133	-0.557***	0.149		
Covariances		Estimate		Estimate		
EOM~~PEA		-0.584***		-0.500		
RTC~~CRN				0.444***		
PEA~~Pe Error				-0.126***		
CRN~~Pe Error				0.235***		
RTC~~Pe Error				0.182**		
SDRTC~~CRN				-0.113**		
RTC~~PEA		-0.110**		-0.067*		
Letter Word~~Passage Comprehension		0.207*		0.153		

***p<.001, **p<.01, *p<.05, +p<.10

Adjusted Measures SEM Model

An Adjusted Measures SEM model including adjusted (predicted) measures of both behavioral and neural measures was run to examine whether accounting for trial features and child demographics in estimates of cognitive performance increases the relation between cognitive control and academic skills. Details on how ERP adjusted measures were calculated are detailed below in the Estimating Neural Measures section. Adjusted ERP measures were added as additional indicators to the finalized one-factor Adjusted Measure Model CFA identified in Aim 2 (see Table 16).

Estimating Neural Measures. Linear regression models were used to estimate adjusted scores for ERPs, accounting for potential measurement bias from differences in child age and data quality. Final models for predicting variance in ERP components can be found in Table 30. Age was not a significant predictor of any of the ERP components (p>.05). Thus, the final model, only included a standardized score of EEG quality as a predictor variable along with a fixed intercept estimate. Lower quality of EEG data was statistically significant and related to more positive values in the Pe Error (β=-2.92, p=.001) and associations with negative values in the ERN (β=0.85, p=.102) and CRN (β=0.73, p=.105) were also near trend levels.

Table 30

Predictors of ERP Amplitudes

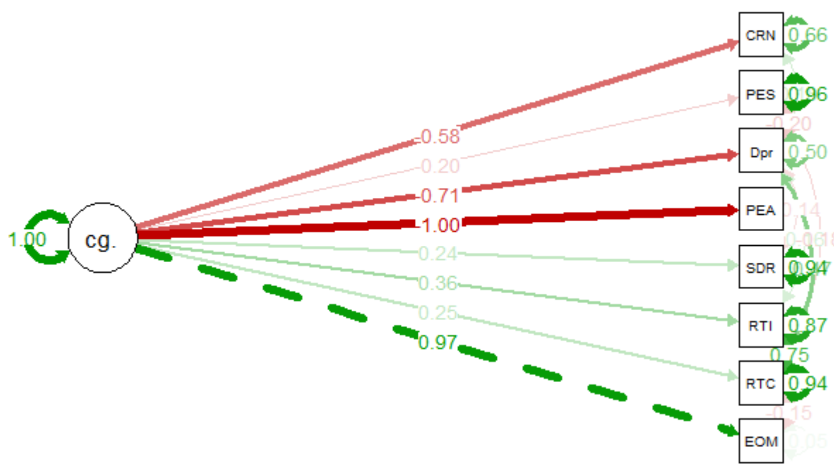
	CRN	ERN	Pe Correct	Pe Error
Intercept	5.056(0.45)***	-6.338(0.52)***	3.666(0.61)***	19.721(0.89)***
EEG Quality	0.730(0.49)	0.849(0.52)	0.281(0.61)	-2.92(0.89)**
Model Statistic	2.65	2.67	0.21	10.78**
Std. Error	7.021	8.095	9.53	13.91
R^2	0.011	0.011	0.001	0.042
Adjusted R^2	0.007	0.007	-0.003	0.038

***p<.001, **p<.01, *p<.05, +p<.10

Estimating Broader One-Factor CFA. In order to examine how ERPs loaded onto a behavioral cognitive control factor when cognitive measures were adjusted for potential measurement bias in child and trial features, adjusted (predicted) measures of ERP components were added to the Adjusted Measure CFA from Aim 2. The ERN and Pe components did not load well onto the factor, and so they were removed from the model. Next, covariances were examined incrementally to try and reach a better model fit. Details of the one-factor CFA model testing can be found in the Appendix B. The final one-factor model is presented below in Figure 11 and Table 31.

Figure 11

Predicted Measure One-factor CFA Model with Neural Indicators



Note: The anchor variable, EOM is indicated with a dotted line. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Table 31

Behavioral and ERP Indicators CFA: Solution for Predicted Indicators

	Latent Estimate	Std. Error	Z-Value	Variance Estimate	Std. Error	Z-Value	λ
EOM	1.000			0.046	0.005	9.29***	0.974
RTC	0.255	0.067	3.83***	0.855	0.069	12.43***	0.249
RTI	0.379	0.068	5.60***	0.829	0.068	12.20***	0.361
SDRTI	0.257	0.074	3.47**	0.936	0.080	11.20***	0.240
PEA	-1.037	0.023	-44.51***	0.000			-1.00
d-prime	-0.757	0.052	-14.67***	0.493	0.049	10.14***	-0.708
PES	-0.214	0.066	-3.27**	0.919	0.074	12.38***	-0.204
CRN	-0.622	0.049	-12.65***	0.661	0.065	10.24***	-0.580
Cognitive Control				0.867	0.103	8.38***	
Covariances							
RTC~~RTI				0.631	0.061	10.31***	
RTC~~d-prime				0.303	0.044	6.88***	
EOM~~RTC				-0.030	0.008	-3.81***	
RTC~~PES				-0.157	0.035	-4.47***	
d-prime~~PES				-0.138	0.044	-3.16**	
SDRTI~~d-prime				-0.093	0.038	-2.42*	
RTI~~d-prime				0.101	0.044	2.28*	
d-prime~~CRN				0.101	0.032	3.16**	

****p*<.001, ***p*<.01, **p*<.05, +*p*<.10

Table 32

Comparisons between One-factor Predicted CFAs

	Behavior λ	Neural λ	Behavior Error Variance	Neural Error Variance
EOM	0.974	0.974	0.052***	0.046***
RTC	0.272	0.249	0.907***	0.855***
RTI	0.321	0.361	0.895***	0.829***
SDRTI	0.191	0.240	0.960***	0.936***
PEA	-1.00	-1.00	0.000***	0.000***
d-prime	-0.709	-0.708	0.493***	0.493***
PES	-0.181	-0.204	0.964***	0.919***
CRN		-0.580		0.661***
Cognitive Control	*unstandardized variance estimate*		0.945***	0.867***
Covariances	Behavior	Neural		
RTC~~RTI	0.662***	0.631***		
RTC~~d-prime	0.298***	0.303***		
EOM~~RTC	-0.064***	-0.030***		
EOM~~RTI	-0.042**			
RTC~~PES	-0.175***	-0.157***		
d-prime~~PES	-0.140**	-0.138**		
RTI~~SDRTI	0.111***			
SDRTI~~d-prime	-0.108**	-0.093*		
RTI~~d-prime	0.078**	0.101*		
EOM~~d-prime	-0.021*			
EOM~~PES	0.025*			
d-prime~~CRN		0.101**		

***p<.001, **p<.01, *p<.05, +p<.10

The final model (χ^2=17.61, *df*=13, *CFI*=0.996, *AIC*=4110.99 *RMSEA*=.044, *SRMR*=.026) included an additional indicator variable (CRN) and reduced the number of covariances to 8 from the 11 included in the behavioral only CFA. Table 32 details the differences between the behavior-only and the one-factor CFA that includes neural indicators. Results indicate that there was little difference in factor loadings or indicator error variance between models. When the CRN was included as an indicator variable, the unstandardized error variance of the cognitive control construct decreased from 0.945 to 0.867, indicating that the CRN decreased the amount of dispersion in the latent factor.

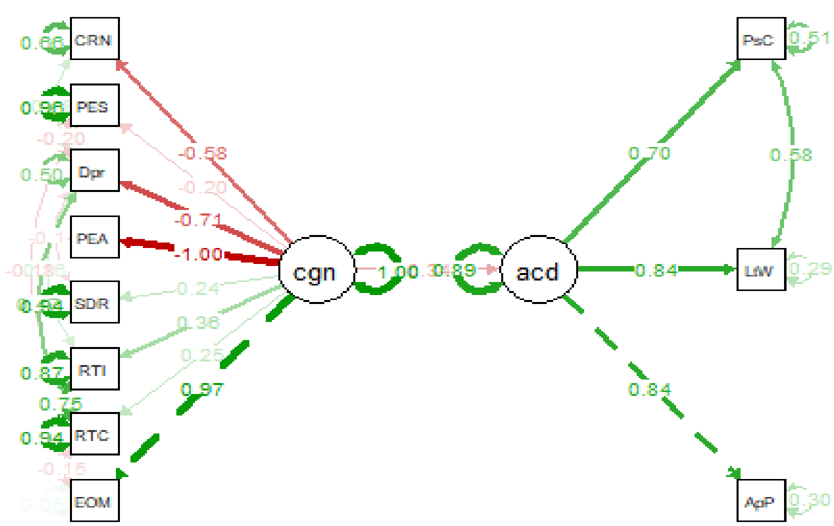
SEM model with Adjusted Measures. Results of the model (χ^2=48.05, *df*=35, *CFI*=0.993, *AIC*=5744.69 *RMSEA*=.039, *SRMR*=.044) are outlined in Figure 12 and Table 33. Results indicate a negative relation between cognitive control and academics, that was the same strength as the association Aim 2 (β=-0.308, *p*<.001). Given the factor loadings, the cognitive control construct is a construct of poor cognitive control – such that higher factor loadings are associated with greater errors of omission and longer and more variable reaction times for both Go and No-Go trials as well as lower d-prime ratios, less PES, and greater (more negative) CRN values. Therefore, better cognitive control skills (negative Z-scores) were associated with higher academic skills (positive Z-scores). Table 34 shows a comparison of factor loadings and error variances between models.

Table 35 shows factor loadings, model fit, and regressions between latent factors for Observed and Adjusted Measures SEM models. Model fit, as indexed by AIC, was better for Adjusted Measures models than Observed Measures models. The cognitive latent factor in the Observed Measures models best described variance in the SDRT indicators (λ_{SDRTC}=0.868,

λ_{SDRTI}=0.876). Whereas, the cognitive latent factor in the Adjusted Measure models best described variance in EOM (λ=0.974). Additionally, the cognitive latent factor in Adjusted Measures model described variance in 7 out of 9 behavioral measures. While, the cognitive latent factor in the Observed Measure model described variance in 5 out of 9 behavioral measures. There were also differences in factor loadings of ERP components, as the CRN (λ=-0.318) and Pe Error (λ=-0.406) loaded onto the cognitive control latent factor in the Observed Measures model but only CRN (λ=-0.580) loaded onto the cognitive latent factor in the Adjusted Measures model. Directionality of factor loadings was consistent across models. All models showed a

Figure 12

Adjusted Measures SEM Model: Behavioral and Neural Indicators of Academic Skills



Note: Anchor variables for each construct are indicated with dotted lines. Green lines are associated with positive factor loadings or covariances and red lines are associated with negative factor loadings or covariances.

