VALIDATING NEURAL MARKERS OF EFFORTFUL CONTROL IN YOUNG CHILDREN

By

Sharon Lee Lo

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Doctor of Philosophy

2017

ABSTRACT

VALIDATING NEURAL MARKERS OF
EFFORTFUL CONTROL IN YOUNG CHILDREN

By

Sharon Lee Lo

Effortful control (EC) in early childhood is a robust predictor of later mental health, social-emotional adjustment, and academic performance. EC is defined as the ability to regulate behavioral and emotional reactivity, and develops most rapidly during preschool-age years, a time recognized as a critical period to intervene and prevent later psychopathology. However, we have little to no knowledge of the basic etiological processes underlying EC in young children, which precludes the development of effective interventions based on mechanisms of action. Mastery of EC skills is reliant on the successful integration of brain processes underlying neural and behavioral systems, and event-related potentials (ERPs) represent a confluence of these systems, making it an ideal tool to understand the neurobehavioral underpinnings of EC in children. Two common conflict-monitoring ERPs known as the error-related negativity (ERN) and N2 are thought to index aspects of EC in children, but the nature and function of these markers is not understood.

The present study used a short-term targeted training of effortful control (EC) as an experimental manipulation to address whether (1) the ERN and N2 are valid markers of EC in young children (aged 4-5 years), and (2) whether training-induced changes in EC as indexed by behavioral and neural measures are associated with changes in social-emotional adjustment problems. While results across behavioral, neurophysiological, and informant-reported methods indicated that EC skills were not substantially impacted by targeted EC training, they also suggested that the ERN and N2 may be dynamic measures of EC skills in young children, and may reflect similar processes in this age period. Future research would benefit from developing measures that

are more sensitive to detecting changes in EC skills during this age period in order to better understand the coherence between brain and behavior processes that support the maturation of EC.

This manuscript is dedicated to the Lo family.
Thank you for always encouraging me to be myself and to do my best.

ACKNOWLEDGEMENTS

I wish to thank my mentors Dr. C. Emily Durbin and Dr. Jason S. Moser for providing guidance and direction throughout my graduate school career, and comments on drafts of this manuscript. I greatly appreciate the research assistants, children, and families who made data collection for the present investigation possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Individual differences in effortful control (EC) are implicated in the development of psychopathologies that have detrimental personal and societal costs (Caspi, 1996; Moffitt et al., 2011). They are more strongly associated with school readiness as compared to IQ and initial academic competencies (Blair & Razza, 2007). EC skills develop most rapidly during early childhood (Rothbart, 1989) and involve emotional and cognitive control processes that regulate behavioral and emotional reactivity (Rothbart, Ellis, Rueda, & Posner, 2003). Variation in EC skills in early childhood is a robust predictor of major life outcomes in adulthood related to physical and mental health, financial status, crime conviction, and academic achievement, even after controlling for IQ and social class (Moffitt et al., 2011). These clear public health implications of targeting EC in childhood have placed pressure on researchers to design effective interventions based on basic mechanisms underlying EC development. Although evidence has shown that mastery of EC skills is dependent on brain processes and integration of underlying behavioral and neural systems (Whittle, Allen, Lubman, & Yücel, 2006), this area of research is underdeveloped in young children. Two event-related potentials (ERPs) assessing conflict-monitoring aspects of EC, the Error-Related Negativity (ERN) and N2, have been assessed in children (Yeung, Botvinick, & Cohen, 2004), but the nature and function of these neural markers is not well understood. The present study aimed to use an experimental manipulation to test the functional validity of proposed neurophysiological markers of EC in young children.

Effortful Control in Childhood

The temperament and cognitive neuroscience literatures have created a number of terms to describe the ability to control and regulate thoughts, emotions, and behaviors. Broadly, regulation occurs through voluntary control such as purposefully inhibiting a response or via non-voluntary means such as fear. Self-regulation refers to the ability to modulate reactivity, and is often used in developmental psychology alongside EC. EC is one of the fundamental dimensions of temperament defined as the capacity to regulate reactive processes, such as emotion, and to purposefully suppress a predominant response and execute a subordinate one (Rothbart, Ellis, Rueda, & Posner, 2003). EC is thought to involve two important facets of self-regulation known as attentional control and inhibitory control. Attentional control refers to the ability to effectively allocate one's attention in a flexible manner while inhibitory control refers to the capacity to regulate and resist pre-potent behavioral impulses or responses.

"Executive attention", "executive control", "cognitive control", and "self-control" are terms commonly used in neurocognitive research that refer to a critical component of a broader construct known as executive functioning. In an attempt to bridge developmental and cognitive neuroscience literatures, the terms "attentional control", "inhibitory control", and "effortful control (EC)" are used. While these terms may have different origins compared to the terms used in the cognitive literature, there does not appear to be a firm distinction between these measures of EC given that they all involve the ability to control dominant responses in favor of selecting less dominant ones. The integration of neural networks responsible for modulating cognition and emotion are thought to play a role in attentional and inhibitory control processes; and the basic link between these processes is the primary contributor to adaptive EC functioning.

Dimensions of effortful control. Increasing evidence suggests that attentional control and inhibitory control represent two separate dimensions of EC (e.g., Carver et al., 2008; Oldehinkel, Hartman, De Winter, Veenstra, & Ormel, 2004), and are found to differentially predict behavior problems and academic performance (Kim, Nordling, Yoon, Boldt, & Kochanska, 2012). However, the ways in which attentional and inhibitory control vary as a function of temperamental reactivity and influence behavioral and social emotional outcomes in children are largely unexplored. Some reviews indicate that attentional and inhibitory control mechanisms may develop over different time courses (see Diamond, 2013 for a review), but no hypotheses have been tested as to how these developmental trajectories may differ.

Correlates of behavioral measures of effortful control. Given the importance of early EC in later life outcomes, researchers have been interested in examining the developmental course of EC, and its contribution to social-emotional and psychological development. Within the first two years of life, regulatory behaviors emerge that are important to the development of EC such as complying to demands and purposefully inhibiting reactive motor responses (Kopp, 1982). EC becomes increasingly coherent across behavioral EC performance from 2 years of age and beyond; and expected gender differences (i.e., girls demonstrating higher EC compared to boys) in behavioral EC performance is observed across a greater number of tasks at 3 years of age (Kochanska et al., 2000). Following toddlerhood, the preschool years are characterized by considerable development in the ability to more flexibly engage in regulatory behaviors to meet different situational demands, and adjust behavior appropriately to changing demands. Children in this developmental period are able to not only inhibit a predominant response, but also execute the subordinate action, whereas their younger peers may only be able to effectively inhibit the dominant action.

Children in the preschool years also tend to experience a substantial increase in social interactions, and their ability to exert control to modulate their automatic behavior and emotional reactivity in these situations has important implications for their social-emotional development. Higher EC in preschool is related to greater social competence, conscience, and empathy (see Eisenberg, Smith, Sadovsky, & Spinrad, 2004). EC may also support socialization by allowing children to attend and respond to the emotional distress of others while regulating their own emotional reactivity. Data from parent- and teacher-reported EC suggests that preschoolers rated higher in attentional control were more likely to cope with anger using non-hostile methods rather than overt aggression (Eisenberg, Fabes, Nyman, Bernzweig, & Pinuelas, 1994). Young children with higher levels of EC exhibit less intense positive and negative emotionality on behavioral tasks designed to elicit variation in emotional reactivity (Kochanska et al., 2000), suggesting that higher EC is related to more effective regulation of both positive and negative affect.

With respect to the role of EC in predicting psychological adjustment outcomes, studies of EC in adults indicate that low EC is associated with both internalizing and externalizing problems (Carver, Johnson, & Joormann, 2008), but this relationship is not as well understood in children. Evidence from both parent-reported and laboratory-assessed temperament measures suggest that high EC may be related to fewer concurrent and later externalizing behaviors (e.g., Kochanska & Knaack, 2004; Wachs & Bates, 2001). Some studies have found that high EC is associated with regulatory behaviors that allow the child to adapt and cope with frustration (Shoda, Mischel, & Peake, 1990), whereas other studies suggest that children with high EC are also high in guilt and shame, a predisposition for developing later depression and anxiety disorders (Rothbart, Ahadi, & Hershey, 1994). These differing associations hark back to the Blocks' constructs of ego control and ego resiliency (Block & Block, 1980), where ego control is defined as a person's typical tendency to restrain behavioral impulses, and ego resiliency is defined as a person's ability to flexibly engage ego control (i.e., reduce or increase such control) when called upon by differing contexts. Another hypothesis is that the association between EC and internalizing behaviors is dissociable by examining attentional control and inhibitory control separately. Specifically, evidence suggests high levels of inhibitory control in early childhood place children at higher risk for developing internalizing psychopathology whereas high levels of attentional control are protective against internalizing psychopathology (White, McDermott, Degnan, Henderson, & Fox, 2011).

Overall, the reviewed results demonstrate that EC is central to the social-emotional and psychological development from a young age. However, when EC is studied using only behavioral or informant-report measures, findings regarding its developmental trajectory and association with important outcomes do not provide enough evidence to elucidate other contributing mechanisms of EC development.

Neurophysiological Indices of Effortful Control in Youth

Evidence from neuroimaging studies and cognitive research points to the "executive attention network" as the neural network that monitors and helps resolve conflict by modulating neighboring attention networks, and this modulation is thought to underlie core functions of EC (Rothbart & Rueda, 2005; Rothbart, Sheese, & Posner, 2007). The lack of research on neurocognitive development in early childhood is a barrier to understanding the behavioral and neurophysiological processes characterizing EC and its development, knowledge that is essential for developing interventions based on the underlying mechanisms associated with mental illness.

An important measure of EC is error commission. Attentional and inhibitory control skills are responsible for our ability to detect and subsequently correct errors. The ERN and N2 are event-related potentials (ERPs) that have been consistently recorded using electroencephalogram (EEG). The ERN follows the commission of an error whereas the N2 precedes a correct response following stimulus presentation and is enhanced on trials involving more conflict. Increasing evidence from source localization and neuroimaging studies suggest that the ERN and N2 are generated from the dorsal ACC (e.g., van Veen & Carter, 2002; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001), providing support that both ERPs may index the efficiency of the executive attention network. The ERN was initially identified in adults appearing as a negative deflection at frontocentral electrodes within approximately 100 ms of an error, and identified as a robust marker of processes related to error correction and suppression (Gehring, Liu, Orr, & Carp, 2012; Yeung & Summerfield, 2012). The ERN effect is assessed by comparing its amplitude following an erroneous response to the amplitude following a correct response. The negative deflection following a correct response is known as the correct-related negativity (CRN), and the difference between the ERN and CRN (ERN minus CRN) is known as the ERN difference (ΔERN).

Unlike the ERN, the N2 is a negativity elicited prior to a response and approximately 300 ms following stimulus presentation. The N2 has been examined using different visual and auditory modalities that produce several different N2 subcomponents. In the present study, I focused on the inhibitory visual N2 (henceforth referred to as the N2) that is maximal at frontocentral sites since it is more distinguishable compared to other N2 components elicited with auditory stimuli and is most often examined in relation to executive attention and EC-related skills. The N2 is typically larger or more negative on correct trials involving greater conflict such as successfully inhibiting a pre-potent response (van Veen & Carter, 2002) or on an incongruent versus congruent trial. The N2 elicited on high conflict trials may be referred to as the "No-Go N2" or "incongruent N2" whereas the N2 elicited on low conflict trials may be referred to as the "Go N2" or "congruent N2". The difference between the high conflict N2 and low conflict N2 (high conflict N2 minus low conflict N2) is known as the ΔN2, and is a measure of the N2 effect between high and low conflict trials. For the purposes of the present study, I will refer to N2 elicited on high conflict trials as the "high conflict N2" and N2 elicited on low conflict trials as the "low conflict N2".

Current research suggests that the ERN and N2 can be elicited in children as young as 3 years of age (Espinet, Anderson, & Zelazo, 2012; Grammer, Carrasco, Gehring, & Morrison, 2014; Lo, Schroder, Moran, Durbin, & Moser, 2015), and differences in latency and amplitude have been reported in children compared to adults (Davies, Segalowitz, & Gavin, 2004). However, despite the increase in studies examining the ERN and N2 in children, very little is understood about the nature and function of these neural markers in development, which precludes our ability to understand the extent to which these differences might characterize changes in EC.

Correlates of neurophysiological measures of effortful control. Early studies on the ERN and N2 in children have primarily focused on clinical populations and samples at high risk for developing psychopathology. The most commonly studied population with EC deficits includes children with attention deficit hyperactivity disorder (ADHD), a disorder characterized by developmentally inappropriate levels of inattention and/or hyperactivity and impulsivity that cause significant impairment in daily functioning (DSM-5; American Psychiatric Association, 2013).

ERN studies examining error processing in children with clinically significant deficits in EC suggest that initial error detection and later error reflection may be diminished. Greater support for a reduced ERN in populations with EC deficits is consistent with the hypotheses derived from the conflict monitoring theory, which states that difficulties maintaining attention toward target stimuli reduces the contrast between error and correct trials, resulting in a smaller ERN (Albrecht et al., 2010; Groen et al., 2008; Groom et al., 2013; Rosch & Hawk, 2013; van Meel, Heslenfeld, Oosterlaan, & Sergeant, 2007). However, there are a number of studies reporting either no group differences or larger ERNs in children with ADHD compared to healthy controls. For example, a study examining the ERN in children with ADHD and comorbid conditions, reported that children with ADHD exhibited a significantly larger ERN compared to healthy controls (Burgio-Murphy et al., 2007). It is notable that each study differed in their use of task parameters (e.g., task type, number of trials, task difficulty, task feedback, instructions, task timing) and ERP scoring methods (i.e., mean versus peak amplitude), which introduces a range of methodological issues that likely contribute to the variability in results (Shiels & Hawk, 2010). Additionally, the heterogeneity of symptoms associated with ADHD may contribute to the inconsistencies in the literature regarding the mean-level differences in the ERN between ADHD populations and controls. Specifically, children with ADHD may exhibit symptoms that are more consistent with a predominantly inattentive profile, a predominantly hyperactive/impulsive profile, or both.

Research on the N2 in clinical populations with deficits in EC began several decades earlier compared to research on the ERN in these populations, and is therefore more extensive. However, findings on N2 variation between clinical and healthy populations are highly inconsistent. Several studies have reported a reduced N2 effect in children with ADHD (e.g., Brandeis et al., 1998; Broyd et al., 2005; Johnstone, Watt, & Dimoska, 2010; Pliszka, Liotti, & Woldorff, 2000), Autism Spectrum Disorder (Tye et al., 2014), and unaffected siblings of children with ADHD (Albrecht et al., 2010) whereas others have reported no N2 differences between clinical populations with EC deficits and healthy controls (e.g., Cao et al., 2013; Fallgatter et al., 2004; Johnstone & Galletta, 2013; Rosch & Hawk, 2013; Spronk, Jonkman, & Kemner, 2008). In addition, there are studies reporting the opposite effect, a larger N2 amplitude in children with ADHD compared to healthy controls (Jonkman, van Melis, Kemner, & Markus, 2007; Senderecka, Grabowska, Szewczyk, Gerc, & Chmylak, 2012; Smith, Johnstone, & Barry, 2004). (See reviews by Barry, Johnstone, & Clarke, 2003; Johnstone, Barry, & Clarke, 2013).

Anxious and depressive populations have also been of interest to researchers given that higher negative emotionality and difficulty regulating such affect is related to poor EC skills in early childhood. The ERN is hypothesized to be larger in adults who must demonstrate increased effort to compensate for the anxious apprehension that interferes with their ability to reengage with task goals following an error (Moser, Moran, Schroder, Donnellan, & Yeung, 2013; Moser et al., 2014). Other accounts theorize that larger ERNs in adults reflects trait-like differences in threat sensitivity. Specifically, clinically significant levels of anxiety are characterized by an increased defensive response to errors and that errors are threatening (Proudfit, Inzlicht, & Mennin, 2013). Overall, studies support an association between a larger ERN and greater anxiety in older children (i.e., Hajcak, Franklin, Foa, & Simons, 2008; Hum, Manassis, & Lewis, 2013; Ladouceur et al., 2006) and in adolescents who were behaviorally inhibited as children (McDermott et al., 2009). However, there is some inconsistency between studies on observing this association using parent-reported measures of anxiety (Bress, Meyer, & Hajcak, 2013; Santesso, Segalowitz, & Schmidt, 2006). Some of this inconsistency may arise from limitations of parent-reported measures that introduce more error, such as the influence of parent characteristics and psychopathology on parent-reported measures (Durbin & Wilson, 2012). Importantly, recent evidence suggests that the relationship between the ERN and anxiety may change over the course of development based on reports that the relationship is opposite in younger children (Lo et al., 2015; Meyer, Weinberg, Klein, & Hajcak, 2012; Torpey et al., 2013).

Fewer studies have investigated the relationship between N2 and affective processes in children. There is some evidence to suggest that children with anxiety disorders exhibit larger N2 amplitudes on high conflict trials (Henderson, 2010) whereas others report that these children do not demonstrate an N2 effect compared to their healthy counterparts such that anxiety disorder groups exhibit a similar N2 response regardless of trial type (Hum et al., 2013). Researchers have hypothesized that a larger N2 indicates that anxious children may allocate more attentional resources toward emotion regulation. Very few studies have examined both the ERN and N2 within the same sample, and of those that did (Hum et al., 2013; Ladouceur, Conway, & Dahl, 2010), researchers did not test the extent to which the N2 and ERN may index similar or different EC processes.