Table 33

Behavioral and ERP Indicators SEM: Solution for Adjusted Indicators

	Latent Estimate	Std.Error	Z-Score	Variance Estimate	Std. Error	Z-Value	Cognitive λ	Academic λ
EOM	1.00			0.046	0.004	11.09***	0.974	
RTC	0.255	0.064	3.99888	0.855	0.074	11.49***	0.249	
RTI	0.379	0.063	6.05***	0.829	0.075	11.09***	0.361	
SDRTI	0.257	0.066	3.86***	0.936	0.084	11.09***	0.240	
PEA	-1.037	0.015	-67.78***	0.000			-1.00	
d-prime	-0.757	0.049	-15.32***	0.493	0.044	11.29***	-0.708	
PES	-0.214	0.066	-3.26**	0.919	0.083	11.09***	-0.204	
CRN	-0.622	0.056	-11.02***	0.661	0.060	11.09***	-0.580	
Applied Problems	1.00			0.307	0.121	2.55*		0.836
Letter Word ID	0.963	0.166	5.82***	0.270	0.112	2.42*		0.843
Passage Comprehension	0.815	0.149	5.48***	0.499	0.095	5.24***		0.699
Cognitive Control				0.867	0.082	10.54***		
Academic Skills				0.634	0.131	4.85***		
Regressions								
Academic ~ Cognitive				-0.308	0.067	-4.62***		
Covariances								
RTC~~RTI				0.631	0.066	9.58***		
RTC~~d-prime				0.303	0.044	6.93***		
EOM~~RTC				-0.030	0.007	-4.22***		
RTC~~PES				0.157	0.038	-4.15***		
d-prime~~PES				-0.138	0.042	-3.27**		
SDRTI~~d-prime				-0.093	0.035	-2.63**		
RTI~~d-prime				0.101	0.039	2.55*		
d-prime~~CRN				0.101	0.030	3.34**		
Passage Comprehension~~ Letter Word ID				0.213	0.098	2.18**		

***p<.001, **p<.01, *p<.05, +p<.10

Table 34

Comparisons between Adjusted Score SEM Models

	Aim 2		Aim 3		Aim 2	Aim 3
	Cognitive λ	Academic λ	Cognitive λ	Academic λ	Error Variance	Error Variance
EOM	0.974		0.974		0.046***	0.046***
RTC	0.244		0.249		0.895***	0.855***
RTI	0.361		0.361		0.830***	0.829***
SDRTI	0.240		0.240		0.936***	0.936***
PEA	-1.00		-1.00		0.000	0.000
d-prime	-0.710		-0.708		0.488***	0.493***
PES	-0.204		-0.204		0.919***	0.919***
CRN			-0.580			0.661***
Applied Problems		0.836		0.836	0.307***	0.307*
Letter Word ID		0.843		0.843	0.270***	0.270*
Passage Comprehension		0.699		0.699	0.499**	0.499***
Cognitive Control					0.867***	0.867***
Academic Skills					0.634***	0.634***
Regressions	*β*	Std. Error	*β*	Std. Error		
Academic ~ Cognitive	-0.308***	0.067	-.308***	0.067		
Covariances		Estimate		Estimate		
RTC~~RTI		0.651***		0.631***		
RTC~~d-prime		0.301***		0.303***		
EOM~~RTC		-0.061***		-0.030***		
EOM~~RTI		-0.0308				
RTC~~PES		-0.172***		0.157***		
d-prime~~PES		-0.133**		-0.138**		
RTI~~SDRTI		0.046				
SDRTI~~d-prime		-0.103**		-0.093**		
RTI~~d-prime		0.085*		0.101*		
EOM~~d-prime		-0.015				
EOM~~PES		0.024+				
d-prime~~CRN				0.101**		
Covariances (cont).		Estimate		Estimate		
Passage Comprehension~~ Letter Word ID		0.213*		0.213**		

***p<.001, **p<.01, *p<.05, +p<.10

Table 35

Comparisons of Observed and Adjusted Measures SEM Models

	Observed Measures Model				Adjusted Measures Model			
	Aim 2		Aim 3		Aim 2		Aim 3	
	Cognitive λ	Academic λ	Cognitive λ	Academic λ	Cognitive λ	Academic λ	Cognitive λ	Academic λ
EOM	0.515		0.515		0.974		0.974	
RTC	0.388		0.380		0.244		0.249	
RTI					0.361		0.361	
SDRTC	0.884		0.868					
SDRTI	0.856		0.876		0.240		0.240	
PEA	-0.315		-0.318		-1.00		-1.00	
d-prime					-0.710		-0.708	
PES					-0.204		-0.204	
CRN			-0.318				-0.580	
Pe Error			-0.406					
Applied Problems		0.812		0.797		0.836		0.836
Letter Word ID		0.852		0.884		0.843		0.843
Passage Comprehension		0.729		0.737		0.699		0.699
Error Variance	*unstand.*	*stand*	*unstand.*	*stand*	*unstand.*	*stand*	*unstand.*	*stand*
Cognitive Control	0.264	1.00	0.240	1.00	0.866	1.00	0.866	1.00
Academic Skills	0.558	0.849	0.575	0.885	0.633	0.885	0.633	0.885
Regressions	*β*	Std. Error	*B*	Std. Error	*β*	Std. Error	*β*	Std. Error
Academic ~ Cognitive	-0.612***	0.133	-0.557***	0.149	-0.308***	0.067	-.308***	0.067
Fit Indices								
χ2	16.01		31.83		29.53		48.05+	
CFI	1.000		0.995		0.996		0.993	
AIC	5385.84		5762.86		5150.51		5744.69	
RMSEA	0.002		0.030		0.034		0.039	
SRMR	0.039		0.043		0.037		0.044	

***p<.001, **p<.01, *p<.05, +p<.10

negative association between cognitive control and academic skills. Factor loadings indicate that the cognitive control latent factor reflects poorer cognitive control. Therefore, greater academic skills were predicted from greater (more negative) cognitive control skills. The strength of associations between the cognitive control and academic latent factors was stronger for Observed Measures models than for Adjusted Measures models. However, because the Adjusted Measures models described variance in more behavioral indicators of Go/No-Go task performance it is difficult to directly compare these values. However, comparisons of the R^2 indices across models suggest that the Observed SEM models accounted for the most variance in literacy subtests of the Woodcock-Johnson. Whereas, the Adjusted SEM models accounted for the most variance in the applied problems subtest. Table 36 highlights the percentage of variance accounted for in each academic measure across traditional multiple regression and SEM models. Traditional regression models showed worse goodness-of-fit as indicated by R^2 between .019 and .122, as compared to the Observed SEM (.535> R^2>.782) and Adjusted SEM (.488> R^2>.711) models.

Table 36

Percent Variance Explained in Academic Subscales: Comparisons across Predictive Models

Analytic Approach		Letter Word ID	Passage Comprehension	Applied Problems
Traditional	Behavioral Indicators	10.2%	8.1%	1.9%
Regression	Behavioral and Neural Indicators	12.2%	*NA*	6.3%
Observed	Behavioral Indicators	74.0%	53.5%	64.8%
SEM	Behavioral and Neural Indicators	78.2%	54.4%	63.5%
Adjusted	Behavioral Indicators	71.1%	48.8%	70.0%
SEM	Behavioral and Neural Indicators	71.1%	48.8%	70.0%

Discussion

Brain and behavior measures of performance on computerized tasks are widely used to quantify cognitive control skills, estimate differences in academic achievement, and investigate the etiology of neurodevelopmental disorders. While computerized tasks afford researchers tight experimental control, varying task features related to task randomization, may bias estimates of accuracy and reaction time measures by influencing performance on individual trials. However, whether trial features have a practically significant impact on performance for these types of forced-response tasks is underexplored. Moreover, studies often report bivariate relations between academic skills and a single score of cognitive control performance. As a result, relations between cognitive control and academic skills are inconsistent and relatively weak.

The current project aimed to investigate relations between academic skills and cognitive control using novel methods for quantifying both overt (observable behavior) and covert (internal cognitive processes indexed by neural measures) cognitive processes. IRT models were leveraged to identify task features that impact children's behavioral performance and accommodate for inter-trial variation in Go/No-Go performance among children. SEM models were then used to estimate relations between latent factors of cognitive control and academics and examine how neural measures loaded onto behavioral factors of cognitive control. Overall, findings indicate that trial randomization does lead to differences in trial difficulty. Additionally, a latent factor of cognitive control incorporating multiple measures of Go/No-Go task performance was more strongly related to academic outcomes than traditional multiple regression approaches. Neural measures did not help explain additional variance in academic skills. However, it is possible that the error-identification mechanisms these neural measures reflect differentially relate to performance for individual children, causing some to increase their task performance, whereas others may disengage due to higher emotional reaction to erring. Analyses exploring underlying factor structure of all Go/No-Go measures indicate that task performance may best be supported by multidimensional theories of cognitive control. Implications for developmental and clinical researchers are discussed below.

Go/No-Go Task Performance

Overall, children performed well on the Go/No-Go task demonstrating low occurrences of errors of omission and moderate occurrences of commission errors. However, there was quite a range in performance, with average EOM ranging from 0% to 61.92% and average ECOM ranging from 5% to 96.92%. Additionally, gender differences emerged in reaction time variables. Specifically, boys had greater variance in average RTC values than girls and also showed greater inter-trial variability in reaction time to Go trials. However, girls showed greater variance in inter-trial reaction times for No-Go trials, indicating perhaps that girls made fewer EOM overall and had slower average reaction time to Go trials. This was accompanied by higher d-prime scores for girls as well, indexing a better ratio of successful hit trials to incorrect false alarm trials. This supports previous work that indicates that average performance for boys and girls is similar in measures of cognitive control, but that within gender, the distribution of performance may be more variable for boys than girls (Montroy et al., 2016).

Findings from Aim 1 of the current study demonstrate that task features can impact children's speed and accuracy on the Go/No-Go task, replicating findings from Durston (Thomas, Worden, et al., 2002; Thomas, Yang, et al., 2002) and Schulz (2009) work in which trial difficulty was experimentally manipulated. Trial-level variance as indicated by the random effect of trial (see Table 7) was greater for Go trials, such that trial order accounted for approximately 2.53% of the variance in RTC and 8.60% of the variance in EOM. Whereas for

No-Go trials, trial order only accounted for 1.39% of variance in RTI and 2.80% of variance in ECOM. This may be explained by the greater frequency of Go versus No-Go trials in the task, by design. Trial order was operationalized as the sequential trial in the game (e.g., 1st trial, 150th trial, 320th trial). It is important to keep in mind that trial features were randomized within children which is why trials were cross-classified among children. So, regardless of whether the trial was preceded by a correct response or an error or any number of Go trials, additional variance was explained simply by which trial a response was being used to estimate expected reaction time and likelihood of erring.

Greater number of preceding Go trials was related to increased reaction times and likelihood of error on both Go and No-Go trials in the current sample. Results indicated that each additional preceding Go trial was related to less than 5 ms difference in RT and less than 1% difference in expected likelihood of error. As such, these effects are not particularly practically significant unless comparing performance on trials with few-to-none preceding Go trials and large amounts of preceding Go trials. The current project utilized data from a task that had large variability in the range of preceding Go trials between participants. Some participants only completed a task with 0 to 8 preceding Go trials, while others completed a task with 0 to 22 preceding Go trials. I suggest that future researchers keep the range of trial difficulty consistent across children completing Go/No-Go tasks in order to minimize any potential biases in performance estimations. Having a moderate range of preceding Go trials may also help ensure that there will be a sufficient amount of error trials for extraction of ERP components, as cognitive tasks with high accuracy rates may not produce a sufficient amount of usable neural data to estimate ERN and Pe Error components.

Additionally, trial performance was impacted by the previous trial accuracy such that post-error trials were associated with slower reaction times to both Go and No-Go stimuli as well as increased likelihood of having an omission error but decreased likelihood of making an error of commission. This finding suggests that children in the current sample may have shown increased orientation to an error and decreased their speed, which in turn lead to less task engagement (greater chance of EOM) or better motor inhibition (less chance of ECOM). Correlations among observed average variables also support this assumption as vigilance in response to Go trials (less EOM) was related to better PEA as measured by the average of Go trial performance following correct versus incorrect trials. However, only about one-fourth of the sample demonstrated post-error increases in accuracy. It may be that children have varied reactions to making an error such that some disengage from the task due to frustration or decreased feelings of competence, while others may increase their effort in order to improve their performance. Competing theories of post-error slowing also support the notion that individuals respond differently to errors (Gehring & Fencsik, 2001; King et al., 2010; Steinborn et al., 2012). Notably, PEA only compares average accuracy on Go trials following error – though does not distinguish between omission and commission errors. It may be that commission errors elicit a different post-error response than omission errors. Future research should investigate whether error type influences post-error adjustments to performance.

Relations between Measures of the Go/No-Go

A one-factor solution of cognitive control derived from the Go/No-Go task demonstrated better goodness-of-fit than a two-factor solution. For the observed model, variance in standard deviations in reaction time measures was best explained by the cognitive control construct. In the adjusted score model, accuracy indicators had the highest factor loadings perhaps due to smaller sample variances in SDRTC and SDRTI. Additionally, standard errors in variances estimates were smaller in the adjusted score model than the observed model. Notably, post-error slowing was not well correlated with any other behavioral measure. This finding emerged despite over half of the children demonstrating a positive PES value, indicating increased reaction times in response to error.

This unexpected finding may be explained by competing theories about the function of PES. One widely held idea is that PES is an index of either a conscious task strategy (Gehring & Fencsik, 2001; MacDonald et al., 2000). Individuals trade off speed for accuracy, slowing down to avoid future error. In contrast, PES has also been thought to reflect a heightened emotional reaction to error, causing decreased reaction times, accuracy, and perhaps disengagement from the task (Notebaert et al., 2009; Steinborn et al., 2012). Perhaps relations with other behavioral variables were not observed in the current analysis because PES indexed different reactions to errors for different children. Competing reactions to error may in turn present with varied relations to accuracy measures which would be masked in average correlations among all children. More work is needed to understand the variance in children's error-related adjustments in reaction time.

Neural indicators did not improve model fit or factor loadings for behavioral indicators. However, variance estimates for cognitive control decreased for both observed and predicted CFA models, indicating that neural measures reduced the distribution of the latent factor, also demonstrated by smaller standard error variances. Notably, the ERN did not load onto the cognitive control construct, even though in the traditional regression approach the ERN was the only significant neural predictor of the letter word ID subscale. The ERN component is perhaps the most frequently reported response-locked ERP component in investigations of cognitive control (Downes et al., 2017; Gehring et al., 2011). Findings from the exploratory EFA highlight that the ERN may index a separate but related construct to cognitive control. Errors of commission was the only behavioral indicator to load onto the ERN factor in the EFA, such that greater (more negative) amplitudes in the ERN were associated with fewer errors of commission. Thus, the ERN may index error salience associated with commission but not omission errors.

The CRN and Pe Error loaded onto cognitive control constructs but in unexpected and contradictory directions. The CRN is thought to index an attention cue to sustain engagement through a task (Maier et al., 2011), whereas the Pe is thought to reflect error salience (Overbeek et al., 2005). Therefore, it was anticipated that greater errors of omission would have been associated with smaller (less negative) CRN values and larger (more positive) Pe values. However, in the current analysis greater errors of omission were associated with larger (more negative) CRN values and smaller (more attenuated) Pe values. Future research should investigate whether there is a non-linear association between behavioral and neural predictors that could explain this unexpected finding.

Overall, findings from the current project demonstrates that observations of multiple cognitive processes can be estimated from performance on one task. Often, children are required to complete a battery of cognitive tasks during participation in a research study. Large batteries can be laborious and result in greater data missingness due to cognitive fatigue or boredom. While there is merit in using multiple tasks to assess cognitive control, many studies limit performance on a task to one score of children’s functioning. Findings from the current study highlight that different behavioral and neural measures derived from the same task appear to reflect different cognitive processes. As such, it is suggested that researchers either create a latent construct of cognitive skills from varying task measures, as done in this study, or limit the generalizability of the measure to the trial (i.e., Go, No-Go) and performance type (i.e., accuracy, reaction time). Future studies that investigate relations among multiple performance measures from two or three cognitive control tasks – each with similar measures of reaction time, accuracy, and ERP components - would greatly increase our understanding of how task specificity may alter relations among behavioral indicators. For example, assessing reaction time and accuracy measures in a Go/No-Go, flanker, and Sternberg task which are thought to be measures of inhibition, attention, and working memory, respectively (Corbin & Marquer, 2013; Servant & Logan, 2019). Findings from such work would clarify whether individual measures represent different cognitive constructs or whether there are more similarities among type of measure (i.e., reaction time or accuracy) across tasks.

The methodological approach chosen for estimating CFAs does present with a few limitations. In the current study, estimated relations between behavioral indicators were completed prior to the addition of neural indicators. However, the exploratory EFA results demonstrated that ERPs and behavioral indicators shared loadings onto multiple underlying cognitive factors. As such, future research may consider estimating cognitive constructs with all neural and behavioral variables at the same time.

Additionally, the goal of the current project was to find models that best fit as many Go/No-Go task measures as possible to match the theory that multiple performance measures on the task reflect related cognitive control skills. Although factor loadings varied in strength among indicator variables, CFA results explained how the underlying cognitive control skill predicted variance in Go/No-Go performance measures. There may in fact be multiple "best fit" models for measures of cognitive control depending on task features and sample populations. Future studies may choose to have a higher factor loading criterion for indicator variables. While these approaches can be applied to any computerized task, I would expect to see differences in which measures of behavior are correlated within task based on task design and sample demographics. For example, children in the current study were in elementary school. Elementary school is a developmental period where cognitive control is rapidly developing and variation between participants is typically larger than adult populations. Therefore, similar investigations with different developmental populations should be pursued in the future to investigate whether relations between these measures differ across development.

Relations between Cognitive and Academic Skills

Relations between cognitive control and academic skills were assessed in three iterations. The traditional regression approach utilized multiple regression models to predict academic outcomes in three separate models predicting literacy or math subtests of the Woodcock-Johnson. This approach weights the predictor variables equally. The SEM model approach regressed a latent factor of cognitive control onto a latent factor of academic skills. SEM models allow for indicator variables to have different factor loadings onto latent constructs. The SEM approach was completed for both observed cognitive control measures and adjusted (predicted) cognitive control measures. Both behavioral and neural dependent variables were assessed as predictors of academic skills.

Comparisons across analytic approaches suggest that single score correlations with academics may be underestimating relations between cognitive and academic skills. Traditional regression models were least able to predict academic outcomes, explaining between 1.9% and 12.2% of variance in academic subscales. Only one behavioral predictor mapped onto each academic measure in the traditional multiple regression approach. D-prime predicted performance on the literacy subscales, while RTC predicted performance on the applied problems subtest. Additionally, ERP components did not relate to the passage comprehension subtest. However, the ERN was significantly associated with letter word identification and the CRN was related to applied problems, though neither demonstrated strong relations with the academic outcome. Both the Observed and Adjusted Measures SEM models were able to predict a moderate to high amount of variance in academic skills. The Observed SEM models best predicted outcomes in the literacy subtests of the Woodcock-Johnson, accounting for approximately 53% to 78% of the observed variance in passage comprehension and letter word identification performance. Whereas, the Adjusted SEM models accounted for the most variance in the applied problems subtest, approximately 70%. However, Observed and Adjusted Measure models showed at most a 7% difference in variance explained. Together, these findings highlight that there may not be one "best choice" cognitive variable that can predict variance in all academic skills. It is likely beneficial to consider performance across multiple measures of cognitive control when predicting academic skills, as academic achievement is dependent upon several cognitive processes that allow children to attend in the classroom and demonstrate positive approaches to learning, such as persistence (Li-Grining et al., 2010; Torgrimson et al., 2021).