Attempting to draw conclusions and make inferences about the functional significance of the ERN and N2 in the studies reviewed above is severely limited by the field's inadequate understanding of the basic developmental changes in these processes across childhood and adolescence. Rather than a review of individual findings across different age groups I conducted a meta-regression analysis on existing studies to evaluate the developmental trajectory of the ERN and N2 (Lo *under review*). A meta-regression analysis of 26 studies on the ERN (N = 1, 519) indicated a 0.02μV increase in ERN amplitude with 1-month increase in age, and of 19 studies on the N2 on high conflict trials ($N = 1, 095$) revealed a 0.02μV reduction in high conflict N2 amplitude with 1-month increase in age. These analyses provide initial support for the hypothesis that the ERN and N2 may measure distinct facets of EC, and that dissociable effects of the ERN and N2 related to EC can be observed in childhood and adolescence. The increase in the ERN with age is consistent with the notion that the ERN is a measure of attentional control. As these skills increase with age, the ability to attend to target stimuli increases, which likely increases the discrepancy between error and correct trials when continued processing post-response occurs and the ERN is produced. The N2 amplitude on high conflict trials decreased with age, which is consistent with the notion that the N2 is a measure of inhibitory control. As inhibitory control skills improve with age, the ability to inhibit motor responses and ignore distracting information increases, which decreases conflict between irrelevant and target stimuli. The diminished conflict between irrelevant and target stimuli leads to a reduced N2 on high conflict trials.

Targeted Training and Intervention Studies

Changes in EC arising from brain maturation and associated changes in the ERN can also be understood through researchers' recent efforts to improve EC skills using training and intervention. Much of research on interventions that aid EC development originate from the neurocognitive development literature and target broader skills related to EC. The most common approach is using a computerized working-memory training known as CogMed (Pearson Education, Upper Saddle River, NJ), which uses computer games that place increasingly greater demands on working memory as the child progresses through the training program. There are several published studies evaluating the training effects of CogMed on executive functioning skills. Of these studies, children included in these training studies are either typically developing (Bergman-Nutley et al., 2011; Thorell, Lindqvist, Bergman-Nutley, Bohlin, & Klingberg, 2009), diagnosed with ADHD (Holmes & Gathercole, 2010; Klingberg et al., 2005), or have low working memory span (Holmes, Gathercole, & Dunning, 2009).

On average, the training program occurs across the course of 6-7 weeks and the amount of training ranges between 6-17 hours across the entire study. It's important to note that their findings, in addition to other training studies, suggest transfer to other cognitive and EC-related skills is quite narrow. In other words, training benefits often do not generalize to performance on unpracticed tasks (Diamond & Lee, 2011; Sala & Gobet, 2017). Other training and intervention programs include computer training targeting attention (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005), aerobic exercise (e.g., Drollette et al., 2014a; Hillman, Buck, Themanson, Pontifex, & Castelli, 2009; Pontifex, Saliba, Raine, Picchietti, & Hillman, 2013; Tuckman & Hinkle, 1986), martial arts and mindfulness (e.g., Flook et al., 2010; Lakes & Hoyt, 2004). The behavioral and cognitive effects of these programs are reviewed elsewhere (Diamond & Lee, 2011), and therefore only studies that included neurophysiological measures will reviewed below.

Only one training study to date has investigated the effect of training skills related to EC on the ERN. The study conducted was a brief aerobic exercise intervention based on increasing evidence that physical exercise improves a EC-related skills, which some have suggested reflects the beneficial effects of physical activity on prefrontal cortex functioning (Hillman, Erickson, & Kramer, 2008; Sibley & Etnier, 2003). Children aged 8 to 10 years who were diagnosed or suspected to have ADHD and children selected as healthy controls engaged in either 20 minutes of moderate-intensity aerobic exercise or 20 minutes of seated reading (Pontifex et al., 2013). All children exhibited enhanced performance on academic achievement tests following exercise as compared to the seated condition. Following the seated reading condition, children with ADHD exhibited a reduced ERN relative to healthy controls. However, following the single bout of exercise, there were no group differences in the ERN. An exercise-induced increase in the ERN was only observed in children with ADHD, suggesting that children with deficits in EC skills may be more likely to exhibit exercised-induced changes in the ERN.

Of the studies examining the effects of training on the N2 amplitude, a majority have focused on using computer-based training on performance of the same task or closely related tasks that tap similar EC skills. However, training conditions between these studies are variable. For instance, one study focused on testing the hypothesis that performance on a single behavioral task (the Dimensional Card Change Sort; DCCS) is dependent on the child's reflective processing, and compared the influence of reflective training on N2 amplitude in 2- to 4-year olds (Espinet, Anderson, & Zelazo, 2013). Reflection training included engaging the child in step-by-step processing of how to complete the DCCS (i.e., asking the child to first name the relevant rule or dimension, providing a correct example, and then assisting the child in choosing the correct sort). Results indicated significant reductions in the N2 amplitude following training compared to practice conditions without reflection feedback. In another training study, investigators tested the impact of a 5-day attention program over the course of 2-3 weeks using various computer tasks on 4-6 year olds (Rueda et al., 2005). Results suggested that training produced EEG patterns in young children that looked more similar to older age group counterparts. A follow-up study extended the training program to 5 weeks and reported no differences in N2 amplitude between condition or time point (Rueda, Checa, & Cómbita, 2012). However, authors suggested that evidence of an earlier and more posterior N2 in the training group reflected more efficient detection of conflict given that adults also exhibit a more posterior N2 effect.

The reviewed studies highlight the inconsistency of researchers' interpretations of which N2 characteristics (amplitude, latency, and source localization) index EC skills. The lack of agreement makes it difficult to compare the results across studies and formulate a priori hypotheses in future studies. Additionally, it is challenging to interpret the functional meaning of these changes in neurophysiological indices when behavioral measures are not included or do not appear to be associated with behavioral measures of EC. However, these investigations support an important notion that the ERN and N2 are plastic during this developmental period, and demonstrate that a brief experimental manipulation designed to improve EC skills has the potential to alter both behavioral and neural measures of EC.

PRESENT STUDY

A short-term training targeting EC provided a novel experimental manipulation for explicitly validating brain measures thought to reflect EC. The approach had two advantages: (1) EC was directly manipulated over a short time period and compared to a passive control group, affording the control of environmental and maturation effects that are not controlled in naturalistic longitudinal studies; (2) findings could inform prevention and intervention development based on neural mechanisms of change. A multimethod approach was used to integrate informant-reported, behavioral, and neurophysiological measures of training-induced changes in EC. Some studies in adults suggest the ERN and N2 share similar underlying conflict monitoring processes (Yeung & Cohen, 2006), but others suggest they are elicited by distinct processes (Mathalon, Whitfield, & Ford, 2003). Even less is known about the ERN and N2 in children, and no known study has examined their convergence. This present investigation was the first to use an experimental manipulation to test the validity of the ERN and N2 as markers of EC, and to use multiple measures of these markers and their associated behavioral correlates to help further isolate the neurobehavioral processes that may underlie training-induced changes in EC.

HYPOTHESES

Demonstrating near transfer effects is the first step to ensuring that any improvement detected is associated with the targeted training rather than with repeated practice, in addition to helping isolate other neural and behavioral processes that may be responsible for changes in far transfer tasks (Shipstead et al., 2012). Therefore, I anticipated that children in the training group would exhibit significantly larger improvements on near transfer behavioral tasks that were practiced during training compared to children in the control group. In terms of neurophysiological measures, I anticipated that children in the training group would exhibit larger ERNs compared to the control group given research suggesting that the ERN increases with age and reflects improvement in self-regulation skills (e.g., Davies, Segalowitz, & Gavin, 2004; Lo, *under review*). A handful of studies have found a reduced N2 amplitude in children following trainings targeting broader executive functioning skills (Espinet et al., 2013; Rueda et al., 2005). Thus, I expected a reduced N2 amplitude in the training group, which would suggest that the ERN and N2 may mark different EC processes (e.g., attentional and inhibitory control). Alternatively, if the ERN and N2 show changes in the same direction, this would suggest they index similar processes. The reviewed literature suggests that EC plays an important role in predicting psychological adjustment outcomes. Therefore, I expected that children participating in targeted training would have greater reductions in parent-reported adjustment problems, and lower mean-levels of teacher-reported adjustment problems at follow-up compared to control.

These hypotheses (see Table C1) were tested further by exploratory analyses examining associations between changes in neural, behavioral, and informant-reported measures of EC to disentangle the overlapping and distinctive facets of EC. I anticipated that children exhibiting the largest changes in the ERN and N2 would also demonstrate the largest changes in behavioral measures of EC such that increases in the ERN would be associated with improvement in behavioral measures of EC, and reduction in the N2 would be associated with improvement in behavioral measures of EC. Similarly, increases in the ERN (and decreases in the N2) would be associated with reduction in informant-reported EC and social-emotional problems.

METHODS

Behavioral and neural processes associated with changes in EC following a short-term intervention in a sample of preschool-age children were investigated. Multiple methods (informant-report, behavioral, neurophysiological) and passive control design were employed to assess training-induced changes and provide preliminary tests of mechanisms involved in training-induced changes. The Michigan State University Institutional Review Board approved all procedures for this research study entitled "Training Effortful Control in Young Children" (Protocol #15-476).

Sample

Children between 4 and 5 years of age who have not yet entered kindergarten from the mid-Michigan area were recruited across three consecutive summers (see Table A1), with equal number of girls and boys. Eligible participants were selected into training or passive control group. In order to ensure adequate sample size in the training group, children and families who were determined eligible first were selected to participate in the training group. Participants in the passive control group were selected to match the age and gender of their counterparts in the training group. Participants in Wave 1 were drawn from a larger longitudinal investigation of child temperament (N = 277) and selected if they completed the neurophysiological portion of the study (see Lo, Schroder, Moran, Durbin, & Moser, 2015). The wave 1 training/control cohort was run in Summer of 2014.

Participants in Wave 2 were recruited using commercial mailing lists and postings on classified advertisement websites. Eligible families completed a parent-report measure online to assess their child's EC skills. To counteract potential ceiling effects reported by previous studies exploring EC training (Klingberg et al., 2005; Tominey & McClelland, 2011), a subset of children who had low-to-moderate levels of EC were recruited for the study. A total of 139 participants completed the survey (42 in Sample 2, and 97 in Sample 3). A total of four training/control cohorts were run for Wave 2 (first cohort in Summer of 2015, and final three cohorts in Summer of 2016).

Each training cohort consisted of 5-6 children. The final sample from Wave 2 included 21 children in the training group and 17 in the passive control group.

The final sample of children (n = 48) had a mean age of 5.14 years (*SD* = 0.37, range: 4.50-5.92). Data on race/ethnicity, marital status, and family income were provided by 89.6% of parents. Of those, the racial and ethnic composition of the child participants was as follows: White European American (74.4.7%), African American (9.3%), bi-racial/multi-racial (9.3%), Hispanic/Latino (7.0%). Fifty percent of identified bi-racial and multi-racial children were endorsed as Asian and White European American, and fifty percent endorsed as Hispanic/Latino and White European American. Yearly family income ranged from $10,000 to greater than $100,000; 14.0% reported income below $20,000, and 20.9% less than $41,000. Highest level of education endorsed by parent ranged from some college to professional doctorate. The education level of parents was as follows: some college (16.3%), technical school (7.0%), Associate's degree (16.3%), Bachelor's degree (18.6%), Master's degree (32.6%), and professional doctorate (9.3%). A majority of parent participants reported their marital status as married (76.7%) while other participants reported either a status as single/never married (11.6%), living with a partner (7.0%), or separated (4.7%).

Inclusion/exclusion criteria. Children in both Wave 1 and 2 were eligible to participate if they resided with at least one English-speaking parent, and were free of significant medical conditions or developmental disabilities that prevented them from completing study assessments. Children were also eligible if they have no known history of epilepsy, head trauma resulting in a loss of consciousness for more than five minutes, and hearing, visual, or physical disabilities that could cause difficulties understanding and/or using a computer. Wave 2 caregivers completed an online screener questionnaire including items from the Behavior Rating Inventory of Executive Function-Preschool Version (BRIEF-P; Gioia, Espy, & Isquith, 2002). Children were eligible to participate in the study if their z-score on the BRIEF-P fell between -1 and 2 based on published normative means for this measure.

Timeline and assessments. At baseline, children were assessed 1-2 weeks prior to completing targeted EC training. Baseline assessments included: 1) 2-hour lab assessment consisting of behavioral measures of EC and 2) a separate 2-hour lab visit completing neurophysiological assessment and measures assessing general cognitive skills and academic achievement. During baseline assessments, parents completed questionnaires assessing their child's social-emotional adjustment. In order to ensure adequate sample size in scheduled the training group, children were selected into the training group on a rolling basis and children selected into the passive control group were age- and sex-matched to their training group peers. Children selected into the targeted training group completed a 5-day targeted EC training. The following week, children completed one 2.5-hour post-training assessment that included behavioral and neurophysiological assessments of EC. The same set of questionnaires completed at baseline were re-administered to parents and children's teachers at 1 month and 3 months following targeted training. Children participating in the passive control group were administered the same assessments, but did not participate in any form of training in between assessment time points.

Measures

Behavioral measures of EC. Children completed a set of tasks (Appendix A) designed to elicit behaviors and emotional responses indicative of individual differences in EC skills and traits, some of which were drawn from the NIH Toolbox for Executive Functioning (Zelazo et al., 2013) (e.g., the Dimensional Change Card Sort; DCCS) and others from standardized temperament batteries (Goldsmith, Reilly, Lemery, Longley, & Prescott, 1995). Tasks were chosen that either assess attentional control skills (e.g., Tower of Hanoi, Shape Stroop) and/or motor inhibition and skills related to inhibitory control (e.g., Walk-a-Line Slowly, Simon Says, Hand Game). Few investigations have evaluated developmental sensitivity of behavioral measures of attentional and inhibitory control skills in preschool-aged children. Therefore, a range of different tasks of varying difficulty (Carlson, 2005) were chosen to develop a comprehensive battery. Parents remained outside of the assessment room and were able to watch their child complete the set of tasks via a live video stream. In between each task, a short play break was used to allow children to return to a baseline affective state. Unless otherwise noted below, higher scores on these behavioral tasks were indicative of better EC skills. Each task is described in detail below.

Behavioral measures of attentional control skills.

Digit Span (Weschler, 2003). Children completed the digit span subtest of the Weschler Intelligence Scale for Children – Fourth Edition (WISC-IV). First, the child was asked to repeat increasingly longer digit strings exactly as read by the examiner. If the child incorrectly repeated both trials of a string length, then the task was discontinued. Next, the child was asked to repeat the digits as read by the examiner but backwards. The same discontinue rules were applied to the backward digit span administration.

Reverse Categorization (Carlson, Mandell, & Williams, 2004). The child was asked to help the experimenter sort small and large blocks into a small and large bucket. First, the experimenter demonstrated placing the small blocks into the small bucket and large blocks into the large bucket in a pseudorandom order for 6 trials. The child was given a turn for 6 trials, similar to the 6 pre-switch testing trials in the DCCS. The child was then instructed to play a "silly game" with the experimenter in which the small blocks were sorted into the large bucket, and the large blocks sorted into the small bucket. Twelve post-switch test trials were administered.

Dimensional Change Card Sort (DCCS; Zelazo, 2006). The DCCS is a common measure sued to assess EC skills in children. The child was instructed to sort two bivalent target cards (e.g., blue rabbit and red boat) based on one dimension (e.g., color). After 6 pre-switch trials, the experimenter instructed the child to sort the target cards based on the second dimension (e.g., shape). If the child correctly identified 5 out of the 6 post-switch trials, the experimenter administered an additional version of the DCCS where target cards with a black border followed sorting rules on one dimension and target cards without a black border followed sorting rules on the second dimension.

Day/Night (Gerstadt, Hong, & Diamond, 1994). The child was shown two different cards – a white card with a yellow sun, and a black card with white stars and moon. The child was instructed to say "day" when presented with the black card and "night" when presented with the white card.

*Shape Stroop (Kochanska et al., 2000)***.** The child was presented with three different fruits (i.e., banana, apple, orange) in two sizes (i.e., small and large). The experimenter asked the child to identify the fruit and size of each. Next, the child was presented with two images – both of which depicted a different small fruit was embedded within a larger fruit (e.g., small apple inside a large orange, and small banana inside a large apple). The child was asked to identify either the small or large version of the fruit (e.g., "show me the *small* apple"). Twelve testing trials were administered and children received a score of 1 for correct responses, 0 for incorrect responses.

Flexible Item Selection Task (FIST; Jacques & Zelazo, 2001). The child was shown a set of 3 pictures and asked to select 2 pictures that matched each other on one dimension (i.e., size, shape, color) and then asked to select a different pair of 2 pictures that matched each other on a dimension different from the first. A total of 15 testing trials were administered. Children were not provided feedback during testing trials. On each testing trial, children who correctly selected the first match received a score of 1. Children received an additional score of 1 if they correctly selected the second match, but only if their first match was correct. This scoring system was developed by Jacques and colleagues (2001) in order to prevent score inflation (i.e., an incorrect match on the first trial would mean that either of the remaining two matches were correct). Scores on second matches (variable named "FIST Select 2") reflect the ability to flexibly shift between thinking about two different concepts (i.e., size vs. shape), which is an aspect related to attentional control skills.

Tower of Hanoi (Welsh, 1999). The Tower of Hanoi task adapted from the classic tower task (Simon, 1975) developed by Welsh (1999) to use with younger children was administered. Detailed administration instructions have been reported in the literature (Carlson, Moses, & Claxton, 2004), so a brief description will be provided here. The child was provided instructions on how to play the "Monkey Jumping Game", in which monkeys (wooden disks) could only jump one at a time and the larger monkey could not go on top of a smaller monkey. The child was required to move the monkeys from peg to peg in order to create the same monkey tower that was constructed in front of the experimenter. The task increased in difficulty in terms of the tower construction and number of monkeys involved. In total, there were six levels of difficulty and performance was scored as the highest level completed (0 to 6).