Although neural and behavioral measures of cognitive control are thought to index similar but distinct cognitive processes, results from the current investigation indicate that there was not much difference in factor or residual variance between SEM models that included only behavioral indicators of cognitive control (Aim 2) or that included behavioral and ERP indicators of cognitive control (Aim 3). Similarly, there were small changes in factor loadings and barely any changes in predictive quality between cognitive control and academics. This may be due to the statistical approach which fit ERP measures onto factors of cognitive control after finalizing the best fit among behavioral indicators. Exploratory analyses indicated that the CRN and Pe Error were most related (though moderately) to accuracy on No-Go trials, as indexed by ECOM. ECOM was not retained as an indicator variable in any of the SEM analyses, likely due to relatively weak associations with other behavioral variables. Additionally, findings showed that the ERN was not well correlated with other measures. Specifically, the ERN only loaded onto a factor comprised of the ERN, CRN, and Pe Error, of which the CRN and Pe Error had relatively weak factor loadings. Yet, the traditional multiple regression approach indicated that the ERN was related to performance on the letter word identification subtest. Future research may consider alternative approaches to mapping relations among Go/No-Go task measures that would allow for academic associations with ECOM and the ERN to be modeled.

Conclusion

In computerized cognitive control tasks have been proposed as more “objective” assessments of clinical impairments to improve diagnostic accuracy, predict response to treatment, and provide an understanding of mechanisms of the disorder – specifically in biopsychological processes that underly observable symptoms (Cuthbert & Insel, 2013; Geraldo et al., 2019). However, previous work indicates that impairments in cognitive control skills are not always apparent in average scores from neuropsychological assessments (Nigg, 2005). It has been suggested that trial-to-trial variability within performance may better help characterize different clinical syndromes (DuPaul & Stoner, 2014). This study provides evidence that both children’s latent abilities and trial order does account for difference in trial-to-trial performance.

The model presented here can be replicated with any computerized task that has trial-level behavioral data. This may be particularly important for clinical research and investigations into potential early predictors of neurodevelopmental disorders. Additionally, findings provide preliminary evidence that adjusted performance scores may better help capture unitary constructs of cognitive control than observed measures, though additional research is needed to support this claim. Moreover, results of the current study emphasize that incorporating multiple measures of cognitive control from even one behavioral task is superior to predicting variance in academic outcome measures than singular cognitive outcome measures. Overall, findings suggest that researchers consider reducing the total amount of cognitive tasks that participants are asked to complete and instead maximize the value of the data that can be extracted from one cognitive task, by including multiple measures of performance in estimations of ability.

Appendix A

Table A1

RT Correct: IRT Factor Estimates by Model

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed Effect (SE)						
(Intercept)	555.66(4.48)***	567.92(4.53)***	577.42(4.33)***	568.63(5.75)***	735.63(28.79)***	729.71(2907)***
Trial Difficulty		-5.06(0.25)***	-5.06(0.25)***	-5.06(0.25)***	-5.06(0.25)***	-4.64(0.26)***
Data			-	-	-	-
Source			107.88(13.45)***	105.77(13.36)***	111.47(12.67)***	111.54(12.79)***
Female				17.44(7.57)*	15.91(7.16)*	16.54(7.23)*
Age					-26.97(4.56)***	-26.54(7.23)
Post-Error						10.84(1.99)***
Random Effect (SD)						
Trial ID	755(27.48)	791.5(28.13)	793.7(28.17)	793.6(28.17)	793.4(28.17)	758.1(27.53)
Subject	5039(70.98)	5048.6(71.05)	4118.4(64.17)	4042.7(63.58)	3598.4(59.99)	3666.2(60.55)
Residual	25739(160.43)	25562.9(159.88)	25562.4(159.88)	25562.4(159.88)	25562.4(159.88)	25564.5(159.89)
Fit Statistics						
AIC	803631.6	803226.3	803170.0	803166.8	803135.8	784257.2
BIC	803677.7	803271.4	803224.2	803230.0	803208.1	784338.3
Log Likelihood	-401811.8	-401608.1	-401579.0	-401576.4	-401559.9	-392119.6
Deviance	803623.6	803216.3	803158.0	803152.8	803119.8	784239.2
Residual (df)	61713	61712	61711	61710	61709	6056
Chi-square (*p*)		407.27(*p*<.001)	58.25(*p*<.001)	5.26(*p*=.022)	32.98(*p*<.001)	

***p<.001, **p<.01, *p<.05, +p<.10

Table A2

RT Error: IRT Factor Estimates by Model

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed Effect (SE)						
(Intercept)	469.77(4.39)***	477.49(4.74)***	481.47(4.85)***	475.35(6.32)***	633.05(31.80)***	621.63(32.14)***
Trial Difficulty		-2.70(0.63)***	-2.69(0.64)***	-2.69(0.63)***	-2.67(0.63)***	-1.48(0.69)*
Data Source			-47.10(14.69)**	-45.22(14.68)**	-50.16(14.14)***	-50.25(14.27)***
Female				12.35(8.23)	10.64(7.91)	11.61(7.98)
Age					-25.47(5.04)***	-25.03(5.09)***
Post-Error						22.01(5.35)***
Random Effect (SD)						
Trial ID	541.4(23.27)	550.1(23.45)	552.8(23.51)	552(23.50)	558.9(23.64)	465.8(21.58)
Subject	3932.1(62.71)	3924.7(62.65)	3751.7(61.25)	3714(60.94)	333.2(57.73)	3383.2(58.17)
Residual	29766.4(172.53)	29698.5(172.33)	29696.5(172.33)	29697(172.33)	29689.2(172.31)	29661.1(172.22)
Fit Statistics						
AIC	166470.1	116453.9	116445.8	116445.6	116423.1	113648.5
BIC	116498.4	116489.4	116488.3	116495.2	116479.7	113712.0
Log Likelihood	-58231.0	-58222.0	-58216.9	-58215.8	-58203.5	-56815.2
Deviance	116462.1	116443.9	116433.8	116431.6	116407.1	113630.5
Residual (df)	8816	8815	8814	8813	8812	8602
Chi-square (*p*)		18.13(*p*<.001)	10.11(*p*=.001)	2.24(*p*=.134)	24.53(*p*<.001)	

***p<.001, **p<.01, *p<.05, +p<.10

Table A3

EOM: IRT Factor Estimates by Model

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed Effect (SE)						
(Intercept)	-2.87(0.07)***	-2.81(0.08)***	-2.87(0.08)***	-2.84(0.11)***	0.62(0.54)	0.03(0.52)
Trial Difficulty		-0.03(0.01)***	-0.03(0.01)***	-0.03(0.01)***	-0.03(0.01)***	0.01(0.01)*
Data Source			0.62(0.25)*	0.61(0.25)*	0.50(0.23)*	0.42(0.22)
Female				-0.06(0.14)	-0.10(0.13)	-0.06(0.13)
Age					-0.56(0.09)***	-0.53(0.08)***
						1.19(0.03)***
Random Effect (SD)						
Trial ID	0.11(0.33)	0.11(0.33)	0.11(0.33)	0.11(0.33)	0.11(0.33)	0.08(0.29)
Subject	1.37(1.17)	1.37(1.17)	1.34(1.16)	1.34(1.16)	1.17(1.08)	1.07(1.03)
Fit Statistics						
AIC	34600.3	34581.9	34577.6	34579.5	34542.0	31943.1
BIC	34627.7	34618.4	34623.3	34634.2	34605.9	32015.9
Log Likelihood	-17297.2	-17286.9	-17283.8	-17283.7	-17264.0	-15963.6
Deviance	34594.3	34573.9	34567.6	34567.5	34528.0	31927.1
Residual (df)	67672	67671	67670	67669	67668	65984
Chi-square (*p*)		20.43(<.001)	6.22(.013)	0.18(.668)	39.45(<.001)	

***p<.001, **p<.01, *p<.05, +p<.10

Table A4

ECOM: IRT Factor Estimates by Model

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed Effect (SE)						
(Intercept)	-0.48(0.05)***	-0.55(0.05)***	-0.58(0.06)***	-0.40(0.08)***	-0.32(0.41)	-0.24(0.42)
Trial Difficulty		0.03(0.01)***	.03(0.01)***	0.03(0.01)***	0.03(0.01)***	0.02(0.01)***
Data Source			0.34(0.19)+	0.30(0.18)	0.29(0.18)	0.29(0.18)
Female				-0.36(0.10)***	-0.36(0.10)***	-0.38(0.10)***
Age					-0.01(0.07)	-0.02(0.07)
Post-Error						-0.10(0.04)*
Random Effect (SD)						
Trial ID	0.02(0.15)	0.02(0.15)	0.02(0.15)	0.02(0.15)	0.02(0.15)	0.02(0.13)
Subject	0.73(0.85)	0.73(0.85)	0.72(0.85)	0.69(0.83)	0.69(0.83)	0.71(0.84)
Fit Statistics						
AIC	27902.4	27879.9	27878.6	27868.3	27870.3	27139.2
BIC	27926.5	27912.0	27918.7	27916.5	27926.5	27203.2
Log Likelihood	-13948.2	-13935.9	-13934.3	-13928.2	-13928.1	-13561.6
Deviance	27896.4	27871.9	27868.6	27856.3	27856.3	27123.2
Residual (df)	22614	22613	22612	22611	22610	22017
Chi-square (*p*)		24.52(<.001)	3.34(.068)	12.22(<.001)	0.04(.845)	

****p*<.001, ***p*<.01, **p*<.05, +*p*<.10

Table A5

Predicted and Observed Measures of Behavior: Descriptive Statistics

	Min		Mean		Max		SD	
Variable	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted
EOM	0.00	0.49	8.93	8.79	61.92	61.10	10.00	9.74
ECOM	5.00	8.42	39.35	39.41	96.92	91.01	18.63	17.30
RT Correct (ms)	362.35	361.42	555.99	554.65	769.76	761.12	71.35	69.83
RT Error (ms)	309.67	338.83	470.01	467.63	667.54	630.59	71.57	56.35
SDRTC (ms)	58.76	23.09	158.83	27.85	295.01	34.55	42.89	1.65
SDRTI (ms)	29.77	5.35	153.75	15.39	352.87	23.12	61.64	2.61
PEA	-70.72	-32.17	-9.57	-10.29	7.58	-0.90	13.31	7.63
PES (ms)	-131.29	-1.29	17.45	18.52	178.91	37.46	51.51	6.64
d-prime	-0.42	-0.15	1.86	1.84	3.74	3.49	0.83	0.76

Appendix B

A series of models were run to estimate one-factor and two-factor CFAs of cognitive control. Model fit criterion included chi-square (*p*>.05), comparative fit index (CFI,>0.900), root mean square approximation (RMSEA, <.05), and standardized room mean square residual (SRMR, <.05). One and two-factor models were compared using AIC criterion. For all models, variables and co-variances were removed/added one-at-a-time to increase goodness of fit. Variables were removed if they demonstrated large negative residual variances and lambda factor loadings -1> λ>1 or fell below +/- 0.100 for lambda factor loadings (*p*>.05). Covariances were then added in incrementally, using the modindices function in lavaan, while also accounting for theoretically relevant residual shared variances (e.g., reaction time variables likely had shared variance relating to motor skills that would not be explained by cognitive control construct) in order to obtain a well-fitting model. Covariances were removed if they were not statistically significant (*p*>.05) unless retention of covariance in the model increased goodness-of-fit.

One-Factor CFA Model Testing: Observed Behavioral Indicators

Ten models in total were run. Goodness-of-fit statistics for each model are presented in Table B1. Model 1 presented with all variables meeting the criterion of a factor loading +/- 0.100, and thus covariances were incrementally added to improve model fit. In Model 2, a covariance between RTC and RTI was added. In Model 3, a covariance between SDRTC and SDRTI was added. However, the factor loading of RTI fell below the loading threshold (λ=0.003, *p*=.909), and was removed as an indicator variable in Model 4. In Model 4, ECOM had a low and statistically non-significant loading (λ=0.058, *p*=.329) and was removed from further analyses. In Model 6, a covariance between EOM and PEA was added. However, PES dipped below the desired threshold (λ=0.012, p=.170) and was removed. In Model 7, d-prime had a large negative variance and high factor loading (λ=-2.697, p=.195) and was removed from additional models. In Model 8, the covariance between SDRTC and SDRTI was no longer significant (p=0.688) and so was removed from Model 9. A final covariance between RTC and PEA was added in Model 10 in order to improve χ^2 and *RMSEA* fit indices.

Table B1

Fit Indices for One-Factor CFA Models of Observed Behavioral Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	18	27	1324.08***	0.342	6673.25	0.409***	0.190
2	19	26	982.23***	0.515	6333.41	0.358***	0.251
3	20	25	745.76***	0.635	6098.94	0.317***	0.229
4	17	19	909.4***	0.458	5780.60	0.404***	0.193
5	15	13	262.30***	0.751	4963.83	0.258***	0.122
6	16	12	140.74***	0.871	4844.27	0.193***	0.196
7	14	7	126.26***	0.880	4027.54	0.244***	0.213
8	12	3	11.21*	0.986	3492.03	0.098+	0.036
9	11	4	11.45*	0.988	3490.28	0.081	0.031
10	12	3	2.21	1.000	3483.04	0.000	0.009

***p<.001, **p<.01, *p<.05, +p<.1

Two-Factor CFA Model Testing: Observed Behavioral Indicators

Ten models in total were run. Goodness-of-fit statistics for each model are presented in Table B2. In Model 1, ECOM had a negative variance of -0.034 that was not statically significant (*p*=.136). As such, ECOM error variance was set to 0 in Model 2. In Model 2, PES showed a low factor loading to Inhibition (λ=0.048, *p*=.415) and was removed from Model 3. In Model 4, a covariance between RTC and RTI was added. In Model 5, a covariance between SDRTC and SDRTI was added. In Model 6, a covariance between RTC and SDRTC was added. In Model 7, a covariance between SDRTI and RTC was added. In Model 8, a covariance between RTI and SDRTI was added. In Model 9, a covariance term between RTI and SDRTC was added. A final covariance between EOM and PEA was added in Model 10 in order to improve X^2 and *RMSEA* fit indices.

Table B2

Fit Indices for Two-Factor CFA Models of Observed Behavioral Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	26	19	616.10***	0.697	5981.28	0.331***	0.147
2	25	20	617.10***	0.697	5980.27	0.323***	0.147
3	24	21	617.77***	0.697	5978.94	0.315***	0.148
4	25	20	391.39***	0.812	5754.56	0.254***	0.122
5	26	19	240.17***	0.888	5605.34	0.201***	0.109
6	27	18	136.76***	0.940	5503.93	0.152***	0.098
7	28	17	116.83***	0.949	5486.01	0.143***	0.086
8	29	16	53.63***	0.981	5424.80	0.091**	0.059
9	30	15	27.88*	0.993	2401.05	0.055	0.030
10	31	14	21.28+	0.996	5396.45	0.043	0.028

***p<.001, **p<.01, *p<.05, +p<.1

One-Factor CFA Model Testing: Predicted Behavioral Indicators

Fifteen models in total were run. Goodness-of-fit statistics for each model are presented in Table B3. In Model 1, ECOM had a lower factor loading (λ=0.010, *p*=.857) and was removed as an indicator from further analyses. In Model 2, SDRTC had a low factor loading (λ=0.098, *p*=.137) and was removed as an indicator variable. In Model 3, PEA had a negative error variance of -0.009 (*p*=.474) and so the error variance of PEA was set to 0 in Model 4. In Model 5, a covariance term between RTC and RTI was added. In Model 6, a covariance term between RTC and d-prime was added. A covariance term between RTC and EOM was added in Model 7. Next, a covariance term between RTI and EOM was added to Model 8. In Model 9, a covariance term between RTC and PES was added. In Model 10, a covariance between d-prime and PES was included. A covariance term between RTI and SDRTI was added to Model 11. In Model 12, a covariance between SDRTI and d-prime was included. In Model 13, a covariance between RTI and d-prime was added. A covariance between EOM and d-prime was added to Model 14. A final covariance term between EOM and PES was added to Model 15.

Table B3

Fit Indices for One-Factor CFA Models of Adjusted Behavioral Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	18	27	1267.22***	0.473	6233.09	0.400***	0.191
2	16	20	464.32***	0.715	5415.64	0.278***	0.153
3	14	14	401.81***	0.742	4600.94	0.311***	0.155
4	13	15	402.15***	0.742	4599.27	0.300***	0.155
5	14	14	175.66***	0.892	4374.79	0.201***	0.088
6	15	13	77.94***	0.957	4279.07	0.132***	0.069
7	16	12	65.51***	0.964	4268.64	0.125***	0.070
8	17	11	54.50***	0.971	4259.32	0.117***	0.068
9	18	10	43.91***	0.977	4251.04	0.109**	0.059
10	19	9	25.95**	0.986	4238.63	0.089*	0.046
11	20	8	25.25**	0.989	4236.38	0.087*	0.042
12	21	7	14.57*	0.995	4227.70	0.061	0.035
13	22	6	11.74	0.996	4226.87	0.058	0.024
14	23	5	7.87	0.998	4224.99	0.045	0.022
15	24	4	4.03	1.000	4223.16	0.005	0.020

****p*<.001, ***p*<.01, **p*<.05, +*p*<.1

Two-Factor CFA Model Testing: Adjusted Behavioral Indicators

Eleven models in total were run. Goodness-of-fit statistics for each model are presented in Table B4. In Model 1, SDRTC had a low factor loading onto the attention construct (λ=0.099, *p*=.140) and was removed from Model 2. In Model 3, a covariance between RTC and RTI was added. A covariance between PES and SDRTC was added to Model 4. In Model 4, ECOM had a low negative variance of -0.014 (*p*=.591). As such, in Model 5, the error variance for ECOM was held constant at 0. In Model 6, a covariance between RTC and PES was added to the model. Next, a covariance between RTC and EOM was added to Model 7. In Model 8, a covariance between EOM and RTI was added. A covariance between SDRTI and RTI was added to Model 9. A covariance between SDRTC and RTC was then added to Model 10. A final covariance between SDRTC and SDRTC was added to Model 11 to reach fit indices thresholds for goodness-of-fit.

Table B4

Fit Indices for Two-Factor CFA Models of Adjusted Behavioral Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	26	19	374.27**	0.849	5356.09	0.255***	0.114
2	25	20	377.08***	0.848	5356.94	0.249***	0.118
3	26	19	112.30***	0.960	5094.16	0.131***	0.075
4	27	18	78.85***	0.974	5062.71	0.109***	0.061
5	26	19	79.34***	0.974	5061.20	0.105***	0.061
6	27	18	59.58***	0.982	5043.44	0.090**	0.059
7	28	17	47.45***	0.987	5033.31	0.079*	0.059
8	29	16	37.56**	0.991	5025.43	0.069	0.058
9	30	15	33.34**	0.992	5023.20	0.065	0.055
10	31	14	29.14*	0.994	5021.00	0.061	0.051
11	32	13	15.83	0.999	5009.69	0.028	0.035

***p<.001, **p<.01, *p<.05, +p<.1

One-Factor CFA Model Testing: Observed Behavioral and ERP Indicators

Ten models in total were run. Goodness-of-fit statistics for each model are presented in Table B5. In Model 1, Pe Correct had a low factor loading (λ=-0.002, *p*=.979) and was removed from the indicator variables in Model 2. In Model 2, the ERN had a low factor loading (λ=-0.018, *p*=.788) and was removed from the indicator variables. A covariance between CRN and RTC was added to Model 4. In Model 5, a covariance between PEA and Pe Error was added. In Model 5, the covariance between RTC and PEA was no longer significant and it was removed from Model 6. A covariance between CRN and Pe Error was then included in Model 7. A covariance between RTC and Pe Error was next added to Model 8. In Model 9, a covariance between SDRTC and CRN was included. The covariance factor between RTC and PEA was added back into Model 10 in order to improve goodness-of-fit in *RMSEA*.