Yes/No (Wolfe & Bell, 2004). The child was instructed to say "no" when the experimenter shook her head yes (i.e., up and down), and to say "yes" when the experimenter shook her head no (i.e., side to side). Two practice trials were administered and included in the overall score if the child responded correctly on the first try, followed by fourteen testing trials for a total of a possible score of 16.

Behavioral measures of inhibitory control skills.

Walk-a-Line Slowly (Murray & Kochanska, 2002). The child was asked to walk on a line taped on the floor. Next, the child was asked to walk on the line again, but this time as slowly as possible on a line taped on the floor, and then as quickly as possible. The child was again instructed to walk on the line as slowly as possible. The length of time for each trial was recorded by two blinded research assistants, and inhibitory control skills on this task were operationalized as the difference between Slow Trial 1 (Baseline: $\alpha = 0.98$; Post-assessment: $\alpha = 0.98$) and Trial 1 (Baseline: $\alpha = 0.99$; Post-assessment: $\alpha = 1.00$).

Simon Says (Strommen, 1973). The child was required to perform various actions (e.g., touch head, touch nose) as instructed by the experimenter, but only if the command was preceded by the phrase "Simon Says". A total of 20 testing trials were administered.

Hand Game (Hughes, 1998). The child was instructed to imitate the hand gesture of the experimenter and then perform the opposite hand gesture. For example, when the experimenter points a finger, the child is instructed to make a fist, and vice versa. Children received a score of 1 if they successfully performed the opposite hand gesture, and a score of 0 if they failed to perform the opposite hand gesture. A total of 15 possible trials were administered. If the child received a score of 1 on 6 consecutive trials, then the task was discontinued and the child was given a score of 1 on the possible remaining trials.

Bear Dragon (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996). The child was introduced to two puppets, a "nice bear" and "naughty dragon", and asked to only follow commands when the bear speaks but not follow demands when the dragon speaks. The experimenter spoke in a high-pitched voice and moves the bear's mouth to state a specific movement (e.g., "touch your head") and then spoke in a low-pitched voice while moving the dragon's mouth saying the next movement (e.g., "touch your nose"). Ten practice trials (5 bear trials, 5 dragon trials in alternating order) were administered with feedback followed by 10 test trials (also administered in alternating order). The child was reminded of the rules after the first 5 trials regardless of performance. The child received a score of 1 if they responded correctly, and 0 if they responded incorrectly.

Knock-Tap (Klenberg, Korkman, & Lahti-Nuuttila, 2001). The child was required to imitate the experimenter's action (i.e., knocking or tapping on the table) for two practice trials, and then instructed to perform the opposite action. For example, when the experimenter knocks on the table, the child must tap the table. A total of 8 testing trials are administered and children received a score of 1 for correct responses, 0 for incorrect responses.

Snack Delay (Goldsmith et al., 1995). The child was instructed to wait until the experimenter rang a bell to eat a piece of candy that was placed under a clear cup. The experimenter followed a systematic series of 8 delay trials that ranged from 10 to 30 seconds. The number of prompts, errors, self-corrects, and touches were coded by two blinded research assistants (Baseline Prompts ICC = 0.99, Post-Assessment Prompts ICC = 0.69; Baseline Errors ICC = 0.99, Post-Assessment Errors ICC = 0.94; Baseline Self-Corrects ICC = 0.98, Post-Assessment Self-Corrects ICC = 0.64; Baseline Touches ICC = 0.89, Post-Assessment Touches ICC = 0.90). Prompts were defined as verbal requests by the child directed at the experimenter (e.g., "go", "do it", "ring it"). If the child ate the piece of candy, this behavior was coded as an error. If the child exhibited behaviors indicative of making an error but then correcting it (e.g., picking up the cup to reveal the candy and placing it back, picking up the candy and then returning it to the plate), this behavior was coded as a self-correct. Touches were defined as physical contact with task-relevant objects (e.g., hitting the bell, touching the experimenter's hand to ring the bell, displacing the cup). Behavioral codes were not mutually exclusive and behaviors could be double coded (e.g., picking up the cup and placing it back would be coded as a self-correct and a touch). A composite score was created by Z-scoring the behavioral codes and averaging the standardized values. The composite score was then reverse scored (multiplied by -1) so that higher composite scores reflected better EC skills on this task.

Head Toes Knees Shoulders (HTKS; McClelland & Cameron, 2012). This task required children to first respond as instructed to a command (e.g., "touch your toes"), then perform the opposite of the original instruction (e.g., touch their head when the command was "touch your toes"). In subsequent trials, additional commands for touching shoulders and knees were added to increase the complexity of the task. A total of 20 test trials were administered and children were given a score of 2 for a correct response, 1 for a self-corrected response, and 0 for an incorrect response.

Whisper Game (Carlson et al., 2004). The child was asked to whisper the names of 10 different popular cartoon characters (e.g., Dora, Big Bird). The experimenter spoke in a whisper throughout the game and instructed the child to whisper the color of the cartoon character if they did not know the character's name. Children were given a score of 0 if they shouted the response, 1 if they used a normal talking voice, and 2 if they whispered.

Transfer type. Near transfer measures included tasks that were practiced during training, and assessed at baseline and post-assessment (see Table D1). For intermediate transfer measures with visual stimuli (i.e., DCCS and Stroop), standard versions (i.e., red rabbits and blue boats; day/night) were used during baseline and post-training assessments, and alternative versions (i.e., yellow flowers and green trucks; grass/snow stroop) were used during targeted training practice.

Far-intermediate transfer behavioral tasks were those not practiced during training, and only administered at baseline and post-assessments. Behavioral performance on computerized tasks (e.g., accuracy, mean reaction time) on Go/No-Go and Flanker tasks were also as performance-based measures of EC at baseline and post-training. All performance-based measures of EC served as dependent variables for within- and between-subjects analyses.

Assessment of cognitive ability and academic achievement. Children completed the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007) to assess their receptive vocabulary. The Block Design and Matrix Reasoning subtests were administered from the Weschler Preschool and Primary Scale of Intelligence (WPPSI-III; Weschler, 2002) to assess their fluid reasoning skills. The Block Span subtest from the Stanford Binet Intelligence Scales for Early Childhood (Early SB5; Roid, 2005) was administered to assess visual-spatial working memory, and the Last-Word subtest was administered to assess verbal working memory. Academic achievement was assessed using the Letter-Word Identification and Applied Problems subtests from the Woodcock Johnson-III Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001). This procedure occurred at the end of the neurophysiological assessment visit.

Informant-reported measures. At baseline, demographic data and future enrollment in school were collected. Parents' responses from the screener questionnaire were used as the baseline measure for the BRIEF-P. The BRIEF-P is a standardized questionnaire that asks informants to rate items related to inhibitory control, attentional shifting, emotion regulation, and planning. Parents completed the Child Behavior Checklist (CBCL) to measure parents' perceptions of their child's psychological adjustment (Achenbach, 1991). This commonly used measure asks parents to rate how often their child exhibits certain behaviors indicative of internalizing problems (e.g., anxiety, depression) and externalizing problems (e.g., hyperactivity, disruptive behavior). The CBCL and BRIEF-P were administered at baseline, 1-month, and 3-month follow-up time points. Teacher versions of these questionnaires were administered at the 3-month follow-up time point to allow sufficient time for teachers to familiarize themselves with the child's behavior in the classroom. A total of 25 teacher-reported questionnaires were completed and used for analysis.

Neurophysiological assessment of EC. An experimenter guided the child through each step of the EEG set-up and electrode application. The parent was permitted to stay in the room to observe EEG electrode application and set up. After set up, parents waited outside of the testing room in an observation room where they could view their child completing tasks through a ceiling camera. The experimenter was present throughout testing, but sat behind the child out of their view and either provided encouragement to complete the task or performance feedback in between task blocks depending on the child's accuracy (see below for description). The children completed a total of two tasks on the computer that were randomly counterbalanced. The child was positioned 17 inches from a 21-inch computer monitor for each task.

Go/No-Go task. Children completed a picture version of the Go/No-Go task called the Zoo Game, which has been used on similar samples as the present study (Grammer et al., 2014). Children were asked to help a zookeeper capture zoo animals that had escaped from their cages.

Children were presented with images of three orangutans who were helping the zookeeper and therefore did not need to be put back in their cages. The child was instructed to press the spacebar quickly and accurately to each animal (Go stimuli) except when the animal was an orangutan (No-Go stimuli), in which case the child was to withhold pressing the spacebar. On each trial, a stimulus of a colorful zoo animal was presented at a central location on the computer monitor. A fixation cross appeared before the stimulus, which accurately to each animal (Go stimuli) except when the animal was an orangutan (No-Go stimuli), in which case the child was to withhold pressing the spacebar. On each trial, a stimulus of a colorful zoo animal was presented at a central location on the computer monitor. A fixation cross appeared before the stimulus, which remained on the screen for 750 ms. Inter-trial intervals (ITI) were set to 500 ms.

The task began with a brief practice block, which consisted of 12 trials (9 Go trials and 3 No-Go trials). The practice block was repeated until the child demonstrated an understanding of the task. After the practice block, children completed 8 blocks that consisted of 40 trials (30 Go trials and 10 No-Go trials), totaling to 320 trials and lasting approximately 20 minutes. Novel sets of animal images balanced for animal size, color, and type were used in each block. Children were given performance feedback after each block of the task that was either related to making too many errors such as, "Remember to watch out for the orangutan friends", or not enough errors such as, "Remember to try and catch the animals even faster next time!" The performance feedback prompts provided for the children was determined by their accuracy in the preceding block, which was automatically calculated to help yield error rates higher than 10% but lower than 35% to ensure adequate accuracy rates and number of useable error trials for stable error-related waveforms. Before the beginning of the task and after blocks 2, 4, 6, 7, and 8, feedback was also provided using a "Zoo Map", which allowed children to track their progress in the task.

Flanker task. A developmentally appropriate version of the Flanker task (Lo et al., 2015) adapted from Rueda, Posner, Rothbart, and Davis-Stober (2004) was administered to children. Flanker stimuli consisted of 5 yellow cartoon fish swimming to the left or right on a blue background (see Figure B1). The child was instructed to focus on responding to the swimming direction of the middle fish/central target stimulus while ignoring the flanking fish stimuli. The task began with a practice block of 20 trials, including 5 congruent left trials (all fish facing to the left), 5 congruent right trials (all fish facing to the right), 5 incongruent left trials (middle fish facing left and flanking fish facing right), and 5 incongruent right trials (middle fish facing right and flanking fish facing left). A fixation cross appeared before each stimulus, which remained on the screen for 750 ms. ITI varied randomly between 700-1200 ms.

The practice block was repeated until the child understood the task. After the practice block, children completed 7 blocks that consisted of 20 trials (5 of each trial type as in the practice block), for a total of 140 trials lasting approximately 15 minutes. Similar to the Go/No-Go task, children were given performance feedback after each block of the task that was either related to making too many errors such as, "Pay attention to Goldie", or not enough errors such as, "Try to go even faster next time!" The performance feedback prompts provided for the children was determined in the same manner described in the Go/No-Go task to ensure adequate accuracy rates and number of useable error trials for stable error-related waveforms. Before the beginning of the task and after blocks 1, 3, 5, and 7, feedback was also provided using a "Treasure Map", which allowed children to track their progress.

Neurophysiological Recording and Data Reduction

All EEG recordings were taken from 64 Ag-AgCl electrodes using the Active Two Biosemi System (BioSemi, Amsterdam, The Netherlands). For EEG data acquisition, electrodes were placed in a stretch-lycra cap according to the 10/20 system with two additional electrodes placed on the left and right mastoids. Electrooculogram activity from eye movements and blinks were recorded at FP1 and three additional electrodes placed 1 cm from the pupil, one placed directly beneath the left pupil and the remaining two placed on the left and right outer canthi. In accordance with BioSemi's design specifications, the Common Mode Sense active electrode and Driven Right Leg passive electrode served as the reference during data acquisition. All EEG signals were digitized with a sampling rate of 512 Hz using ActiView software (BioSemi).

EEG data were re-referenced to the numeric mean of the mastoids and band-pass filtered with cutoffs of 0.1 and 30 Hz (12 dB/oct rolloff). All trials were also corrected for eye movements and blinks using the method developed by Gratton, Coles, and Donchin (1983). A computer-based algorithm was used to detect physiological artifacts such that individual trials were rejected if there was a voltage step greater than 50 μV between sampling points, a voltage difference of more than 200 μV within a trial, or a maximum voltage difference less than 0.5 μV within a trial.

Time-locked error-related event-related potentials. Based on visual inspection of the grand average ERN and previous published reports of the ERN in young samples (Grammer et al., 2014; Lo et al., 2015), the ERN in the Go/No-Go and Flanker tasks were quantified using average amplitude measures relative to a pre-response baseline of -150 to -50 along the midline (Fz, FCz, Cz, CPz, and Pz). The mean amplitude of the ERN was computed on error trials in a window 50 ms around the peak following the response. The CRN was computed on correct response trials. The ERN difference (ΔERN) was calculated by subtracting the CRN from the ERN waveforms.

Stimulus-locked event-related potentials. Based on visual inspection of the grand average N2 and published reports of the N2 in children (Pontifex et al., 2011; Rueda et al., 2005), the N2 in the Go/No-Go and Flanker tasks were quantified using average amplitude measures relative to a pre-response baseline of -100 to 0 along the midline (Fz, FCz, Cz, CPz, and Pz). The mean amplitude of the high conflict N2 component was evaluated as the largest negative going peak within a 325 to 550 ms latency window for the Go/No-Go task on No-Go trials and 250 to 550 ms latency window for the Flanker task on incongruent trials in order to capture the full negative going waveform. The low conflict N2 was the same measure, but computed on Go trials for the Go/No-Go task and congruent trials for the Flanker task.

Training Procedures

Targeted EC trainings took place over the course of 5 consecutive weekdays (Monday–Friday) from 9AM to 12PM the summer immediately prior to children's enrollment in kindergarten. All tasks consisted of child participants interacting with experimenters and other participants in small and large group settings. Chosen activities and tasks were drawn from the literature on training attentional and inhibitory domains of EC and school readiness.

Training structure. Each day of targeted training was structured similarly such that the morning began with a warm-up activity (e.g., singing a welcome song) followed by a "small group" activity (see Table E1 for sample schedule). Each small group consisted of 3 children and one main counselor who provided instructions and reflection training (Espinet et al., 2013) on the chosen activity. Reflection training is a feedback procedure in which participants are trained on correcting their error when a mistake is made, in contrast to corrective feedback where the participant is only told if their response is right or wrong, but not told why or how to correct their error. This type of feedback procedure is thought to work as one of the active ingredients in improving skills related to EC. In small groups, activities were chosen from the attentional control domain (2 tasks: DCCS and Stroop) or inhibitory control domain (4 tasks: Simon Says, Whisper Game, Hand Game, and Walk-a-Line). Small group activities took approximately 15 minutes and were followed by transitioning to a "large group" activity. Large group activities involved all 6 children and tapped into response inhibition and working memory using either oral and/or visual cues (5 tasks: Red Light/Purple Light, Freeze Game, Jumping Game, Color-Matching Freeze, and Drum Beats). Large group tasks were chosen based on literature supporting their ease of implementation in classrooms with groups of young children with varying skills related to EC (Schmitt, McClelland, Tominey, & Acock, 2014; Tominey & McClelland, 2011).

Large group activities took approximately 10 minutes followed by a different small group activity. The order in which specific small or large group activities were chosen were unique from day-to-day. Each training day had a scheduled break (i.e., snack and free play) that occurred after a total of 3 small group and 2 large group transitions, leaving 1 small group and 1 large group transition for the end of the training session. A positive group reinforcement method (i.e., marble jar) was implemented during the training program such that children engaging in on-task behavior (e.g., transitioning from small and large group activities after being prompted once), were given marbles to place in a jar. Once the jar was full, children earned a party at the end of the training program.

RESULTS

Based on intent-to-treat principles to address non-random attrition, all participants were included in analyses whenever possible or appropriate. Child participants were analyzed irrespective of training performance or attendance. Exceptions included analyses involving difference scores where both baseline and post-assessment scores were required. Moreover, children with mean ERN amplitudes greater than 3 SDs from the sample mean on any of the midline electrode sites were identified as outliers and removed from analyses (Go/No-Go task: n = 3; Flanker task: n =5). Children with mean N2 amplitudes greater than 3 SDs from the sample mean on any of the midline electrode sites were identified as outliers and removed from analyses (Go/No-Go task: n = 2; Flanker task: n =4).

Power Analyses

Mixed between-within subjects analysis of variances (ANOVAs) were conducted to evaluate time (baseline vs. post-assessment) and time by group (training vs. control) interaction effects for behavioral, neural, and parent-reported measures. When using the entire sample (N = 48), I was able to detect a significant large effect ($\eta^2_p = 0.14$) with 78% power, a significant moderate ($\eta^2_p = 0.06$) with 40% power, and a significant small ($\eta^2_p = 0.01$) with 10% power. Independent t-tests were conducted to evaluate the differences in teacher-reported concerns between training and the control group. I was able to detect a significant large effect ($d = 0.8$) with 77% power, a significant moderate effect ($d = 0.5$) with 40% power, and a significant small effect ($d = 0.2$) with 10% power. Bivariate correlations were conducted to address exploratory aims of understanding correlated changes across method. Using the entire sample, I was able to detect a significant large effect ($r = 0.5$) with 97% power, a significant moderate effect ($r = 0.3$) with 57% power, and a significant small effect ($r = 0.1$) with 10% power. Given these power analyses, results discussed below will only focus on effects that were moderate in size (e.g., $\eta^2_p = 0.06$, $d = 0.5$, $r = 0.3$) and large in size (e.g., $\eta^2_p = 0.14$, $d = 0.8$, $r =$

0.5) regardless of whether they reached statistical significance. If effects discussed did not meet statistical significance testing standards (□ $= 0.05$), it is noted in the text and indicated in the tables. Trend-level effects are defined as effects with $0.05 < p \leq 0.10$.