Table B5

Fit Indices for One-Factor CFA Models of Observed Behavioral and Neural Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	20	25	264.11***	0.701	5649.61	0.197***	0.136
2	18	18	188.02***	0.768	4948.50	0.196***	0.121
3	16	12	139.87***	0.815	4247.45	0.208***	0.113
4	17	11	50.77***	0.942	4160.36	0.121***	0.070
5	18	10	35.58***	0.963	4147.16	0.102**	0.067
6	17	11	39.07***	0.959	4148.66	0.102**	0.070
7	18	10	31.91***	0.968	4143.49	0.094	0.059
8	19	9	23.51**	0.979	4137.10	0.081	0.041
9	20	8	13.17	0.993	4128.75	0.051	0.033
10	21	7	8.60	0.998	4126.19	0.031	0.025

***p<.001, **p<.01, *p<.05, +p<.1

Two-Factor CFA Model Testing: Observed Behavioral and ERP Indicators

Twelve models in total were run. Goodness-of-fit statistics for each model are presented in Table B6. In Model 1, Pe Correct did not significantly load onto either construct of attention (λ=0.035, p=.598) or inhibition (λ=-0.070, p=.262) and was removed from the indicator variables in Model 2. In Model 3, the ERN was removed from the attention construct (λ=0.024, p=.717) and a covariance between Pe Error and the ERN was added. A covariance between the ERN and CRN was included in Model 4. In Model 5, a covariance between the CRN and Pe Error was added. Next, a covariance between the CRN and SDRTC was included in Model 6. A covariance between the Pe Error and PEA was then included in Model 7. A covariance between RTC and EOM was added in Model 8. In Model 8, the error variance of d-prime was negative, -0.010 (p=.618). As such the error variance of d-prime was held constant at 0 in Model 9. A covariance between CRN and RTI was then added to Model 10. A covariance between RTC and the CRN was included in Model 11. Last, a covariance between the ERN and RTC was included in Model 12 to reach the threshold for goodness-of-fit.

Table B6

Fit Indices for Two-Factor CFA Models of Observed Behavioral and Neural Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	43	48	280.11***	0.887	7176.88	0.140***	0.090
2	40	38	195.93***	0.921	6475.14	0.130***	0.077
3	40	38	172.18***	0.933	6451.39	0.120***	0.071
4	41	37	118.90***	0.959	6400.12	0.095***	0.056
5	42	36	107.24***	0.964	6390.46	0.090**	0.052
6	43	35	90.24***	0.972	6375.45	0.080**	0.051
7	44	34	79.13***	0.977	6366.35	0.073*	0.048
8	45	33	62.24**	0.985	6351.45	0.060	0.048
9	44	34	62.64**	0.986	6349.85	0.059	0.048
10	45	33	57.11**	0.988	6346.32	0.054	0.045
11	46	32	46.65*	0.993	6337.87	0.043	0.040
12	47	31	43.76+	0.994	6336.98	0.041	0.038

***p<.001, **p<.01, *p<.05, +p<.1

One-Factor CFA Model Testing: Adjusted Behavioral and ERP Indicators

Eight models in total were run. Goodness-of-fit statistics for each model are presented in Table B7. Model 1 did not converge and received an error notification that the sample covariance matrix was not positive-definitive. In Model 2, removing the ERN and Pe Correct indicator variables allowed the model to converge and the sample covariance matrix was a positive-definitive. Pe Error was removed as an indicator variable in Model 3, as it had a low factor loading (λ=0.043, *p*=.480). In Model 4, a covariance between the CRN and d-prime was added to the model. In Model 5, the covariance between RTI and SDRTI was removed because it was no longer significant (*p*=.165). Similarly, in Model 6 the covariance between EOM and d-prime was removed (*p*=.104). In Model 7, the covariance between EOM and PES was removed (*p*=.126). Last, the covariance between RTI and EOM (*p*=.064) was removed in Model 8.

Table B7

Fit Indices for One-Factor CFA Models of Adjusted Behavioral and Neural Indicators

Model	# Model Parameters	*df*	χ^2	*CFI*	*AIC*	*RMSEA*	*SRMR*
1	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
2	28	17	392.09***	0.791	4817.66	0.299***	0.114
3	26	10	19.14*	0.994	4117.00	0.061	0.025
4	27	9	7.32	1.000	4107.18	0.000	0.021
5	26	10	8.93	1.000	4106.79	0.000	0.023
6	25	11	11.92	0.999	4107.78	0.018	0.023
7	24	12	14.35	0.998	4108.21	0.028	0.024
8	23	13	19.13	0.996	4110.99	0.044	0.026

***p<.001, **p<.01, *p<.05, +p<.10

References

Ahmed, S. F., Grammer, J., & Morrison, F. (2021). Cognition in context: Validating group-based executive function assessments in young children. *Journal of Experimental Child Psychology*, *208*. http://dx.doi.org/10.1016/j.jecp.2021.105131

Albert, W. D., Hanson, J. L., Skinner, A. T., Dodge, K. A., Steinberg, L., Deater-Deckard, K., Bornstein, M. H., & Lansford, J. E. (2020). Individual differences in executive function partially explain the socioeconomic gradient in middle-school academic achievement. *Developmental Science*. http://dx.doi.org/10.1111/desc.12937

Albrecht, B., Brandeis, D., Uebel, H., Heinrich, H., Mueller, U. C., Hasselhorn, M., Steinhausen, H.-C., Rothenberger, A., & Banaschewski, T. (2008). Action monitoring in boys with attention-deficit/hyperactivity disorder, their nonaffected siblings, and normal control subjects: Evidence for an endophenotype. *Biological Psychiatry*, *64*(7), 615–625. https://doi.org/10.1016/j.biopsych.2007.12.016

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *CHILD NEUROPSYCHOLOGY*, *8*(2), 71–82.

Anderson, V. (2002). Executive Function in Children: Introduction. *Child Neuropsychology*, *8*(2), 69–70. https://doi.org/10.1076/chin.8.2.69.8725

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bekker, E. M., Kenemans, J. L., & Verbaten, M. N. (2004). Electrophysiological correlates of attention, inhibition, sensitivity and bias in a continuous performance task. *Clinical Neurophysiology*, *115*(9), 2001–2013. https://doi.org/10.1016/j.clinph.2004.04.008

Beretvas, S. N., & Murphy, DanielL. (2013). An Evaluation of Information Criteria Use for Correct Cross-Classified Random Effects Model Selection. *Journal of Experimental Education*, *81*(4), 429–463. https://doi.org/10.1080/00220973.2012.745467

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. PsycINFO. https://doi.org/10.1016/j.lindif.2011.01.007

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology*, *20*(3), 821–843. https://doi.org/10.1017/S0954579408000394

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*(2), 111–127. https://doi.org/10.1037/0003-066X.57.2.111

Blair, C., & Raver, C. C. (2015, January 5). *School Readiness and Self-Regulation: A Developmental Psychobiological Approach* [Review-article]. Http://Dx.Doi.Org/10.1146/Annurev-Psych-010814-015221. http://www.annualreviews.org/doi/10.1146/annurev-psych-010814-015221

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663. PsycINFO. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Bradley-Johnson, S., Morgan, S. K., & Nutkins, C. (2004). Book Review: The Woodcock-Johnson Tests of Achievement: Third Edition. *Journal of Psychoeducational Assessment*, *22*(3), 261–274. https://doi.org/10.1177/073428290402200305

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228. http://dx.doi.org/10.1080/87565640801982312

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Clawson, A., Clayson, P. E., Keith, C. M., Catron, C., & Larson, M. J. (2017). Conflict and performance monitoring throughout the lifespan: An event-related potential (ERP) and temporospatial component analysis. *Biological Psychology*, *124*, 87–99. https://doi.org/10.1016/j.biopsycho.2017.01.012

Corbin, L., & Marquer, J. (2013). Is Sternberg's Memory Scanning Task Really a Short-Term Memory Task? *Swiss Journal of Psychology*, *72*(4), 181–196. https://doi.org/10.1024/1421-0185/a000112

Criaud, M., & Boulinguez, P. (2013). Have we been asking the right questions when assessing response inhibition in go/no-go tasks with fMRI? A meta-analysis and critical review. *Neuroscience and Biobehavioral Reviews*, *37*(1), 11–23. https://doi.org/10.1016/j.neubiorev.2012.11.003

Crone, E. A., & Steinbeis, N. (2017). Neural perspectives on cognitive control development during childhood and adolescence. *Trends in Cognitive Sciences*, *21*(3), 205–215. http://dx.doi.org/10.1016/j.tics.2017.01.003

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*(1), 126. https://doi.org/10.1186/1741-7015-11-126

Danielmeier, C., & Ullsperger, M. (2011). Post-Error Adjustments. *Frontiers in Psychology*, *2*, 233. https://doi.org/10.3389/fpsyg.2011.00233

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of error-monitoring event-related potentials in adolescents. In R. E. Dahl & L. P. Spear (Eds.), *Adolescent Brain Development: Vulnerabilities and Opportunities* (Vol. 1021, pp. 324–328). New York Acad Sciences. https://doi.org/10.1196/annals.1308.039

De Franchis, V., Usai, M. C., Link to external site, this link will open in a new window, Viterbori, P., & Traverso, L. (2017). Preschool executive functioning and literacy achievement in Grades 1 and 3 of primary school: A longitudinal study. *Learning and Individual Differences*, *54*, 184–195. http://dx.doi.org/10.1016/j.lindif.2017.01.026

Delorme, R., Goussé, V., Roy, I., Trandafir, A., Mathieu, F., Mouren-Siméoni, M.-C., Betancur, C., & Leboyer, M. (2007). Shared executive dysfunctions in unaffected relatives of patients with autism and obsessive-compulsive disorder. *European Psychiatry*, *22*(1), 32–38. PsycINFO. https://doi.org/10.1016/j.eurpsy.2006.05.002

Di Gregorio, F., Steinhauser, M., & Maier, M. E. (2016). Error-related brain activity and error awareness in an error classification paradigm. *Neuroimage*, *139*, 202–210. https://doi.org/10.1016/j.neuroimage.2016.05.074

Dixon, W. (1960). Simplified Estimation from Censored Normal Samples. *Annals of Mathematical Statistics*, *31*(2), 385–391. https://doi.org/10.1214/aoms/1177705900

Downes, M., Bathelt, J., & de Haan, M. (2017). Event-related potential measures of executive functioning from preschool to adolescence. *Developmental Medicine and Child Neurology*, *59*(6), 581–590. https://doi.org/10.1111/dmcn.13395

Downing, K. M. (2016). An examination of three theoretical models of executive functioning [ProQuest Information & Learning (US)]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering* (Vol. 77, Issues 2-B(E), p. No Pagination Specified). https://www.proquest.com/psycinfo/docview/1823912978/abstract/FE654472C222480CPQ/32

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. http://dx.doi.org/10.1037/bul0000192

DuPaul, G. J., & Stoner, Gary. (2014). *ADHD in the schools: Assessment and intervention strategies* (Third edition., Vol. 1–1 online resource (xviii, 366 pages)). The Guilford Press. http://site.ebrary.com/id/10895713

Durston, S., Thomas, K. M., Worden, M. S., Yang, Y., & Casey, B. J. (2002). The Effect of Preceding Context on Inhibition: An Event-Related fMRI Study. *NeuroImage*, *16*(2), 449–453. https://doi.org/10.1006/nimg.2002.1074

Durston, S., Thomas, K. M., Yang, Y., Uluğ, A. M., Zimmerman, R. D., & Casey, B. j. (2002). A neural basis for the development of inhibitory control. *Developmental Science*, *5*(4), F9–F16. https://doi.org/10.1111/1467-7687.00235

Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P. P. P., Hawkins, G. E., Heathcote, A., Holmes, W. R., Krypotos, A.-M., Kupitz, C. N., Leite, F. P., Lerche, V., Lin, Y.-S., Logan, G. D., Palmeri, T. J., Starns, J. J., Trueblood, J. S., van Maanen, L., … Donkin, C. (2019). The Quality of Response Time Data Inference: A Blinded, Collaborative Assessment of the Validity of Cognitive Models. *Psychonomic Bulletin & Review*, *26*(4), 1051–1069. https://doi.org/10.3758/s13423-017-1417-2

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*(5), 1043–1053. https://doi.org/10.1111/j.1469-8986.2009.00838.x

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465–486. https://doi.org/10.1207/s15326942dn2601_6

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of Crossmodal Divided Attention on Late Erp Components .2. Error Processing in Choice Reaction Tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Fishbein, D. H., Domitrovich, C., Williams, J., Gitukui, S., Guthrie, C., Shapiro, D., & Greenberg, M. (2016). Short-Term Intervention Effects of the PATHS Curriculum in

Young Low-Income Children: Capitalizing on Plasticity. *The Journal of Primary Prevention*, *37*(6), 493–511. https://doi.org/10.1007/s10935-016-0452-5

Flynn, T. M. (1975). Behavioral Components of School Readiness. *The Journal of Experimental Education*, *44*(1), 40–45. https://doi.org/10.1080/00220973.1975.11011510

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gashaj, V., Link to external site, this link will open in a new window, Oberer, N., Mast, F. W., & Roebers, C. M. (2019). The relation between executive functions, fine motor skills, and basic numerical skills and their relevance for later mathematics achievement. *Early Education and Development*, *30*(7), 913–926. http://dx.doi.org/10.1080/10409289.2018.1539556

Gavin, W. J., & Davies, P. L. (2007). *Obtaining Reliable Psychophysiological Data with Child Participants: Methodological Considerations* (pp. 424–448). Cambridge University Press. https://doi.org/10.1017/CBO9780511499791.016

Gehring, W. J., & Fencsik, D. E. (2001). Functions of the Medial Frontal Cortex in the Processing of Conflict and Errors. *Journal of Neuroscience*, *21*(23), 9430–9437. https://doi.org/10.1523/JNEUROSCI.21-23-09430.2001

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2011, December 15). *The Error-Related Negativity (ERN/Ne)*. The Oxford Handbook of Event-Related Potential Components. https://doi.org/10.1093/oxfordhb/9780195374148.013.0120

Geraldo, A., Azeredo, A., Pasion, R., Dores, A. R., & Barbosa, F. (2019). Fostering advances to neuropsychological assessment based on the Research Domain Criteria: The bridge between cognitive functioning and physiology. *The Clinical Neuropsychologist*, *33*(2), 327–356. https://doi.org/10.1080/13854046.2018.1523467

Grammer, J. K., Carrasco, M., Gehring, W. J., & Morrison, F. J. (2014). Age-related changes in error processing in young children: A school-based investigation. *Developmental Cognitive Neuroscience*, *9*, 93–105. https://doi.org/10.1016/j.dcn.2014.02.001

Grammer, J. K., Gehring, W. J., & Morrison, F. J. (2018). Associations between developmental changes in error-related brain activity and executive functions in early childhood. *Psychophysiology*, *55*(3), 1–13. http://dx.doi.org/10.1111/psyp.13040

Grolnick, W. S., & Ryan, R. M. (1989). Parent Styles Associated With Children's Self-Regulation and Competence in School. *Journal of Educational Psychology*, *81*(2), 143–154. https://doi.org/10.1037/0022-0663.81.2.143

Gruetzmann, R., Riesel, A., Klawohn, J., Kathmann, N., & Endrass, T. (2014). Complementary modulation of N2 and CRN by conflict frequency. *Psychophysiology*, *51*(8), 761–772. https://doi.org/10.1111/psyp.12222

Hanna, G. L., Liu, Y., Rough, H. E., Surapaneni, M., Hanna, B. S., Arnold, P. D., & Gehring, W. J. (2020). A Diagnostic Biomarker for Pediatric Generalized Anxiety Disorder Using the Error-Related Negativity. *Child Psychiatry & Human Development*, *51*(5), 827–838. https://doi.org/10.1007/s10578-020-01021-5

Hillman, C. H., Pontifex, M. B., Motl, R. W., O'Leary, K. C., Johnson, C. R., Scudder, M. R., Raine, L. B., & Castelli, D. M. (2012). From ERPs to Academics. *Developmental Cognitive Neuroscience*, *2*(Suppl 1), S90–S98. https://doi.org/10.1016/j.dcn.2011.07.004

Hogan, A. M., Vargha-Khadem, F., Kirkham, F. J., & Baldeweg, T. (2005). Maturation of action monitoring from adolescence to adulthood: An ERP study. *Developmental Science, 8*(6), 525–534. https://doi.org/10.1111/j.1467-7687.2005.00444.x

Howard, S. J., Johnson, J., & Pascual-Leone, J. (2014). Clarifying inhibitory control: Diversity and development of attentional inhibition. *Cognitive Development*, *31*, 1–21. https://doi.org/10.1016/j.cogdev.2014.03.001

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Lawrence Erlbaum Associates.

Huang-Pollock, C. L., & Karalunas, S. L. (2010). Working memory demands impair skill acquisition in children with ADHD. *Journal of Abnormal Psychology*, *119*(1), 174–185. http://dx.doi.org/10.1037/a0017862

Hunter, S. J., & Sparrow, E. P. (2012). *Executive function and dysfunction: Identification, assessment and treatment*. Cambridge University Press.

Jaffe, L.E. (2009). Supplementary manual. Woodcock-Johnson III Tests of Achievement – Braille Adaptation. Louisville, KY: American Printing House for the Blind. Adapted with permission from Woodcock–Johnson III Tests of Achievement Normative Update by N. Mather & R. W. Woodcock, 2001, 2007, Rolling Meadows, IL: Riverside.

Jentzsch, I., & Dudschig, C. (2009). Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Quarterly Journal of Experimental Psychology (2006)*, *62*(2), 209–218. https://doi.org/10.1080/17470210802240655

Kamata, A. (2001). Item Analysis by the Hierarchical Generalized Linear Model. *Journal of Educational Measurement*, *38*(1), 79–93.

Kato, Y., Endo, H., & Kizuka, T. (2009). Mental fatigue and impaired response processes: Event-related brain potentials in a Go/NoGo task. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2008.12.008

Kim, M. H., Grammer, J. K., Marulis, L. M., Carrasco, M., Morrison, F. J., & Gehring, W. J. (2016). Early math and reading achievement are associated with the error positivity. *Developmental Cognitive Neuroscience*, *22*, 18–26. PsycINFO. https://doi.org/10.1016/j.dcn.2016.09.002

King, J. A., Korb, F. M., Cramon, D. Y. von, & Ullsperger, M. (2010). Post-Error Behavioral Adjustments Are Facilitated by Activation and Suppression of Task-Relevant and Task-Irrelevant Information Processing. *Journal of Neuroscience*, *30*(38), 12759–12769. https://doi.org/10.1523/JNEUROSCI.3274-10.2010

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, *46*(5), 1062–1077. https://doi.org/10.1037/a0020066

Lo, S. L. (2018). A meta-analytic review of the event-related potentials (ERN and N2) in childhood and adolescence: Providing a developmental perspective on the conflict monitoring theory. *Developmental Review*, *48*, 82–112. https://doi.org/10.1016/j.dr.2018.03.005

Luu, P., Flaisch, T., & Tucker, D. M. (2000). Medial frontal cortex in action monitoring. *Journal of Neuroscience*, *20*(1), 464–469. https://doi.org/10.1523/JNEUROSCI.20-01-00464.2000

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science (New York, N.Y.)*, *288*(5472), 1835–1838. https://doi.org/10.1126/science.288.5472.1835

Maier, M. E., & Steinhauser, M. (2016). Error significance but not error expectancy predicts error-related negativities for different error types. *Behavioural Brain Research*, *297*, 259–267. https://doi.org/10.1016/j.bbr.2015.10.031

Maier, M. E., Yeung, N., & Steinhauser, M. (2011). Error-related brain activity and adjustments of selective attention following errors. *NeuroImage*, *56*(4), 2339–2347. https://doi.org/10.1016/j.neuroimage.2011.03.083

Mather, N. (2001). Examiner's manual: Woodcock-Johnson III tests of achievement. *Itasca, IL: Riverside Publishing*.