Randomization Checks

Performance on measures of cognitive ability and academic achievement are presented in Table A2. Independent t-test results indicated a trend-level difference between groups in matrix reasoning that was moderate in effect size such that the training group scored higher on this subtest compared to the control group ($t(41) = 1.76$, $p = 0.09$, $d = 0.54$). However, no significant differences were observed between training and the control group in the block design subtest ($t(40) = 1.03$, $p = 0.31$, $d = 0.33$), which measures similar perceptual and fluid reasoning skills as matrix reasoning. Additionally, no significant baseline differences were observed for receptive vocabulary skills ($t(38) = 1.02$, $p = 0.32$, $d = 0.32$), working memory (Nonverbal WM: $t(36) = 0.60$, $p = 0.55$, $d = 0.20$; Verbal WM: $t(36) = 0.42$, $p = 0.68$, $d = 0.14$), or academic achievement (Letter Word: $t(36) = 0.76$, $p = 0.45$, $d = 0.24$; Applied Problem: $t(36) = 0.69$, $p = 0.49$, $d = 0.23$).

In terms of baseline differences in parent-reported measures, there were no statistically significant differences in parent-reported social-emotional problems ($0.38 < p$s < 0.81; $0.07 < d$s < 0.27). EC skills reported on the BRIEF were not statistically different between groups ($0.08 < p$s < 0.94; $0.02 < d$s < 0.58). However, the training group had more parent-reported problems related to inhibitory control compared to controls, and this trend-level effect was moderate in size ($t(36) = 1.78$, $p = 0.08$, $d = 0.58$). No baseline differences in behavioral task performance were observed between groups ($0.13 < p$s < 0.99; $0.02 < d$s < 0.49). There were no baseline differences in ERN as measured by the Flanker ($t(40) = 0.90$, $p = 0.38$, $d = 0.29$) and Go/No-Go ($t(42) = 0.68$, $p = 0.50$, $d = 0.20$) tasks. The control group had a significantly larger high conflict N2 ($t(40) = 3.05$, $p = .004$, $d = 0.98$) compared to the training group on the Flanker task, but did not differ in ΔN2 ($t(40) = 0.56$,

$p = 0.58$, $d = 0.17$). There were no baseline differences in N2 as measured by the Go/No-Go task ($t(43) = 1.59$, $p = 0.12$, $d = 0.47$) or ΔN2 ($t(41) = 0.83$, $p = 0.41$, $d = 0.30$).

Behavioral Measures of EC

Ceiling effects were observed for the following behavioral tasks: Hand Game, Bear Dragon, Whisper, Knock-Tap, Reverse Categorization, and Yes/No. Therefore, the distribution of these variables was negatively skewed. Reflected logarithmic transformations were conducted for these variables, but the transformations were not successful in normalizing the distributions. Thus, non-transformed values were used for all analyses. Mean-level changes across near, intermediate, and far-intermediate transfer tasks are presented in Table A3. Graphical depictions of mean-level changes in behavioral tasks assessing attentional control skills and inhibitory control skills are presented in Figures 2 and 3, respectively.

As mentioned previously, very few studies have evaluated the developmental sensitivity of behavioral measures of EC in preschool-aged children (Carlson, 2005). Therefore, examining the impact of training on each behavioral task rather than creating a composite score was important to filling this gap in the literature. Additionally, a priori hypotheses regarding effect on transfer type (i.e., prediction that children would demonstrate the greatest improvements on near transfer tasks) were of interest in analyses.

Near transfer tasks. To test the effects of time (baseline vs. post) on near transfer task performance by group (training vs. control), mixed repeated measure ANOVAs were conducted (see Table A4). A significant main effect of time indicated improvement in Simon Says ($F(1, 40) = 5.42$, $p = 0.03$, $\eta^2_p = 0.12$). Nonsignificant, moderate effects of time were observed for Walk-a-Line ($F(1, 40) = 2.42$, $p = 0.13$, $\eta^2_p = 0.06$) and Backward Digit Span tasks ($F(1, 40) = 2.18$, $p = 0.15$, $\eta^2_p = 0.06$) such that performance increased with time. Based on descriptive statistics (see Table A3 and Figure B2), these changes appeared to be driven by the training group since improvements were moderate in effect size as compared to changes in the control group, which exhibited either no change or change that was small in effect. However, contrary to hypotheses, no statistically significant time by group interaction effects were observed.

Intermediate transfer tasks. Mixed repeated measure ANOVA results indicated a main effect of time on the Bear-Dragon ($F(1, 35) = 5.18$, $p = 0.03$, $\eta^2_p = 0.13$), DCCS ($F(1, 39) = 8.27$, $p = .007$, $\eta^2_p = 0.18$), and Day/Night tasks ($F(1, 35) = 7.05$, $p = 0.01$, $\eta^2_p = 0.17$). Specifically, children exhibited a significant reduction the Bear-Dragon task at post-assessment, but this did not significantly differ by group ($F(1, 35) = 1.84$, $p = 0.18$, $\eta^2_p = 0.05$). It is important to note that there was low variability in performance on this task and evidence of ceiling effects. In other words, a majority of children achieved the maximum score, and only four children declined in their performance on this task. Both groups demonstrated significantly higher scores at post-assessment compared to baseline for DCCS and Day/Night tasks. The training group exhibited significantly greater improvement on the Day/Night task compared to the control group ($F(1, 35) = 5.45$, $p = 0.03$, $\eta^2_p = 0.14$), but DCCS improvement was not different between groups ($F(1, 39) = 0.21$, $p = 0.65$, $\eta^2_p = 0.01$).

Far-intermediate transfer tasks. In terms of performance on far-intermediate transfer tasks, post-assessment scores on the Head-Toes-Knees-Shoulder task ($F(1, 40) = 7.02$, $p = 0.01$, $\eta^2_p = 0.15$) and FIST Selection 2 variable ($F(1, 34) = 9.53$, $p < .001$, $\eta^2_p = 0.22$) were significantly higher compared to baseline for both groups. A nonsignificant, trend-level effect of time was also observed for the Tower of Hanoi task such that performance increased over time ($F(1, 35) = 3.61$, $p = 0.07$), $\eta^2_p = 0.22$). Mixed repeated measure ANOVA results indicated that the training group exhibited significantly greater improvement in the Head-Toes-Knees-Shoulder task ($F(1, 40) = 5.40$, $p = 0.03$,

$\eta^2_p = 0.12$) whereas the control group demonstrated a nonsignificant, trend-level improvement on the FIST Selection 2 score that was moderate in effect size.

Interim summary of results from behavioral measures of EC. A priori predictions stated that the training group would demonstrate the greatest improvements in near transfer tasks. Contrary to hypotheses, the training group did not exhibit improvements that were dependent on type of transfer task. The largest, and only significant time by group effects, were observed in the Day/Night intermediate transfer task and HTKS far-intermediate transfer task such that the training group demonstrated significantly larger improvements on these tasks.

Neurophysiological Measures of EC

Behavioral performance. Behavioral performance on the Go/No-Go and Flanker tasks are presented in Table A5. Since the Go/No-Go and Flanker tasks were not practiced during training and no skills related to improving accuracy on computer tasks were targeted during training, behavioral performance on these tasks was considered a far transfer measure. I anticipated that children in the training group would demonstrate improved behavioral performance in both the Go/No-Go and Flanker tasks compared to children in the control group.

Effects of training across time (baseline vs. post) and by group (training vs. control) on behavioral performance measures were evaluated using mixed repeated measure ANOVAs (see Table A6). For the Go/No-Go task, results indicated a significant effect of time for reaction time on correct ($F(1, 45) = 4.73$, $p = 0.04$, $\eta^2_p = 0.10$), post-error ($F(1, 45) = 8.88$, $p = .005$, $\eta^2_p = 0.17$), and post-correct trials ($F(1, 45) = 9.79$, $p = .003$, $\eta^2_p = 0.18$) such that all children exhibited significant reductions in reaction time at post-assessment compared to baseline. However, these reductions did not significantly differ between groups. In the Flanker task, results indicated a significant effect of time for number of errors ($F(1, 45) = 38.84$, $p < .001$, $\eta^2_p = 0.46$) and correct trials ($F(1, 45) = 22.90$, $p < .001$, $\eta^2_p = 0.34$), total accuracy ($F(1, 45) = 27.86$, $p < .001$, $\eta^2_p = 0.38$), and post-error accuracy

($F(1, 45) = 55.64$, $p < .001$, $\eta^2_p = 0.55$) such that all children exhibited significant increases in total accuracy (fewer error trials and more correct trials) and improvements in post-error accuracy. Similar to the Go/No-Go task, there were no significant time by group interaction effects.

These results suggest that children demonstrated greater improvements on the Flanker task compared to the Go/No-Go task. It is important to acknowledge these differences given that these tasks may have challenged different skills related to EC. Additional analyses suggested that overall, children exhibited higher total accuracy on the Go/No-Go compared to the Flanker task (baseline: $t(46) = 11.09$, $p < .001$; post-assessment: $t(46) = 7.03$, $p < .001$). These results suggested that the Flanker task was more challenging for children in this developmental age group. Therefore, it is likely that behavioral performance indices from the Flanker task captured greater variability in individual differences in these measures compared to the Go/No-Go task and may explain the greater improvements observed in the Flanker task. In other words, measures that have greater variability also allow for more opportunities for change.

ERN. The location of maximal amplitude for the ERN was evaluated using a 2 (trial type: error; correct) X 5 (site: Fz; FCz; Cz; CPz; Pz) repeated measures ANOVA model and follow-up paired sample t-tests. A detailed review of results supporting the use of ERN analyses at FCz is provided in Appendix F. The response-locked waveforms from the Go/No-Go and Flanker tasks can be seen in Figure B4.

Descriptive statistics (see Table A7) suggested a pattern of larger Go/No-Go and Flanker ERN over time, but contrary to stated hypotheses, no significant effects of time or time by group interactions were observed (see Table A8).

N2. The location of maximal amplitude for the N2 was evaluated using a 2 (trial type: error; correct) X 5 (site: Fz; FCz; Cz; CPz; Pz) repeated measures ANOVA model and follow-up paired sample t-tests. A detailed review of results supporting the use of N2 analyses at FCz is provided in

Appendix F. The stimulus-locked waveforms from the Go/No-Go and Flanker tasks can be seen in Figure B5.

Contrary to a priori predictions, mixed repeated measure ANOVA results indicated that the Go/No-Go high conflict N2 ($F(1, 42) = 5.62, p = 0.02, \eta^2_p = 0.12$) and ΔN2 ($F(1, 42) = 14.45, p < .001, \eta^2_p = 0.26$) was more negative (larger) at post-assessment (see Table A8). However, this increase in N2 amplitude was not significantly different between groups (N2: $F(1, 42) = 0.32, p = 0.57, \eta^2_p = 0.01$; ΔN2: $F(1, 42) = 1.42, p = 0.24, \eta^2_p = 0.03$). In contrast, the high conflict N2 assessed by the Flanker task was more positive at post-assessment ($F(1, 37) = 4.16, p = 0.05, \eta^2_p = 0.10$). A significant time by group interaction suggested that this reduction in the N2 amplitude was specific to children in the control group ($F(1, 37) = 5.69, p = 0.02, \eta^2_p = 0.13$) whereas the children in the training group did not exhibit any significant changes in Flanker N2 amplitude.

Interim summary of results from neurophysiological measures of EC. A priori predictions stated that children in the training group would exhibit larger increases in the ERN amplitude, and greater reductions in the N2 amplitude compared to controls. Additionally, since current literature is unequivocal about the relationship between the ERN and N2, particularly in young children, it was predicted that if the ERN and N2 did exhibit changes in the same direction, this result would suggest that they index similar processes. No significant time by group interactions were observed for ERN measures. A significant time by group interaction was observed for the Flanker N2 such that the control group demonstrated a greater reduction in the high conflict N2 compared to the training group. Given the baseline difference in the high conflict N2 (control group exhibited a larger Flanker N2 at baseline), it is likely that this result reflects regression to the mean. Both groups demonstrated significant increases in the Go/No-Go N2 measures over time. Taken together, these results are inconsistent with predicted hypotheses.

Informant-reported Measures of EC

Parent report of EC skills. A priori predictions stated that the training group would exhibit reductions in parent-reported problems both those related to EC skills and social-emotional functioning. Graphical depictions of mean-level changes in parent-reported EC skills are presented in Figure B6. Mixed repeated measure ANOVA results suggested a main effect of time from baseline to 1-month follow-up such that significant reductions in parent-reported problems related to EC skills were observed (see Table A10). A significant time by group interaction was observed for inhibitory control skills ($F(1, 34) = 7.64$, $p = .009$, $\eta^2_p = 0.18$) such that the training group exhibited a greater reduction in parent-reported problems with inhibition compared to the control group. Nonsignificant, trend-level time by group interactions that were moderate in effect were observed for emotional control skills ($F(1, 34) = 2.89$, $p = 0.10$, $\eta^2_p = 0.08$) and total score on the BRIEF ($F(1, 34) = 3.73$, $p = 0.06$, $\eta^2_p = 0.10$). These results suggested that the training group demonstrated increases in general EC skills and, specifically, the ability to engage in inhibitory control and emotion regulation skills 1 month following the training compared to the control group.

Mixed repeated measure ANOVAs were conducted to test whether these effects were maintained 3 months following the training. Results indicated no statistically significant main effects of time across all scales such that parent-reported EC skills were comparable at 1 month and 3 months after training. A nonsignificant, trend-level increase in problems with working memory was observed that varied by group such that the training group exhibited an increase in working memory difficulties ($F(1, 31) = 2.90$, $p = 0.10$, $\eta^2_p = 0.09$). Similarly, a nonsignificant moderate increase in problems with planning and organization skills was observed in the training group compared to control ($F(1, 31) = 1.97$, $p = 0.17$, $\eta^2_p = 0.06$).

Parent report of social-emotional functioning. Graphical depictions of mean-level changes in parent-reported social-emotional adjustment problems are presented in Figure B7. Mixed repeated measure ANOVA results indicated significant reductions in all scales related to social-emotional problem behaviors from baseline to 1-month follow-up (see Table A10). Nonsignificant, trend-level time by group interactions were observed for anxious/depressed problems ($F(1, 41) = 3.53$, $p = 0.07$, $\eta^2_p = 0.08$), social problems ($F(1, 41) = 3.66$, $p = 0.06$, $\eta^2_p = 0.08$), and attention problems ($F(1, 41) = 3.77$, $p = 0.06$, $\eta^2_p = 0.08$), such that the training group exhibited greater reductions in these areas compared to the control group. Similarly, nonsignificant, but moderate effects of time by group interactions were observed for internalizing problems ($F(1, 41) = 2.74$, $p = 0.11$, $\eta^2_p = 0.06$) and externalizing problems ($F(1, 41) = 2.38$, $p = 0.13$, $\eta^2_p = 0.06$) such that the training group exhibited greater reductions in these areas compared to the control group.

Mixed repeated measure ANOVAs were conducted to test whether effects were maintained 3 months following the training. Results suggested significant reductions in withdrawn/depressed behaviors ($F(1, 31) = 4.54$, $p = 0.04$, $\eta^2_p = 0.13$), social problems ($F(1, 31) = 8.56$, $p = 0.01$, $\eta^2_p = 0.22$), and internalizing behaviors ($F(1, 31) = 4.14$, $p = 0.05$, $\eta^2_p = 0.12$) from 1-month to 3-months post-assessment. A nonsignificant, trend-level reduction was also observed for parent-reported anxious/depressed problems ($F(1, 31) = 3.05$, $p = 0.09$, $\eta^2_p = 0.09$). These reductions did not significantly differ by group.

Teacher report of EC skills. A priori predictions stated that children in the training group would have lower meal-levels of teacher reported problems. Independent t-test analyses indicated that teacher-reported BRIEF scales at 3 months follow-up were not statistically significantly different between the training and control group ($0.31 < p$s < 0.87; $0.07 < d$s < 0.39) (Table A11).

Teacher report of social-emotional functioning. Independent t-test analyses indicated that teacher-reported CBCL scales at 3 months follow-up were not statistically significantly different between training and the control group ($0.15 < p$s < 0.99; $0.00 < d$s < 0.66) (Table A11). Nonsignificant, but moderate effects were observed for externalizing problems ($t(25) = 1.24$, $p = 0.23$, $d = -0.51$) such that the control had fewer teacher-reported problems with externalizing behaviors. In contrast, the training group had nonsignificant, moderate effects of lower teacher-reported difficulties with withdrawn/depressed ($t(25) = 1.16$, $p = 0.26$, $d = 0.51$) and somatic problem behaviors ($t(25) = 1.49$, $p = 0.15$, $d = 0.55$) compared to the control group.

Interim summary of results from informant-reported measures of EC. I hypothesized that children participating in targeted training would have greater reductions in parent-reported problems, and lower mean-levels of teacher reported problems at follow-up compared to control. While not statistically significant, moderate effects of time by group interactions were detected for greater reductions in parent-reported problems with inhibitory and emotional control in the training group compared to the control group. Similarly, nonsignificant, moderate effects of time by group interactions were observed for internalizing and externalizing behaviors such that the training group demonstrated greater reductions in these problems compared to the control group. These results were largely consistent with initial predictions. However, in contrast to hypotheses, teacher-reported concerns at 3 months follow-up were not statistically different between groups. Nonsignificant, moderate effects were observed for fewer withdrawn/depressed and somatic problems in the training group, and higher externalizing problems in the training group compared to the control group.