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, *101*(3), 689–704. PsycINFO. https://doi.org/10.1037/a0014240

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947–959. PsycINFO. https://doi.org/10.1037/0012-1649.43.4.947

McClelland, M. M., Geldhof, G. J., Cameron, C. E., & Wanless, S. B. (2015). Development and Self-Regulation. In *Handbook of Child Psychology and Developmental Science* (pp. 1–43). American Cancer Society. https://doi.org/10.1002/9781118963418.childpsy114

Miller, M. R., Müller, U., Giesbrecht, G. F., Carpendale, J. I., & Kerns, K. A. (2013). The contribution of executive function and social understanding to preschoolers' letter and math skills. *Cognitive Development*, *28*(4), 331–349. http://dx.doi.org/10.1016/j.cogdev.2012.10.005

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. https://doi.org/10.1177/0963721411429458

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. https://doi.org/10.1006/cogp.1999.0734

Montroy, J. J., Bowles, R. P., Skibbe, L. E., McClelland, M. M., & Morrison, F. J. (2016). The development of self-regulation across early childhood. *Developmental Psychology*, *52*(11), 1744–1762. https://doi.org/10.1037/dev0000159

Motes, M. A., Link to external site, this link will open in a new window, Spence, J. S., Brier, M. R., Chiang, H.-S., Link to external site, this link will open in a new window, DeLaRosa, B. L., Eroh, J., Maguire, M. J., Link to external site, this link will open in a new window, Mudar, R. A., Tillman, G. D., Kraut, M. A., & Hart, J. (2018). Conjoint differences in inhibitory control and processing speed in childhood to older adult cohorts: Discriminant functions from a Go/No-Go task. *Psychology and Aging*, *33*(7), 1070–1078. http://dx.doi.org/10.1037/pag0000299

Murphy, D. L., & Beretvas, S. N. (2015). A comparison of teacher effectiveness measures calculated using three multilevel models for raters effects. *Applied Measurement in Education*, *28*(3), 219–236. http://dx.doi.org/10.1080/08957347.2015.1042158

Musser, E. D., & Raiker, J. S. (2019). Attention-deficit/hyperactivity disorder: An integrated developmental psychopathology and Research Domain Criteria (RDoC) approach. *Comprehensive Psychiatry*, *90*, 65–72. https://doi.org/10.1016/j.comppsych.2018.12.016

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*(5), 752–760.

Nigg, J. T. (2005). Neuropsychologic Theory and Findings in Attention-Deficit/Hyperactivity Disorder: The State of the Field and Salient Challenges for the Coming Decade. *Biological Psychiatry*, *57*(11), 1424–1435. https://doi.org/10.1016/j.biopsych.2004.11.011

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing—On the functional significance of the Pe Vis-a-vis the ERN/Ne. *Journal of Psychophysiology*, *19*(4), 319–329. https://doi.org/10.1027/0269-8803.19.4.319

Pontifex, M. B., Scudder, M. R., Brown, M. L., O'Leary, K. C., Wu, C.-T., Themanson, J. R., & Hillman, C. H. (2010). On the number of trials necessary for stabilization of error-related brain activity across the life span. *Psychophysiology*, *47*(4), 767–773. https://doi.org/10.1111/j.1469-8986.2010.00974.x

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods. 2nd ed* (Vol. 1). Sage.

Richardson, C., Anderson, M., Reid, C. L., & Fox, A. M. (2011). Neural indicators of error processing and intraindividual variability in reaction time in 7 and 9 year-olds (vol 53, pg 256, 2011). *Developmental Psychobiology*, *53*(8), 839–839. https://doi.org/10.1002/dev.20601

Ridderinkhof, K. R. (2002). Micro- and macro-adjustments of task set: Activation and suppression in conflict tasks. *Psychological Research*, *66*(4), 312–323. https://doi.org/10.1007/s00426-002-0104-7

Roebers, C. M., Röthlisberger, M., Cimeli, P., Michel, E., & Neuenschwander, R. (2011). School enrolment and executive functioning: A longitudinal perspective on developmental changes, the influence of learning context, and the prediction of pre-academic skills. *European Journal of Developmental Psychology*, *8*(5), 526–540. https://doi.org/10.1080/17405629.2011.571841

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48, 1-36.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Santesso, D. L., & Segalowitz, S. J. (2008). Developmental differences in error-related ERPs in middle- to late-adolescent males. *Developmental Psychology*, *44*(1), 205–217. https://doi.org/10.1037/0012-1649.44.1.205

Schroder, H. S., Fisher, M. E., Lin, Y., Lo, S. L., Danovitch, J. H., & Moser, J. S. (2017). Neural evidence for enhanced attention to mistakes among school-aged children with a growth mindset. *Developmental Cognitive Neuroscience*, *24*, 42–50. http://dx.doi.org/10.1016/j.dcn.2017.01.004

Schulz, K. P., Clerkin, S. M., Halperin, J. M., Newcorn, J. H., Tang, C. Y., & Fan, J. (2009). Dissociable neural effects of stimulus valence and preceding context during the inhibition of responses to emotional faces. *Human Brain Mapping*, *30*(9), 2821–2833. https://doi.org/10.1002/hbm.20706

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

Senderecka, M., Grabowska, A., Szewczyk, J., Gerc, K., & Chmylak, R. (2012). Response inhibition of children with ADHD in the stop-signal task: An event-related potential study. *International Journal of Psychophysiology*, *85*(1), 93–105. https://doi.org/10.1016/j.ijpsycho.2011.05.007

Servant, M., & Logan, G. D. (2019). Dynamics of attentional focusing in the Eriksen flanker task. *Attention, Perception, & Psychophysics*, *81*(8), 2710–2721. https://doi.org/10.3758/s13414-019-01796-3

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology (2006)*, *59*(4), 745–759. https://doi.org/10.1080/17470210500162854

Steinborn, M. B., Flehmig, H. C., Bratzke, D., & Schröter, H. (2012). Error reactivity in self-paced performance: Highly-accurate individuals exhibit largest post-error slowing. *Quarterly Journal of Experimental Psychology (2006)*, *65*(4), 624–631. https://doi.org/10.1080/17470218.2012.660962

Taylor, B. K., Gavin, W. J., Grimm, K. J., Passantino, D. E., & Davies, P. L. (2018). Modeling the interrelationships between brain activity and trait attention measures to predict individual differences in reaction times in children during a Go/No-Go task. *Neuropsychologia*, *109*, 222–231. https://doi.org/10.1016/j.neuropsychologia.2017.12.025

Thurm, F., Link to external site, this link will open in a new window, Li, S.-C., & Hämmerer, D. (2020). Maturation- and aging-related differences in electrophysiological correlates of error detection and error awareness. *Neuropsychologia*, *143*. http://dx.doi.org/10.1016/j.neuropsychologia.2020.107476

Torgrimson, S. J., Tan, P. Z., & Grammer, J. K. (2021). Associations among response inhibition, motivational beliefs, and task persistence in early elementary school. *Journal of Experimental Child Psychology*, *208*, 105141. https://doi.org/10.1016/j.jecp.2021.105141

Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A., & Klein, D. N. (2012). Electrocortical and behavioral measures of response monitoring in young children during a go/no-go task. *Developmental Psychobiology*, *54*(2), 139–150. https://doi.org/10.1002/dev.20590

Tottenham, N., Hare, T., & Casey, B. J. (2011). Behavioral Assessment of Emotion Discrimination, Emotion Regulation, and Cognitive Control in Childhood, Adolescence, and Adulthood. *Frontiers in Psychology*, *2*, 39. https://doi.org/10.3389/fpsyg.2011.00039

van Meel, C. S., Heslenfeld, D. J., Rommelse, N. N., Oosterlaan, J., & Sergeant, J. A. (2012). Developmental trajectories of neural mechanisms supporting conflict and error processing in middle childhood. *Developmental Neuropsychology*, *37*(4), 358–378. https://doi.org/10.1080/87565641.2011.653062

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Votruba, K. L., & Langenecker, S. A. (2013). Factor structure, construct validity, and age- and education-based normative data for the Parametric Go/No-Go Test. *Journal of Clinical and Experimental Neuropsychology*, *35*(2), 132–146. https://doi.org/10.1080/13803395.2012.758239

Wauthia, E., & Rossignol, M. (2016). Emotional Processing and Attention Control Impairments in Children with Anxiety: An Integrative Review of Event-Related Potentials Findings. *Frontiers in Psychology*, *7*, 562. https://doi.org/10.3389/fpsyg.2016.00562

Weigard, A., Soules, M., Ferris, B., Zucker, R. A., Sripada, C., & Heitzeg, M. (2020). Cognitive Modeling Informs Interpretation of Go/No-Go Task-Related Neural Activations and Their Links to Externalizing Psychopathology. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(5), 530–541. https://doi.org/10.1016/j.bpsc.2019.11.013

Weston, R., & Gore, P. A. (2006). A Brief Guide to Structural Equation Modeling. *The Counseling Psychologist*, *34*(5), 719–751. https://doi.org/10.1177/0011000006286345

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-old children. *Journal of Experimental Child Psychology*, *108*(3), 436–452. https://doi.org/10.1016/j.jecp.2010.08.008

Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2007). Developmental changes in error monitoring: An event-related potential study. *Neuropsychologia*, *45*(8), 1649–1657. https://doi.org/10.1016/j.neuropsychologia.2007.01.004

Wild-Wall, N., Oades, R. D., Schmidt-Wessels, M., Christiansen, H., & Falkenstein, M. (2009). Neural activity associated with executive functions in adolescents with attention-deficit/hyperactivity disorder (ADHD). *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *74*(1), 19–27. https://doi.org/10.1016/j.ijpsycho.2009.06.003

Winter, W., & Sheridan, M. (2014). Previous reward decreases errors of commission on later "No-Go" trials in children 4 to 12 years of age: Evidence for a context monitoring account. *Developmental Science*, *17*(5), 797–807. https://doi.org/10.1111/desc.12168

Wöstmann, N. M., Aichert, D. S., Costa, A., Rubia, K., Möller, H.-J., & Ettinger, U. (2013). Reliability and plasticity of response inhibition and interference control. *Brain and Cognition*, *81*(1), 82–94. https://doi.org/10.1016/j.bandc.2012.09.010

Wright, L., Lipszyc, J., Dupuis, A., Thayapararajah, S. W., & Schachar, R. (2014). Response Inhibition and Psychopathology: A Meta-Analysis of Go/No-Go Task Performance. *Journal of Abnormal Psychology*, *123*(2), 429–439. https://doi.org/10.1037/a0036295

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931–959. https://doi.org/10.1037/0033-295x.111.4.931

Zhai, F., Brooks-Gunn, J., & Waldfogel, J. (2011). Head Start and urban children's school readiness: A birth cohort study in 18 cities. *Developmental Psychology*, *47*(1), 134–152. http://dx.doi.org/10.1037/a0020784

Zhang, J.-S., Wang, Y., Cai, R.-G., & Yan, C.-H. (2009). The brain regulation mechanism of error monitoring in impulsive children with ADHD-An analysis of error related potentials. *Neuroscience Letters*, *460*(1), 11–15. https://doi.org/10.1016/j.neulet.2009.05.027