Associations Between Behavioral, Neural, and Parent-reported Measures of EC

In order to better understand the extent to which the ERN and N2 are measures of EC, an exploratory aim was to examine how changes in the ERN and N2 were associated with observed changes in behavioral and informant-reported measures of EC. Given the sample size, I was underpowered to evaluate whether these associations were moderated by group. Therefore, exploratory analyses focused on understanding the individual differences in change that occurred in neural measures of EC across time and whether these changes cohered with behavioral and informant-reported measures of EC. I anticipated that greater improvements in behavioral measures of EC would be associated with greater changes in ERP measures (increases in ERN and decreases in the N2 if they index distinct facets of EC, or in similar directions if they index similar EC skills). Similarly, greater reduction in parent-reported adjustment and EC problems would be associated with greater changes in ERP measures. Lastly, greater improvements in behavioral measures of EC would be associated with greater reductions in parent-reported adjustment and EC problems. While I did not have hypotheses about specific measures within methodology, I anticipated that correlated change across method would be stronger between methods that were closer in level of measurement. In other words, the strength of correlated change would be observed in the following order from weakest to strongest: parent-reported and neural measures of EC; behavioral and neural measures of EC; and parent-reported and behavioral measures of EC.

Neurophysiological and behavioral measures of EC.

Behavioral tasks tapping attentional control skills. Bivariate correlations were conducted between the change scores (post-assessment minus baseline) of neurophysiological measures of EC and behavioral measures of attentional control (Table A12). The Go/No-Go ERN was significantly associated with changes in total FIST score ($r = -0.35$, $p = .05$) such that greater increases in the Go/No-Go ERN were associated with improvements in the FIST task (Figure H1). Surprisingly, improvement in the DCCS Border measure was associated with reductions in the Flanker ERN ($r = 0.35$, $p = .05$) (Figure H2), a similar nonsignificant trend was observed with improvement in the Yes/No task ($r = 0.30$, $p = 0.08$). In terms of the N2, there were no significant associations between the Go/No-Go N2 measures and behavioral performance on attentional control tasks. Reductions in the Flanker ΔN2 were associated with improvement in the DCCS task ($r = 0.35$, $p = 0.04$) (Figure H3), and a similar nonsignificant trend was observed with improvement in the Tower of Hanoi task ($r = 0.32$, $p = 0.08$). Overall, changes in neurophysiological measures of EC were not associated with any near transfer attentional control tasks. Furthermore, changes in the Go/No-Go and Flanker ERN had contrasting associations with improvements in attentional control task performance.

Behavioral tasks tapping inhibitory control skills. No significant associations between changes in neurophysiological measures of EC and behavioral measures of inhibitory control skills were observed (Table A13). Nonsignificant trend-level associations were detected between reductions in the Go/No-Go ΔERN and Flanker ΔN2 and improvement in the HTKS trial measure ($r = 0.28$, $p = 0.09$ $r = 0.29$, $p = 0.09$, respectively) (Figures H4-H5).

Neurophysiological measures of EC and parent-reported measures.

Parent report of EC skills. Bivariate correlations were conducted between the change scores of neurophysiological measures of EC and parent report of EC skills (1-month follow-up minus baseline scores) assessed by the BRIEF (Table A14). No significant or trend-level associations between changes in neurophysiological measures and parent report of EC were observed.

Parent report of social-emotional behaviors. Bivariate correlations were conducted between the change scores of neurophysiological measures of EC and parent report of social-emotional functioning (1-month follow-up minus baseline scores) assessed by the CBCL (Table A15). Increases in the high conflict N2 assessed by the Go/No-Go task were significantly associated with decreases parent-reported social problems ($r = 0.32$, $p = 0.05$) (Figure H6). Nonsignificant, trend-level associations were observed between increases in the Go/No-Go ERN and fewer parent-reported social problems ($r = 0.27$, p =0.10) and rule-breaking behaviors ($r = 0.27$, p =0.10) (Figures

H7-H8). In contrast, nonsignificant, trend-level decreases in the Flanker ΔN2 were associated with fewer parent-reported problems with rule breaking (r = -0.29, p =0.09) (Figure H9).

Behavioral and parent-reported measures of EC and social-emotional behaviors.

Behavioral tasks tapping attentional control skills and parent-reported measures. With a few exceptions, a majority of changes in behavioral tasks assessing attentional control skills were unrelated to changes in parent-reported BRIEF and CBCL scores (Table A18). Improvement in the Day/Night task was significantly associated with reduction in parent-reported CBCL problems related to rule-breaking (r = -0.41, p = 0.01) and aggression (r = -0.41, p = 0.03). In contrast, improvement in the FIST task was associated with more reported problems related to working memory (r = 0.48, p = .003) and attentional shifting (r = 0.45, p = 0.01).

Behavioral tasks tapping inhibitory control skills and parent-reported measures. A majority of changes in behavioral tasks assessing inhibitory control skills were unrelated to changes in parent-reported BRIEF and CBCL scores (Table A19). Improvement on the Walk-A-Line task was significantly associated with reductions in parent-reported problems with inhibition, overall EC skills, social concerns, attention, and aggression (0.32 < rs < 0.40; p < .05). Surprisingly, improvement in Simon Says was significantly associated with greater parent-reported problems with aggression (r = 0.43, p = .006). In contrast, improvement in the Hand Game was significantly associated with reductions in problems with attentional shifting (r = -0.44, p = .007) and overall EC skills (r = -0.36, p = 0.03). All other changes in behavioral tasks assessing inhibitory control skills were unrelated to changes in parent-reported EC.

Interim summary of associations across methodologies. As an exploratory aim, I did not have hypotheses regarding associations between specific measures. However, broader a priori hypotheses stated that greater changes in neural indicators of EC would be associated with greater changes in behavioral measures of EC and parent-reported adjustment.

Contrary to hypotheses, the pattern of correlated change associations between neural and behavioral measures of EC (attentional and inhibitory control tasks) were difficult to discern given their inconsistency. Overall, the results suggested that changes in neural measures of EC were unrelated to changes in attentional and inhibitory control behavioral tasks.

Similar to the lack of correlated changes between neural and behavioral measures of EC, no consistent pattern of associations was observed between changes in neural and parent-reported measures. These results suggested that changes in parent-reported concerns during this time are likely independent from changes in the ERN and N2.

With a few exceptions, a majority of changes in behavioral measures of EC were unrelated to changes in parent-reported measures of EC and social-emotional adjustment. Taken together, these results indicate that there are substantial individual differences in the type of change that occurred over this time period. While there were significant changes (primarily in behavioral measures of EC and parent-reported concerns) reported during this time period, these changes appeared to occur exclusive of one another.

DISCUSSION

The purpose of the present study was to better understand the neurobehavioral processes that may underlie training-induced changes in EC. This investigation was the first to use a short-term training to explicitly test the functional validity of the ERN and N2 as markers of EC, and to evaluate training-induced changes in EC using a multimethod approach. The primary focus of the investigation was not necessarily to develop or assess the effectiveness of an EC intervention, but rather to use an innovative method of evaluating the validity of the ERN and N2 as markers of EC. The method of testing the validity of the ERN and N2 was dependent on assessing the training effects, which was similar to the methods used to assess the effectiveness of an intervention.

Overall, effects of training-induced changes in EC were small. This contrasts results from previous training studies using similar measures (Rueda, Checa, & Cómbita, 2012; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009) that reported changes in behavioral and neural measures that were larger in effect (d = 0.99 and d = 1.36, respectively). The present results are consistent with more recent reviews demonstrating that computer-based cognitive training targeting working memory and other EC-related skills do not produce significant effects of improved cognitive abilities, attention, EC-related skills, behavior, or academic functioning in children (Cortese et al., 2015; Melby-Lervåg & Hulme, 2013; Rapport, Orban, Kofler, Friedman, 2013) and young adults (Redick et al., 2013). A recent meta-analysis also indicated that age was a significant moderator of the effects of working memory training on near transfer tasks such that smaller effects were observed in children compared to young and elderly adults (Melby-Lervåg, Redick, & Hulme, 2016). A summary of the results from each methodology will be presented below, in addition to a synthesis and discussion of these results.

Main Effects Within Method Type

Behavioral Measures of EC. Less than half of the selected behavioral tasks demonstrated significant changes from baseline to post-assessment. Contrary to hypotheses, children did not demonstrate the greatest improvements on near transfer compared to intermediate or far-intermediate transfer tasks. Significant improvements in behavioral task performance in the training group compared to the control group were observed for an intermediate transfer attentional control task (Day/Night) and a far-intermediate transfer inhibitory control task (HTKS). Changes in performance on behavioral tasks did not favor tasks assessing more attentional control skills or inhibitory control skills. This result may reflect the notion that training targeted both attentional and inhibitory control skills related to EC. Alternatively, as suggested by previous research, these facets of EC may be best represented as a unitary factor in preschool (Wiebe, Epsy, & Charak, 2008; Willoughby, Blair, Wirth, & Greenberg, 2012). The demonstrated intermediate and far-intermediate transfer is also inconsistent with transfer theory predictions (Klauer, 2001), which limit the extent of potential transfer to the relationship with near transfer improvement (i.e., using the near transfer effect size as an estimate). It should be noted that the lack of near transfer improvement makes it difficult to interpret whether any other observed changes reflect random or systematic effects, rather than the training itself.

Neurophysiological Measures of EC. Children demonstrated improvement on several behavioral performance indices assessed using the Go/No-Go and Flanker tasks. Children exhibited greater improvements in the Flanker task compared to the Go/No-Go task. It is possible that children reached a ceiling effect in their behavioral performance on the Go/No-Go task, whereas the Flanker task was more challenging and therefore there was more to improve on with time. In addition, children tended to become faster with time. Children exhibited faster reaction times on error and correct trials on the Flanker task whereas they demonstrated faster reaction times on post-error and post-correct trials on the Go/No-Go task. Given that these changes in behavioral performance across time was not statistically significant between groups, these changes are likely attributable to maturation and/or retest effects.

Changes in neurophysiological measures of EC were observed in N2 measures. Specifically, children demonstrated reductions in the Flanker high conflict N2 over time. This effect was driven by the significant reduction in the high conflict N2 in the control group. Children also demonstrated an increase in the Go/No-Go high conflict N2, but this change did not significantly differ by group. Increases were observed in the Flanker ERN, but these changes were not specific to group condition. These results are inconsistent with a priori predictions, which stated that the training group would exhibit a larger ERN and reduced N2. A priori hypotheses suggested that the ERN and N2 may index different facets of EC (attentional control and inhibitory control, respectively). However, given the present results there are several possible explanations. First, the ERN and N2 may instead index similar EC-related skills. Evidence that the Go/No-Go ERN and Go/No-Go N2 were related at both baseline and post-assessments (see Tables H1-H2) support the notion that ERN and N2 may measure similar processes. However, the Flanker ERN and Flanker N2 were not related at baseline or post-assessments (see Tables I3-I4), indicating that these neural indices were not reliably associated with one another within this task. Second, improvements in behavioral measures of EC over time may not be supported by changes in the ERN and N2. Third, more advanced data reduction and analytic techniques may be required to address the high variability in ERP amplitude observed in this sample. There are a number of factors that contributed to the large variance of these measures, such as the variability in ERP amplitude latency often reported in young children. More advanced analytic techniques that could account for latency differences would be helpful in future investigations. Fourth, the ERN and N2 may not be sensitive enough to measure EC-related skills, specifically attentional and inhibitory control, in this developmental period.

While the present findings are inconsistent with a priori predictions, the results indicate that ERP measures, particularly the N2, are quite dynamic during this developmental time period. This contrasts research in adults that suggests that these ERP measures are stable over time (Weinberg & Hajcak, 2011). In addition, results support the notion that the Go/No-Go and Flanker tasks elicit differences in behavioral performance that may manifest at the neurophysiological level in young children, and highlight the importance of using multiple tasks to elicit the same ERPs of interest. The multitude of differences between the Go/No-Go and Flanker tasks such as ERP amplitude, ERP change, and behavioral performance further demonstrate the value of using multiple cognitive tasks to better understand the functioning meaning of the ERN and N2 in young children. Results from the present study support the notion that the Go/No-Go and Flanker tasks likely require children to engage in different types of EC skills that manifest at behavioral and neural levels. This contrasts research in adults suggesting that the same ERP assessed across different types of cognitive tasks indexes the same skill or neural process across tasks (Riesel, Weinberg, Endrass, Meyer, & Hajcak, 2013). Lastly, the present findings provide insight into how little we understand about measures of neural processes such as the ERN and N2 in young children. The ERN and N2 may not be the optimal measures to use as primary outcome variables for future training studies using children in this age range. Until we understand more about how to better analyze and elicit ERPs in young children, funding agencies and other stakeholders should be cautious interpreting intervention effects based on neural measures of change in early childhood.

Informant-reported Measures. As predicted, the training group demonstrated significant reductions in parent-reported inhibition problems compared to the control group. Nonsignificant, trend-level reductions were also observed for parent-reported emotional control problems. Follow-up exploratory analyses examined the extent to which training effects on parent-reported EC problems were maintained at 3-months follow-up. Reductions were maintained in the training group, but the control group also demonstrated reductions such that group differences in parent-reported BRIEF scores were not observed. Given that the training group had a nonsignificant, but moderate effect of higher parent-reported inhibition problems at baseline, it is possible that the present results reflect regression to the mean.

Results also indicated significant reductions in parent-reported social-emotional concerns 1-month following training. These reductions were not statistically significant between groups, but the training group demonstrated trend-level, moderate reductions in parent-reported anxious/depressed problems, social problems, and attention problems compared to the control group. Follow-up exploratory analyses examined the extent to which training effects on parent-reported social-emotional concerns were maintained at 3-months follow-up. While training effects were maintained, the control group demonstrated reductions in parent-reported social-emotional concerns. Thus, no significant differences between the training and control group were observed. These results also support evidence of regression to the mean. Additionally, because parents were not blinded to group assignment, it is possible that the initial reduction in parent-reported social emotional problems in the training group reflected placebo or demand characteristics effects. In other words, parents of children in the training group may have believed that participating in a week-long training targeting EC skills would likely benefit their child, and this belief may have biased their reports such that the reduction in problem behaviors was overestimated.

Contrary to a priori predictions, there were no statistically significant group differences in teacher-reported EC skills and social-emotional functioning 3 months following training. Nonsignificant, moderate effects of higher reported externalizing behaviors in the training group were observed. In contrast, nonsignificant, moderate effects of lower reported withdrawn/depressed and somatic complaint problems in the training group were also observed. There are several explanations for these inconsistent results. First, exploratory analyses revealed that while effects of reduced parent-reported concerns were maintained 3 months after training, the training and the control group no longer differed in parent-reported problems with EC skills of social-emotional functioning. Teacher-reported measures were collected 3 months after training in order to permit enough time for teachers to become familiar with their students. Therefore, it is possible that the lack of difference between training and the control group in terms of teacher-reported problems reflects the similar lack of difference in parent-reported concerns between groups at the 3-month follow-up assessment. Second, the teacher-reported results may provide further support that the greater reduction in parent-reported problems for the training group observed 1 month following training is attributable to demand characteristics. When recruited, parents of children in the training group were told that their children were participating in a "camp" that targeted self-regulation and EC skills important for kindergarten. Therefore, it is possible that parents anticipated that their children should improve in these dimensions following the training.

Associations Between Changes in Multimethod Measures Across Time

Behavioral and Neurophysiological Measures of EC. A majority of correlated changes in ERP and behavioral task measures were unrelated. Of the five associations that were significant, the patterns across behavioral tasks assessing either attentional or inhibitory control skills were not consistent with hypotheses. Specifically, *decreases* rather than predicted increases in ERP measures were associated with improvements in behavioral task scores, with one exception. Increases in the Go/No-Go ERN were associated with improvements in the FIST task. Of the significant correlated change associations, changes in the ERN and N2 occurred in similar directions, lending some support to the notion that they may index similar processes during this developmental period. No patterns emerged that were specific to behavioral tasks that assessed attentional versus inhibitory control skills.

There are a number of possible explanations and implications of these results. First, the limited variability of the behavioral measures of EC due to ceiling effects likely reduced the ability to detect effects. Second, it is possible that changes in behavioral measures of EC are not dependent on changes in the ERN and N2, and vice versa. Changes in behavioral and neural measures of EC that occurred from baseline to post-assessment may occur at different rates. The maturation of these processes may manifest differently not only by time, but also by method, and it would be important for future research to understand whether change assessed by one type of method is dependent on changes in another (e.g., neural changes supporting behavioral changes, vice versa, or neither).

Parent-Reported Measures of Adjustment and Neurophysiological Measures of EC. The unexpected results of correlated change between ERP measures and behavioral task measures discussed above is further complicated by the correlated changes observed between ERP and parent-reported measures. First, a majority of correlated changes in ERP and parent-reported measures (BRIEF and CBCL) were unrelated. Of the four associations that were significant, some were consistent with hypotheses while others were not. Specifically, increases in the Go/No-Go ERN were associated with fewer parent-reported social problems and rule-breaking behaviors. Similar to the Go/No-Go ERN, increases in the Go/No-Go N2 were associated with fewer parent-reported social problems. However, reductions in the Go/No-Go N2 were associated with rule-breaking behaviors. No correlated change associations were observed between ERP measures and parent-reported BRIEF scores. While I was underpowered to conduct further moderator analyses to examine whether correlated change associations differed by condition, scatterplots (Figures H1-H9) indicated that associations were similar between conditions.

Behavioral Measures of EC and Parent-Reported Measures of Adjustment. Children who made greater gains in behavioral task performance were anticipated to also exhibit the greatest reductions in parent-reported concerns. Similar to the other reported correlated changes across method type, changes in behavioral task scores were unrelated to changes in parent-reported concerns. Of the five behavioral tasks that exhibited significant change correlations scores, most were in the expected direction. Specifically, increases in behavioral task score were associated with decreases in parent-reported concerns. Exceptions included the FIST and Simon Says. Improvement in the Walk-a-Line task was reliably associated with lower parent-reported problems across several BRIEF and CBCL scales. It is possible that the consistency of associations observed within this task reflects the high variability of the way performance is measured. Specifically, Walk-a-Line uses length time (seconds) as a measure of inhibitory control. Therefore, this task allows for more variability in terms of score compared to a task that has a restricted range (such as a total possible score from 0-8 like the Yes/No task).

Limitations

An ongoing issue in the developmental literature, particularly for early childhood research, is how to resolve the conflict of using developmentally appropriate measures to elicit target skills of interest and simultaneously ensuring the validity of such measures to assess these skills. Furthermore, a fundamental measurement problem in EC research is the conflation between specific processes of interest and foundational skills that are central to those specific processes that cannot be disentangled given the demands of a chosen EC task (Epsy, Clark, Garza, Nelson, James, & Choi, 2016). In the present study, this problem made it challenging to determine the meaning of null and non-confirmatory results. For example, there are several interpretations regarding the lack of dissociation between the ERN and N2. First, attentional control and inhibitory control skills may not be discrete constructs during preschool-age. Second, attentional and inhibitory skills may be distinguishable, but researchers have yet to devise measures that are developmentally sensitive enough to detect the differences. And third, the ERN and N2 may measures similar processes related to EC. Another significant limitation is that the present study failed to find evidence of near transfer improvement, which renders it unfeasible to determine whether any other observed changes reflected random or systematic effects, rather than training effects. It is also possible that initial a priori predictions of what constitutes near versus far transfer change were incorrect, and the processes that changed as a result of the training were different than anticipated.

Another complication of the current study is that children exhibited substantial variability in ERP amplitude such that standard deviations were significantly larger average amplitudes, which meant that detecting any significant effects would require a much larger sample. Important to the study was understanding why children might have such significant variability in these ERP measures. Therefore, it was critical to keep as many children as possible in analyses while still attempting to adhere to typical parameters adopted in the current literature (i.e., behavioral accuracy cutoffs or specific number of useable trials). There are a number of methodological concerns in using ERP measures in this age range, with some researchers suggesting that they may be unreliable, and thus an invalid assessment of EC-related skills. However, examining the unreliability of the measure may be critical to understanding how these processes unfold. Unreliability can come in a number of forms such as a lack of consistency in latency and peak of an ERP from trial to trial, or lack of reliability from one time-point to another. Both forms may be quite meaningful, particularly in tracking the development of the ERN and N2. This issue impacted the results of the current study because average ERP amplitudes likely did not capture the exact window of where the ERP occurred across all child participants. In other words, one child may have had a much earlier latency of onset of an ERP compared to another child that had a much later latency. As a result, the chosen time window during which the ERP was averaged may not accurately reflect the process of engaging in EC skills for that child. Ideally, one could use measurement techniques to select time windows specific to each child's peak ERP amplitude. Unfortunately, the analytic techniques required to analyze the ERP data in this manner were out of the scope of the present investigation. However, one would anticipate that variability across trial and lack of reliability decreases with increasing age.

Despite attempts at randomization, there were several notable baseline differences between training and the control group that may have impacted the present results. Specifically, the control group had a significantly larger Flanker N2 at baseline, and the effect of time by group observed for the Flanker N2 was driven by the reduction in the N2 in the control group. It is likely that the significant time by group interaction reflected regression to the mean for children in the control group. Another limitation that impacted the study was the lack of blinded group assignment. Thus, problems with demand characteristics were of concern. While significant reductions in parent-reported concerns at 1-month follow-up appeared specific to the training group, there were two alternative explanations to suggest demand characteristics played a role. First, treatment gains in the training group were not significantly different from that of the control group at 3-months follow-up, perhaps suggesting that parents initially overestimated the reduction in problem behaviors given assumptions that participation in targeted training would increase EC skills. Second, teacher-reported results indicated no significant differences between the training and control group. Teachers were blind to conditions, so their responses were likely not impacted by demand characteristics. Unfortunately, this limitation was specific to the study design and could not be corrected post-hoc. Similarly, problems with ceiling effects were uncorrectable using statistical analyses. On several behavioral tasks, children had already mastered the task at baseline, reducing the likelihood that these tasks could capture variability of individual differences and that children would change in their performance from baseline to post-assessment. However, given that there is limited research on the level of task difficulty and sensitivity of multiple behavioral tasks to assess EC skills in this age range, the ceiling effects observed in the present study are useful for future research. Specifically, if researchers are interested in evaluating the effects of a targeted EC training, they may wish to select tasks that demonstrated greater variability and change within the current sample (e.g., FIST, HTKS, Simon Says) to maximize the likelihood of capturing individual differences in EC skills at this age.

Lastly, the main limitation of using a passive control group rather than an active control group is that the study design does not control for other factors related to targeted training such as engagement with training staff and additional visits to the laboratory space. Attempts were made to minimize the impact of this limitation, such as ensuring that training staff did not act as experimenters in baseline and post-assessments. Additionally, using a passive control group is a critical first step for evaluating the effect of interventions prior to using an active control group. Specifically, a passive control group controls for any changes due to maturation and practice; as both intervention and active control groups are subject to maturation and practice effects, these must be ruled out prior to comparing an intervention to an active control group. Results from the present study failed to establish a proof-of-concept for predicted training effects on the ERN and N2, and therefore, suggest that a more rigorous active control design would produce no effect of training on these neurophysiological indices of interest.

Future Directions

Despite the increased prominence of EC-related constructs in the literature, particularly in research focusing on the early development of EC-related skills and their implication for later life outcomes, there are few studies that provide any data or recommendations about measurements to assess these skills in preschool-age children. One of the first comprehensive investigations in children aged 2 to 6 was conducted by Carlson (2005). Carlson demonstrated the importance of systematically examining the task difficulty of commonly used measures of EC-related skills across different age groups in this period. The battery of behavioral tasks in the present study were selected to encompass varying levels of task difficulty. The reasons for this were threefold: 1) it was unknown how children selected for poor-to-average EC skills would perform; 2) it was unknown how performance would change in response to targeted EC training; and 3) some newer behavioral tasks had not been investigated alongside other commonly used tasks to assess EC in order to compare differences in task performance in this age group. Findings from the present study suggest that more research is needed to identify more sensitive behavioral measures of EC skills in young children. This research could help better guide researchers in selecting tasks for detecting the development of early EC skills and predicting later adjustment outcomes.

Similarly, more research is needed to understand the development of associated neural changes that support behavioral changes in EC skills. Specific to ERPs, future studies should use multiple developmental cognitive tasks within the same sample to elicit ERPs of interest. Findings from the present study suggest that the ERN and N2 elicited across the Go/No-Go and Flanker tasks may capture different aspects of EC in young children. It is likely that the task demands of each of these tasks differ, with the Go/No-Go task engaging more inhibitory control processes and the Flanker task engaging more attentional control processes. Examining how these task distinctions might contribute to differences reflected in the ERP amplitudes and associations to behavioral correlates are important to understanding the functional significance of these ERPs in young children. Additionally, it is recommended that future developmental ERP studies focus on devising a standard set of procedures and practical guidelines for data acquisition and analysis in children. For example, variability in temporal consistency and latency jitter likely had an impact on the variability observed in the ERN and N2 amplitudes in the present results. Woody filtering techniques (Woody, 1967) time shifts an ERP waveform on each segment to match the averaged ERP template, and then averages the re-aligned segments (see Lin et al., 2015 for example). Time frequency analyses can then be used to examine the extent to which signals in the theta band (frequency band of the ERN and N2) are synchronized, and the impact of latency jitter on signal intensity and amplitude. Within-subject analyses can also be used to better understand trial-to-trial variability in behavioral performance and the impact changes in an individual's behavioral performance may have on their own ERN and N2 components. It would be useful for these types of analytic techniques to be outlined in the literature as they relate to ERP assessments in young children in order to understand the high variability in ERP amplitude and latency often reported in early childhood samples.

Lastly, more effort should be made to publish negative and/or non-confirmatory effects, particularly in intervention and training research. Evidence suggests that the current scientific literature is complicated by the tendency to publish positive results (Franco, Malhortra, & Simonovits, 2016), thereby overestimating the true effects of certain phenomena. If researchers and policy makers want to develop effective methods of increasing EC skills in children that are supported by empirical evidence, then it is vital that we begin to systematically understand the true effects of all trainings targeting these skills so that future studies can make modifications guided by evidence. Results from the present investigation highlight an important point of consideration for the field. Specifically, there may be a number of ways to improve EC skills in children that do not necessarily involve training the ability to control behavioral and emotional reactivity. Critical to determining more effective strategies to improve EC is understanding what factors are primarily responsible for the maturation of EC skills in children. In other words, what makes children who are high in EC good at regulating their behaviors and emotions? Do children high in EC simply engage in inhibition or attentional shifting with higher frequency? Do they have more chances to practice engaging in their EC skills? The developmental literature has yet to answer these questions despite extensive knowledge that early EC skills are important to later success in adulthood. Recent research in adults in the EC training literature suggest that persons high in EC-related skills report lower frequency of engaging in EC behaviors such as resisting impulses (Imhoff, Schmidt, and Gerstenberg, 2014), indicating that individuals who are high in EC are not exhibiting increased control of resisting temptations, but rather a decreased experience of temptation overall. Therefore, researchers suggest that improving EC skills may be better served by training individuals to actively avoid temptation instead of reactively attempting to inhibit their impulses.

Conclusions

Identifying valid markers of EC in young children is important to not only further our understanding of the neural processes that support the development of EC skills, but also learn how to better assess changes in these skills that are either associated with maturation or targeted training. Results of the present study suggested that the targeted EC training did not substantially impact the ERN and N2 amplitude. However, it is premature to conclude that the present results support the notion that inducing change in early EC skills is unattainable given the study's design and methodological limitations and weaknesses within the field at large. Instead, the results support that the ERN and N2 are dynamic measures of EC skills in young children. There was some evidence to suggest that the ERN and N2 reflect similar processes and may not be dissociable in this age period. Changes in behavioral measures of EC observed were more common for tasks that demonstrated greater variability in terms of individual differences in task performance. However, given the inconsistent pattern of associations across method and measures, it is unclear what factors contributed to these changes. Significant changes in both neural and behavioral processes have been identified in this developmental period (Casey, Tottenham, Liston, & Durston, 2005; Casey & Somerville, 2011) as it relates to the development of EC skills. Therefore, future research would benefit from continuing to employ multimethod assessment of early EC skills and to develop measures that are sensitive enough to study change within this age period to better understand the coherence between brain and behavior processes that support the maturation of EC.

APPENDICES

APPENDIX A

TABLES

Table A1
Sample Size, Age, and Sex of Total Sample Presented by Wave and Sample Number

		Training	Control	Age (SD)	% Males
Wave 1	Sample 1	5	5	5.09 (0.48)	40.0
Wave 2	Sample 2	6	3	5.11 (0.33)	33.3
	Sample 3	15	14	5.18 (0.34)	62.1
Total		26	22	5.14 (0.37)	52.0

Table A2

Performance on Cognitive and Academic Achievement Measures

	Training		Control		Total		Effect Size
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*
PPVT	111.11	10.27	107.29	13.14	109.10	11.87	0.32
WPPSI-IV Block	8.28	4.08	7.0	3.76	7.76	3.96	0.33
WPPSI-IV Matrix	11.42	2.47	10.0	2.78	10.86	2.66	0.54
SB-5 Nonverbal WM	9.10	2.41	8.56	2.99	8.86	2.65	0.20
SB-5 Verbal WM	10.67	1.39	10.41	2.29	10.55	1.83	0.14
WJ-III Letter Word	106.00	15.82	111.41	27.59	108.42	21.73	0.24
WJ-III Applied Problem	102.86	18.00	107.13	19.29	104.70	18.43	0.23

Note. PPVT = Peabody Picture Vocabulary Test. WPPSI-IV = Weschler Preschool and Primary Scale of Intelligence, 4th Edition. SB-5 = Stanford Binet Intelligence Scales for Early Childhood, 5th Edition. WM = Working Memory. WJ = Woodcock Johnson-III Tests of Achievement.

Table A3

Descriptive Statistics of Behavioral Task Performance Across Different Transfer Levels in Training and Control Groups from Baseline to Post-Assessment

	Training Baseline		Training Post		Effect Size	Control Baseline		Control Post		Effect Size
Near Transfer	*M*	*SD*	*M*	*SD*	*d*	*M*	*SD*	*M*	*SD*	*d*
Walk-a-line	14.18	12.69	28.81	43.02	0.46	14.50	9.62	15.96	14.35	0.12
Simon Says	14.85	4.26	17.04	2.99	0.60	15.94	3.04	16.56	2.53	0.22
Hand Game	14.48	1.97	14.71	0.56	0.16	14.69	0.60	14.63	0.50	-0.11
Backward Digit Span	3.05	1.94	3.71	1.59	0.37	3.38	1.96	3.44	1.93	0.03
Reverse Categorization	11.57	0.87	10.86	2.83	-0.34	11.73	0.59	11.53	0.92	-0.26
Intermediate Transfer										
Bear/Dragon	10.0	0.00	9.95	0.22	-0.32	9.56	1.26	9.38	1.50	-0.13
Knock-Tap	7.38	1.77	7.90	0.30	0.41	7.69	0.48	7.69	0.60	0.00
DCCS Border	6.96	2.71	8.65	2.72	0.62	6.71	2.70	7.50	1.83	0.34
DCCS Total	17.12	5.41	19.96	3.62	0.62	17.31	4.62	19.38	1.78	0.59
Day/Night	13.57	2.42	15.52	1.36	0.99	13.50	3.37	13.63	4.41	0.03
Shape Stroop	11.24	1.48	11.38	1.32	0.10	11.13	1.63	11.25	1.61	0.07
Far-intermediate Transfer										
Snack Delay	0.09	0.61	0.15	0.39	0.12	-0.11	0.69	-0.18	1.02	-0.08
HTKS trials	7.23	7.48	11.69	8.37	0.56	7.50	8.08	6.88	7.43	-0.08
HTKS total	21.69	11.95	29.31	11.45	0.65	23.69	9.52	24.19	9.78	0.05
Whisper Game	28.00	4.36	28.62	5.00	0.13	28.76	4.35	30.00	0.00	0.50
FIST Select 2	8.80	3.66	9.95	3.61	0.32	10.00	3.44	14.38	6.80	0.81
FIST Total	22.35	4.76	23.60	5.05	0.25	23.75	4.84	24.63	5.68	0.17
Tower of Hanoi	3.62	1.53	3.86	1.98	0.14	3.13	1.93	4.19	1.91	0.55
Yes/No	14.29	3.35	14.71	2.47	0.14	15.06	1.44	15.25	1.13	0.15

Note. Italicized text represents tasks that more directly tap attentional control skills. Regular text represents tasks that more directly tap inhibitory control skills. DCCS = Dimensional Card Change Sort. HTKS = Head-Toes-Knees-Shoulders. FIST = Flexible Item Shift Task.

Table A4
Mixed Repeated Measure ANOVA Results for Behavioral Task Measures

	Effect of Time			Time by Group		
Near Transfer	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Walk-a-line	2.42	0.13	0.06	1.62	0.21	0.04
Simon Says	5.42	0.03*	0.12	1.68	0.20	0.04
Hand Game	0.10	0.75	0.00	0.30	0.59	0.01
Backward Digit Span	2.18	0.15	0.06	1.50	0.23	0.04
Reverse Categorization	1.29	0.26	0.04	0.41	0.53	0.01
Intermediate Transfer						
Bear/Dragon	5.18	0.03*	0.13	1.84	0.18	0.05
Knock-Tap	1.30	0.26	0.04	1.30	0.26	0.04
DCCS Border	3.38	0.08†	0.09	0.45	0.51	0.01
DCCS Total	8.27	.007***	0.18	0.21	0.65	0.01
Day/Night	7.05	0.01**	0.17	5.45	0.03*	0.14
Shape Stroop	0.13	0.72	0.00	0.00	0.98	0.00
Far-intermediate Transfer						
Snack Delay	0.01	0.93	0.00	0.19	0.66	0.01
HTKS trials	2.74	0.11	0.06	4.81	0.03*	0.11
HTKS total	7.02	0.01**	0.15	5.40	0.03*	0.12
Whisper Game	0.42	0.52	0.01	0.12	0.73	0.00
FIST Select 2	9.53	.000***	0.22	3.25	0.08†	0.09
FIST Total	0.97	0.33	0.03	0.03	0.86	0.00
Tower of Hanoi	3.61	0.07†	0.09	1.45	0.24	0.04
Yes/No	0.36	0.55	0.01	0.06	0.82	0.00

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. *Italicized text* represents tasks that more directly tap attentional control skills. Regular text represents tasks that more directly tap inhibitory control skills. DCCS = Dimensional Card Change Sort. HTKS = Head-Toes-Knees-Shoulders. FIST = Flexible Item Shift Task.

Table A5

Descriptive Statistics of Behavioral Performance on Go/No-Go and Flanker Tasks in Training and Control Groups from Baseline to Post-Assessment

	Training Baseline		Training Post		Effect Size	Control Baseline		Control Post		Effect Size
Go/No-Go Task	*M*	*SD*	*M*	*SD*	*d*	*M*	*SD*	*M*	*SD*	*d*
Error No-Go trials	23.73	8.21	22.81	10.84	-0.10	26.48	9.15	25.24	10.43	-0.13
Correct Go trials	167.62	38.14	167.46	37.39	0.00	174.05	31.74	172.95	43.71	-0.03
Total accuracy (%)	85.01	7.78	84.46	7.33	-0.07	84.03	6.09	82.69	9.58	-0.17
RT error (ms)	507.93	83.87	505.43	80.54	-0.03	523.81	75.86	522.30	90.16	-0.02
RT correct (ms)	622.92	71.51	603.86	80.27	-0.25	602.02	56.60	583.44	80.06	-0.27
Post-error accuracy (%)	78.14	15.26	80.23	11.86	0.15	82.20	12.81	80.77	13.57	-0.11
RT post error (ms)	549.93	70.64	522.16	57.78	-0.43	540.64	57.98	509.70	67.85	-0.49
RT post correct (ms)	540.86	42.96	522.69	45.30	-0.41	528.23	40.44	509.67	37.61	-0.48
Flanker Task										
Error trials	36.20	17.67	23.84	13.46	-0.79	39.80	22.56	28.25	23.52	-0.50
Correct trials	96.28	42.09	118.48	57.87	0.44	87.75	41.36	112.55	56.76	0.50
Total accuracy (%)	63.71	17.01	73.47	14.79	0.61	60.23	17.95	68.37	18.88	0.44
RT error (ms)	621.98	92.13	614.36	83.20	-0.09	635.26	79.70	623.36	108.71	-0.12
RT correct (ms)	701.91	63.96	705.73	66.77	0.06	691.51	72.94	669.65	88.92	-0.27
Post-error accuracy (%)	61.92	12.26	75.59	15.71	0.97	58.92	12.40	74.94	15.19	1.16
RT post error (ms)	554.67	44.41	574.06	48.30	0.42	547.03	60.50	546.89	67.15	0.00
RT post correct (ms)	572.67	38.98	589.90	32.60	0.48	564.10	40.21	567.63	60.73	0.07

Note. RT = reaction time. Post error trials were defined as trials that were identified as errors (either a false alarm on the Go/No-Go task or error on the Flanker trial) followed by a correct response. The reaction time on post error trials was measured after the error itself. Post correct trials were defined as trials that were identified as correct responses followed by another correct response. The reaction time on post correct trials was measured after the first correct response.

Table A6
Mixed Repeated Measure ANOVA results for Behavioral Performance on Go/No-Go and Flanker Tasks

	Effect of Time			Time by Group		
Go/No-Go Task	F	p	η^2_p	F	p	η^2_p
Error No-Go trials	0.63	0.43	0.01	0.01	0.91	0.00
Correct Go trials	0.02	0.89	0.00	0.01	0.91	0.00
Total accuracy (%)	0.95	0.34	0.02	0.16	0.69	0.00
RT error (ms)	0.04	0.84	0.00	0.00	0.96	0.00
RT correct (ms)	4.73	0.04^{*}	0.10	0.00	0.98	0.00
Post error accuracy (%)	0.03	0.87	0.00	0.73	0.40	0.02
RT post error (ms)	8.88	$.005^{***}$	0.17	0.10	0.94	0.00
RT post correct (ms)	9.79	$.003^{***}$	0.18	0.00	0.97	0.00
Flanker Task						
Error trials	38.84	$.000^{***}$	0.46	0.01	0.94	0.00
Correct trials	22.90	$.000^{***}$	0.34	0.06	0.81	0.00
Total accuracy (%)	27.86	$.000^{***}$	0.38	0.36	0.55	0.01
RT error (ms)	0.31	0.58	0.01	0.12	0.73	0.00
RT correct (ms)	0.09	0.77	0.00	2.17	0.15	0.05
Post error accuracy (%)	55.64	$.000^{***}$	0.55	0.34	0.57	0.01
RT post error (ms)	1.99	0.17	0.04	0.96	0.33	0.02
RT post correct (ms)	3.89	$0.06^{\dagger}$	0.08	1.50	0.23	0.03

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. RT = reaction time.

Table A7

Descriptive Statistics of Response-Locked and Stimulus-Locked Event-Related Potentials on Go/No-Go and Flanker Tasks in Training and Control Groups from Baseline to Post-Assessment

	Training Baseline		Training Post		Effect Size	Control Baseline		Control Post		Effect Size
Go/No-Go	*M*	*SD*	*M*	*SD*	*d*	*M*	*SD*	*M*	*SD*	*d*
ERN	-1.74	5.04	-3.41	5.55	-0.32	-2.93	6.85	-2.64	5.00	0.05
ΔERN	-1.62	3.71	-2.50	4.19	-0.22	-0.90	3.90	-1.14	7.81	-0.04
High Conflict N2	-19.31	5.12	-21.80	6.86	-0.41	-22.11	6.73	-23.07	6.57	-0.14
ΔN2	0.74	3.41	-3.02	4.92	-0.89	-0.28	3.50	-2.08	4.96	-0.42
Fish Flanker										
ERN	-1.23	3.67	-1.93	4.08	-0.18	-0.27	2.95	-1.82	2.76	-0.54
ΔERN	-1.90	3.37	-2.62	4.30	-0.19	-0.87	3.60	-2.20	3.42	-0.38
High Conflict N2	-0.09	3.92	-0.72	3.13	-0.18	-3.42	2.82	-0.64	3.62	0.86
ΔN2	-0.62	2.97	-0.75	2.13	-0.05	-1.27	4.55	-0.09	2.99	0.31

Note. All mean ERP amplitudes are reported at FCz. ERN = Error-related Negativity. ΔERN = ERN Difference (ERN – Correct Related Negativity). ΔN2 = N2 Difference (High Conflict N2 – Low Conflict N2).

Table A8
Mixed Repeated Measure ANOVA Results for Event-Related Potentials Elicited by Go/No-Go and Flanker Tasks

	Effect of Time			Time by Group		
Go No-Go Task	*F*	*p*	η^2_p	*F*	*p*	η^2_p
ERN	0.42	0.52	0.01	0.94	0.34	0.02
ΔERN	0.25	0.62	0.01	0.49	0.49	0.01
High Conflict N2	5.61	0.02*	0.12	0.32	0.57	0.01
ΔN2	14.45	0.00***	0.26	1.42	0.24	0.03
Fish Flanker Task						
ERN	1.90	0.18	0.05	0.38	0.54	0.01
ΔERN	1.71	0.20	0.04	0.16	0.70	0.00
High Conflict N2	4.16	0.05*	0.10	5.69	0.02*	0.13
ΔN2	0.67	0.42	0.02	0.56	0.46	0.02

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. ERN = Error-related Negativity. ΔERN = ERN Difference (ERN – Correct Related Negativity). ΔN2 = N2 Difference (High Conflict N2 – Low Conflict N2).

Table A9

Descriptive Statistics of Parent-Reported Behavior Rating Inventory of Executive Function and Child Behavior Checklist Raw Scores in Training and Control Groups from Baseline to 1-Month Follow-Up

	Training Baseline		Training 1-month		Effect Size	Control Baseline		Control 1-month		Effect Size
BRIEF	*M*	*SD*	*M*	*SD*	*d*	*M*	*SD*	*M*	*SD*	*d*
Inhibit	29.52	5.95	22.57	3.92	-1.38	26.24	5.31	23.27	6.80	-0.49
Shift	16.38	3.85	13.71	2.80	-0.79	16.29	3.44	14.07	3.95	-0.60
Emotional Control	19.24	3.99	14.67	2.67	-1.35	18.53	3.56	16.73	5.55	-0.39
Working Memory	18.19	3.86	14.43	2.87	-1.11	17.65	3.62	15.20	4.04	-0.64
Plan/Organize	18.10	2.64	14.67	2.29	-1.39	17.82	2.90	14.93	2.96	-0.99
Total GEC	101.43	14.54	80.05	10.13	-1.71	96.53	12.51	84.20	18.93	-0.77
CBCL										
Anxious/Depressed	0.46	0.39	0.22	0.22	-0.76	0.37	0.27	0.31	0.27	-0.22
Withdrawn/Depressed	0.22	0.19	0.11	0.12	-0.69	0.25	0.24	0.16	0.19	-0.42
Somatic Problems	0.19	0.20	0.07	0.09	-0.77	0.16	0.23	0.10	0.14	-0.32
Social Problems	0.35	0.27	0.18	0.14	-0.79	0.31	0.22	0.28	0.25	-0.13
Attention Problems	0.54	0.35	0.32	0.27	-0.70	0.47	0.27	0.36	0.32	-0.37
Rule-Breaking	0.23	0.13	0.14	0.13	-0.69	0.21	0.17	0.19	0.24	-0.10
Aggressive Behaviors	0.42	0.30	0.24	0.18	-0.73	0.40	0.31	0.33	0.33	-0.22
Internalizing Scale	0.29	0.23	0.13	0.12	-0.87	0.26	0.20	0.19	0.16	-0.39
Externalizing Scale	0.32	0.20	0.19	0.14	-0.75	0.30	0.23	0.26	0.28	-0.16

Note. BRIEF = Behavior Rating Inventory of Executive Function, Preschool Version. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. Internalizing Scale = Sum score of Anxious/Depressed, Withdrawn/Depressed, and Somatic Problems. Externalizing Scale = Sum score of Rule-Breaking and Aggressive Behaviors.

Table A10

Mixed Repeated Measure ANOVA Results for Parent-Reported Behavior Rating Inventory of Executive Function and Child Behavior Checklist

	Baseline to 1-month						1-month to 3-month					
	Effect of Time			Time by Group			Effect of Time			Time by Group		
BRIEF	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Inhibit	35.19	.000***	0.51	7.64	0.01**	0.18	0.01	0.92	0.00	0.28	0.60	0.01
Shift	22.29	.000***	0.40	0.46	0.51	0.01	0.40	0.53	0.01	0.40	0.53	0.01
Emotional Control	18.86	.000***	0.36	2.89	0.10†	0.08	0.50	0.49	0.02	0.02	0.89	0.00
Working Memory	27.28	.000***	0.45	1.50	0.23	0.04	2.90	0.10†	0.09	0.47	0.50	0.02
Plan/Organize	39.71	.000***	0.54	0.32	0.58	0.01	1.97	0.17	0.06	0.52	0.48	0.02
GEC	44.45	.000***	0.57	3.77	0.06†	0.10	0.35	0.56	0.01	0.18	0.67	0.01
CBCL												
Anxious/Depressed	10.09	.003***	0.20	3.53	0.07†	0.08	3.05	0.09†	0.09	0.08	0.78	0.00
Withdrawn/Depressed	11.78	.001***	0.22	0.60	0.44	0.01	4.54	0.04*	0.13	0.90	0.35	0.03
Somatic Problems	10.19	.003***	0.20	0.52	0.48	0.01	1.18	0.29	0.04	0.61	0.44	0.02
Social Problems	8.88	.005***	0.18	3.66	0.06†	0.08	8.56	0.01**	0.22	0.54	0.47	0.02
Attention Problems	25.41	.000***	0.38	3.77	0.06†	0.08	0.54	0.47	0.02	0.06	0.82	0.00
Rule Breaking	7.38	0.01**	0.15	1.87	0.18	0.04	0.10	0.76	0.00	1.13	0.30	0.04
Aggressive	18.16	.000***	0.31	1.82	0.18	0.04	0.05	0.83	0.00	0.09	0.77	0.00
Internalizing Scale	17.96	.000***	0.31	2.74	0.11	0.06	4.14	0.05*	0.12	0.01	0.92	0.00
Externalizing Scale	17.18	.000***	0.30	2.38	0.13	0.06	0.13	0.72	0.00	0.13	0.72	0.00

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Function, Preschool Version. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. Internalizing Scale = Sum score of Anxious/Depressed, Withdrawn/Depressed, and Somatic Problems. Externalizing Scale = Sum score of Rule-Breaking and Aggressive Behaviors.

Table A11

Descriptive Statistics of Teacher-Reported Behavior Rating Inventory of Executive Function and Child Behavior Checklist Raw Scores at 3-Month Follow-up

	Training 3-month		Control 3-month		Effect Size
BRIEF	*M*	*SD*	*M*	*SD*	*d*
Inhibit	20.47	7.79	19.90	4.23	-0.09
Shift	11.76	2.59	13.10	4.09	0.39
Emotional Control	13.00	4.66	13.30	4.32	0.07
Working Memory	13.00	3.43	13.70	2.21	0.24
Plan/Organize	12.53	3.92	12.30	1.64	-0.08
Total	70.76	20.50	72.30	11.25	0.09
CBCL					
Anxious/Depressed	0.20	0.29	0.20	0.25	0.00
Withdrawn/Depressed	0.15	0.26	0.27	0.21	0.51
Somatic Problems	0.02	0.05	0.06	0.09	0.55
Social Problems	0.13	0.23	0.13	0.19	0.00
Attention Problems	0.35	0.42	0.29	0.37	-0.15
Rule Breaking	0.06	0.11	0.04	0.09	-0.20
Aggressive	0.14	0.18	0.05	0.07	-0.66
Internalizing Scale	0.12	0.16	0.18	0.13	0.41
Externalizing Scale	0.10	0.13	0.05	0.05	-0.51

Note. BRIEF = Behavior Rating Inventory of Executive Function, Preschool Version. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. Internalizing Scale = Sum score of Anxious/Depressed, Withdrawn/Depressed, and Somatic Problems. Externalizing Scale = Sum score of Rule-Breaking and Aggressive Behaviors.

Table A12

Bivariate Correlations Between Changes in Neurophysiological Measures of EC and Changes in Behavioral Measures of Attentional Control from Baseline to Post-Assessment

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.96**	--						
3. HC N2	0.16	0.11	--					
4. ΔN2	0.11	0.07	0.85**	--				
Fish Flanker								
5. ERN	0.14	0.13	0.09	0.16	--			
6. ΔERN	0.13	0.10	0.19	0.22	0.91**	--		
7. HC N2	0.01	-0.01	0.22	0.18	0.01	0.06	--	
8. ΔN2	-0.16	-0.17	0.03	0.07	0.19	0.19	0.72**	--
9. Backward Digit Span	0.06	0.04	0.17	0.19	-0.14	-0.04	-0.11	-0.09
10. Reverse Categorization	0.07	0.06	0.23	0.13	-0.21	-0.10	0.03	-0.03
11. DCCS Border	0.22	0.19	0.05	0.05	0.35*	0.31†	0.05	0.12
12. DCCS Total	0.16	0.16	0.22	0.17	0.25	0.20	0.20	0.35*
13. Day/Night	-0.19	-0.20	-0.05	-0.05	0.28	0.25	-0.03	0.09
14. Shape Stroop	0.04	0.00	0.03	-0.07	0.01	0.13	0.03	-0.15
15. FIST Select 2	-0.02	-0.06	-0.10	-0.19	-0.08	-0.19	0.15	0.16
16. FIST Total	-0.35*	-0.30†	-0.07	-0.07	0.06	0.02	0.06	0.08
17. Tower of Hanoi	0.15	0.11	0.27	0.05	-0.10	-0.23	0.22	0.32†
18. Yes/No	-0.06	-0.14	0.16	0.15	0.30†	0.26	0.17	0.10

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Shift Task.

Table A13
Bivariate Correlations Between Changes in Neurophysiological Measures of EC and Changes in Behavioral Measures of Inhibitory Control from Baseline to Post-Assessment

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.96**	--						
3. HC N2	0.16	0.11	--					
4. ΔN2	0.11	0.07	0.85**	--				
Fish Flanker								
5. ERN	0.14	0.13	0.09	0.16	--			
6. ΔERN	0.13	0.10	0.19	0.22	0.91**	--		
7. HC N2	0.01	-0.01	0.22	0.18	0.01	0.06	--	
8. ΔN2	-0.16	-0.17	0.03	0.07	0.19	0.19	0.72**	--
9. Walk-a-line	-0.02	0.04	-0.14	-0.07	0.11	0.03	-0.02	0.03
10. Simon Says	-0.11	-0.03	-0.10	-0.17	-0.18	-0.19	-0.10	0.03
11. Hand Game	0.28	0.17	-0.01	-0.10	0.00	0.14	0.02	-0.05
12. *Bear Dragon*	0.19	0.21	-0.05	0.04	0.12	0.10	-0.16	-0.21
13. *Knock-Tap*	-0.18	-0.17	0.06	0.18	0.08	-0.08	-0.26	-0.03
14. *Snack Delay*	-0.15	-0.15	-0.12	-0.05	0.06	0.02	-0.09	-0.19
15. HTKS trials	0.26	0.28†	0.12	-0.03	0.01	0.08	0.05	0.29†
16. HTKS total	0.25	0.24	0.02	-0.16	0.01	0.04	-0.07	0.13
17. Whisper Game	0.05	0.10	0.11	0.06	-0.06	0.00	-0.13	0.00

Note. ***p < .001. **p < 0.01. *p < 0.05. †p < 0.10. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. HTKS = Head-Toes-Knees-Shoulders task.

Table A14

Bivariate Correlations Between Changes in Neurophysiological Measures of EC and Changes in Parent-Reported Behavior Rating Inventory of Executive Function Scores

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.96**	--						
3. HC N2	0.16	0.11	--					
4. ΔN2	0.11	0.07	0.85**	--				
Fish Flanker								
5. ERN	0.14	0.13	0.09	0.16	--			
6. ΔERN	0.13	0.10	0.19	0.22	0.91**	--		
7. HC N2	0.01	-0.01	0.22	0.18	0.01	0.06	--	
8. ΔN2	-0.16	-0.17	0.03	0.07	0.19	0.19	0.72**	--
9. BRIEF Inhibit	-0.21	-0.17	-0.02	0.06	-0.18	-0.18	-0.28	-0.17
10. BRIEF Shift	-0.03	0.04	-0.05	0.03	0.20	0.19	-0.05	0.02
11. BRIEF Emotional Control	-0.00	-0.01	0.15	0.25	0.04	-0.03	-0.13	-0.04
12. BRIEF Working Memory	-0.14	-0.06	-0.21	-0.22	-0.21	-0.21	-0.04	-0.05
13. BRIEF Plan/Organize	-0.25	-0.23	-0.13	0.02	0.02	0.04	-0.16	-0.06
14. BRIEF Total GEC	-0.15	-0.12	-0.05	0.05	-0.03	-0.08	-0.19	-0.09

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Function, Preschool Version. GEC = Global Executive Composite. BRIEF change scores were calculated from baseline to 1-month follow-up.

Table A15

Bivariate Correlations Between Changes in Neurophysiological Measures of EC and Changes in Parent-Reported Child Behavior Checklist Scores

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.96**	--						
3. HC N2	0.16	0.11	--					
4. ΔN2	0.11	0.07	0.85**	--				
Fish Flanker								
5. ERN	0.14	0.13	0.09	0.16	--			
6. ΔERN	0.13	0.10	0.19	0.22	0.91**	--		
7. HC N2	0.01	-0.01	0.22	0.18	0.01	0.06	--	
8. ΔN2	-0.16	-0.17	0.03	0.07	0.19	0.19	0.72**	--
9. CBCL AnxDep	-0.03	0.02	-0.04	-0.05	-0.05	-0.07	-0.08	-0.01
10. CBCL WithDep	0.02	0.02	0.05	-0.03	0.19	0.16	0.06	-0.14
11. CBCL Somatic	-0.08	-0.13	0.10	0.02	0.02	-0.08	0.15	0.24
12. CBCL Soc	0.27†	0.22	0.32*	0.21	-0.07	-0.07	0.07	0.01
13. CBCL Attn	-0.07	-0.07	0.11	0.17	-0.02	-0.08	-0.10	-0.10
14. CBCL Ruleb	0.27†	0.29†	0.13	0.19	-0.04	0.00	-0.09	-0.29†
15. CBCL Agg	-0.04	-0.03	0.14	0.14	-0.15	-0.19	-0.15	-0.02
16. CBCL Int	-0.05	-0.03	0.02	-0.03	0.04	-0.02	0.03	0.03
17. CBCL Ext	0.09	0.11	0.15	0.18	-0.12	-0.13	-0.14	-0.15

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale. CBCL change scores were calculated from baseline to 1-month follow-up.

Table A16

Bivariate Correlations Between Changes in Behavioral Measures of Attentional Control and Changes in Measures of Inhibitory Control from Baseline to Post-Assessment

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. BDS	--																	
2. RC	0.12	--																
3. DCCS Border	-0.29†	-0.23	--															
4. DCCS Total	-0.28†	-0.14	0.89**	--														
5. Day/Night	0.02	0.01	0.09	0.07	--													
6. Shape Stroop	0.15	-0.02	-0.32†	-0.23	0.29†	--												
7. FIST Select 2	-0.03	-0.02	0.14	0.16	-0.22	-0.32†	--											
8. FIST Total	-0.03	0.10	-0.20	0.07	0.16	0.14	0.24	--										
9. Tower of Hanoi	-0.09	0.32†	-0.06	0.03	-0.20	-0.20	0.15	-0.12	--									
10. Yes/No	-0.06	-0.02	-0.01	-0.03	0.36*	0.32†	-0.11	-0.11	0.06	--								
11. Walk-a-line	-0.07	-0.22	-0.11	-0.11	0.11	0.13	-0.02	-0.04	0.01	0.22	--							
12. Simon Says	0.30†	0.09	-0.16	-0.13	0.07	-0.04	0.02	0.08	0.01	-0.12	-0.28†	--						
13. Hand Game	-0.17	0.10	0.36*	0.24	0.00	-0.01	-0.04	-0.34*	0.20	0.03	-0.18	-0.16	--					
14. Bear Dragon	-0.14	-0.02	0.24	0.26	-0.15	0.06	0.01	-0.15	-0.32*	-0.05	0.03	-0.31†	0.02	--				
15. Knock-Tap	-0.07	-0.05	-0.02	0.34*	0.21	-0.32†	0.14	0.20	0.09	-0.04	-0.18	0.05	0.10	0.08	--			
*16. Snack Delay**	-0.03	-0.09	-0.28	-0.32†	0.14	0.42*	-0.28†	0.04	-0.16	0.12	0.09	-0.06	-0.01	0.04	-0.10	--		
17. HTKS trials	0.04	0.09	0.26	0.30†	-0.12	-0.10	0.00	-0.26	0.28†	0.20	0.17	0.15	0.31†	0.15	-0.02	-0.39*	--	
18. HTKS total	0.00	0.13	0.30†	0.32*	0.09	0.06	0.11	0.01	0.18	0.15	0.03	0.27†	0.23	0.16	0.21	-0.33*	0.72**	--
19. Whisper Game	0.11	0.33†	-0.29	0.09	0.25	-0.16	0.19	0.12	0.26	-0.04	0.10	0.19	0.00	-0.13	-0.20	0.19	0.02	0.22

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. First 10 tasks listed are behavioral tasks designed to elicit attentional control skills. Variables 11-19 represent behavioral tasks designed to elicit inhibitory control skills. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. *Snack delay scores are reverse coded so that higher scores mean better effortful control skills. BDS = Backward Digit Span. RC = Reverse Categorization. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Selection Task. HTKS = Head-Toes-Knees-Shoulders task.

Table A17

Bivariate Correlations Between Changes in Parent-Reported Behavior Rating Inventory of Executive Functions and Child Behavior Checklist Scores from Baseline to 1-Month Follow-Up

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. BRIEF Inhibit	--													
2. BRIEF Shift	0.42*	--												
3. BRIEF EmoC	0.66**	0.69**	--											
4. BRIEF WM	0.68**	0.42*	0.43**	--										
5. BRIEF PlanO	0.56**	0.33*	0.37*	0.74**	--									
6. BRIEF GEC	0.88**	0.70**	0.82**	0.81**	0.73**	--								
7. CBCL AnxDep	0.28	0.06	0.13	0.21	0.25	0.24	--							
8. CBCL WithDep	0.10	0.18	0.06	0.06	0.00	0.10	0.60**	--						
9. CBCL Somatic	0.09	0.03	0.28†	0.03	0.01	0.13	0.27†	0.18	--					
10. CBCL Soc	0.25	0.13	0.28†	0.08	0.02	0.22	0.47**	0.51**	0.24	--				
11. CBCL Attn	0.44**	0.08	0.20	0.26	0.24	0.33*	0.47**	0.36*	0.22	0.60**	--			
12. CBCL Ruleb	0.22	0.10	0.16	0.17	0.08	0.19	0.34*	0.46*	-0.11	0.58**	0.46**	--		
13. CBCL Agg	0.44**	0.04	0.26	0.18	0.02	0.28†	0.37*	0.38*	0.29†	0.61**	0.56**	0.55**	--	
14. CBCL Int	0.23	0.10	0.20	0.16	0.15	0.22	0.90**	0.76**	0.58**	0.54**	0.48**	0.32*	0.45**	--
15. CBCL Ext	0.40**	0.07	0.25	0.20	0.05	0.28†	0.40**	0.46**	0.15	0.67**	0.59**	0.82**	0.93**	0.45**

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Functions. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organize. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale.

Table A18
Bivariate Correlations Between Changes in Behavioral Measures of Attentional Control and Changes in Parent-Reported Behavior Rating Inventory of Executive Functions and Child Behavior Checklist Scores

	16	17	18	19	20	21	22	23	24	25
1. BRIEF Inhibit	0.07	-0.15	-0.13	0.02	-0.22	0.15	0.15	0.33†	-0.10	-0.15
2. BRIEF Shift	0.00	0.05	0.01	0.25	-0.03	-0.04	0.10	0.42*	-0.06	0.10
3. BRIEF EmoC	0.08	-0.05	0.03	0.27	-0.06	0.14	-0.01	0.24	-0.11	0.22
4. BRIEF WM	0.11	0.10	-0.06	0.13	-0.09	0.10	0.13	0.48**	-0.12	-0.13
5. BRIEF PlanO	0.08	-0.06	0.02	0.10	-0.03	-0.08	0.11	0.27	-0.29†	-0.07
6. BRIEF GEC	0.09	-0.04	-0.05	0.18	-0.13	0.08	0.25	0.44**	-0.16	-0.07
7. CBCL AnxDep	0.17	-0.03	-0.15	-0.19	-0.08	-0.14	-0.12	-0.10	-0.07	0.15
8. CBCL WithDep	0.30†	0.12	-0.18	-0.28	-0.15	-0.16	-0.11	-0.05	0.02	-0.13
9. CBCL Somatic	-0.02	-0.25	-0.01	-0.01	0.15	-0.12	-0.24	-0.24	0.15	0.15
10. CBCL Soc	0.17	0.04	0.13	0.10	-0.24	-0.11	-0.17	-0.04	0.24	-0.14
11. CBCL Attn	0.06	-0.05	-0.06	-0.01	-0.11	0.14	0.13	0.22	-0.20	-0.20
12. CBCL Ruleb	0.32†	0.12	-0.26	-0.14	-0.41*	-0.23	0.01	0.07	0.02	-0.23
13. CBCL Agg	0.24	-0.19	-0.04	0.00	-0.36*	0.08	0.07	0.13	0.22	-0.11
14. CBCL Int	0.19	-0.07	-0.15	-0.21	-0.04	-0.18	-0.20	-0.17	0.02	0.11
15. CBCL Ext	0.31	-0.07	-0.14	-0.06	-0.43*	-0.06	0.05	0.12	0.15	-0.18
16. BDS	--	--	--	--	--	--	--	--	--	--
17. RC	--	--	--	--	--	--	--	--	--	--
18. DCCS Border	--	--	--	--	--	--	--	--	--	--
19. DCCS Total	--	--	--	--	--	--	--	--	--	--
20. Day/Night	--	--	--	--	--	--	--	--	--	--
21. Shape Stroop	--	--	--	--	--	--	--	--	--	--
22. FIST Select 2	--	--	--	--	--	--	--	--	--	--
23. FIST Total	--	--	--	--	--	--	--	--	--	--
24. Tower of Hanoi	--	--	--	--	--	--	--	--	--	--
25. Yes/No	--	--	--	--	--	--	--	--	--	--

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. Bivariate correlations between changes in behavioral measures of attentional control skills are presented in Table A16, and therefore not listed in the present table. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. BRIEF = Behavior Rating Inventory of Executive Functions. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organize. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale. BDS = Backward Digit Span. RC = Reverse Categorization. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Selection Task.

Table A19

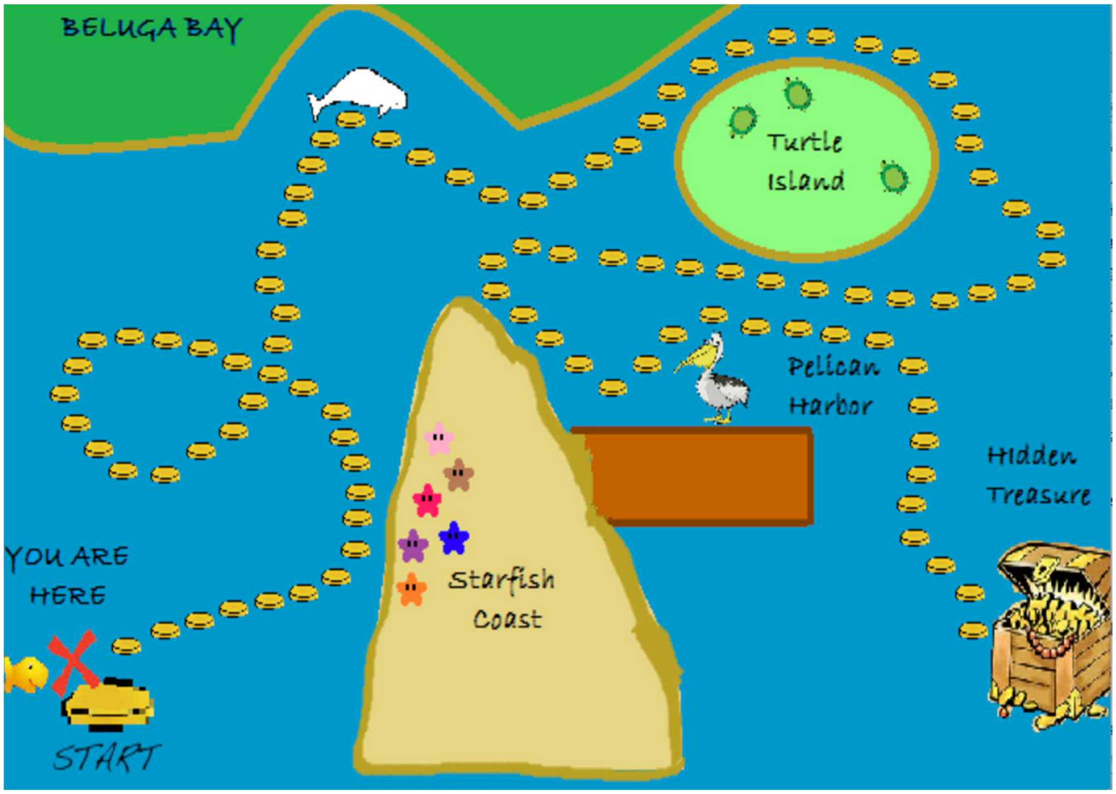
Bivariate Correlations Between Changes in Behavioral Measures of Inhibitory Control and Changes in Parent-Reported Behavior Rating Inventory of Executive Functions and Child Behavior Checklist Scores

	16	17	18	19	20	21	22	23	24
1. BRIEF Inhibit	-0.35*	0.11	-0.26	-0.08	0.17	-0.05	-0.28†	-0.01	0.29†
2. BRIEF Shift	-0.08	-0.02	-0.44**	-0.05	0.09	-0.21	-0.17	-0.07	0.00
3. BRIEF EmoC	-0.31†	-0.08	-0.27	0.12	0.18	-0.22	-0.19	0.07	0.21
4. BRIEF WM	-0.27	0.17	-0.31†	-0.02	0.14	0.10	-0.19	0.00	0.26
5. BRIEF PlanO	-0.28	-0.04	-0.21	-0.09	0.19	0.04	-0.25	-0.21	0.11
6. BRIEF GEC	-0.34*	0.04	-0.36*	-0.02	0.20	-0.09	-0.27†	-0.04	0.24
7. CBCL AnxDep	0.11	0.16	-0.23	-0.20	-0.07	0.11	-0.05	0.07	0.21
8. CBCL WithDep	0.00	0.28†	-0.16	-0.17	-0.19	0.04	-0.13	0.00	0.19
9. CBCL Somatic	-0.03	-0.10	0.12	-0.12	-0.02	0.18	-0.04	0.04	-0.01
10. CBCL Soc	-0.38*	0.18	0.07	-0.27	-0.24	-0.06	-0.01	0.15	0.20
11. CBCL Attn	-0.40*	0.12	-0.26	-0.05	-0.01	0.17	-0.31†	-0.07	0.31†
12. CBCL Ruleb	-0.32†	0.23	-0.09	-0.05	-0.13	0.15	-0.08	-0.06	0.20
13. CBCL Agg	-0.32*	0.43**	-0.12	-0.10	-0.01	-0.03	0.16	0.26	0.23
14. CBCL Int	0.05	0.15	-0.14	-0.22	-0.11	0.14	-0.08	0.05	0.18
15. CBCL Ext	-0.37*	0.39*	-0.12	-0.09	-0.07	0.05	0.07	0.15	0.25
16. Walk-a-line	--	--	--	--	--	--	--	--	--
17. Simon Says	--	--	--	--	--	--	--	--	--
18. Hand Game	--	--	--	--	--	--	--	--	--
19. *Bear Dragon*	--	--	--	--	--	--	--	--	--
20. *Knock-Tap*	--	--	--	--	--	--	--	--	--
21. *Snack Delay**	--	--	--	--	--	--	--	--	--
22. HTKS trials	--	--	--	--	--	--	--	--	--
23. HTKS total	--	--	--	--	--	--	--	--	--
24. Whisper Game	--	--	--	--	--	--	--	--	--

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. Bivariate correlations between changes in behavioral measures of inhibitory control skills are presented in Table A16, and therefore not listed in the present table. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. BRIEF = Behavior Rating Inventory of Executive Functions. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organize. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale. *Snack delay scores are reverse coded so that higher scores mean better effortful control skills. HTKS = Head-Toes-Knees-Shoulders task.

APPENDIX B

FIGURES



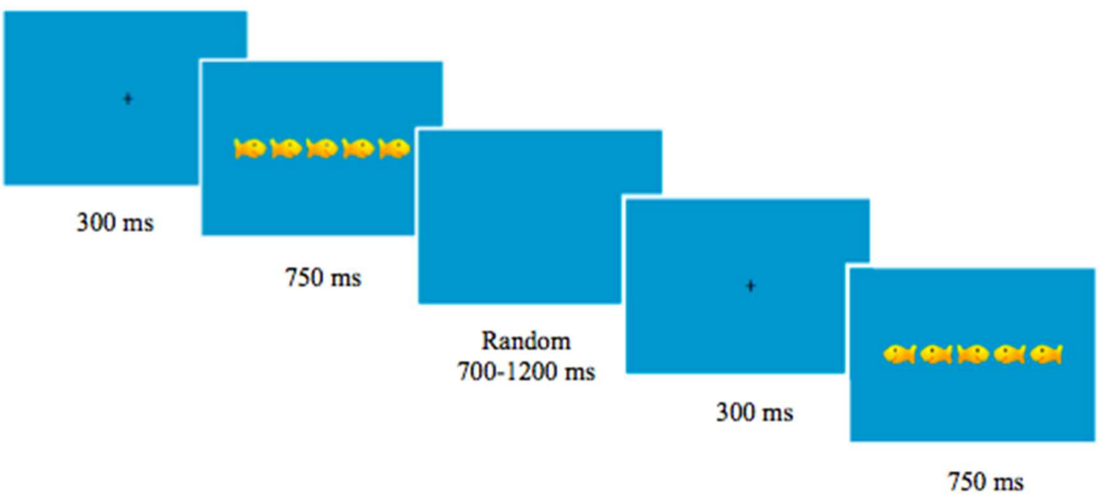


Figure B1. Top: Treasure map slide on the Flanker task. Bottom: Sample trials of the Flanker task

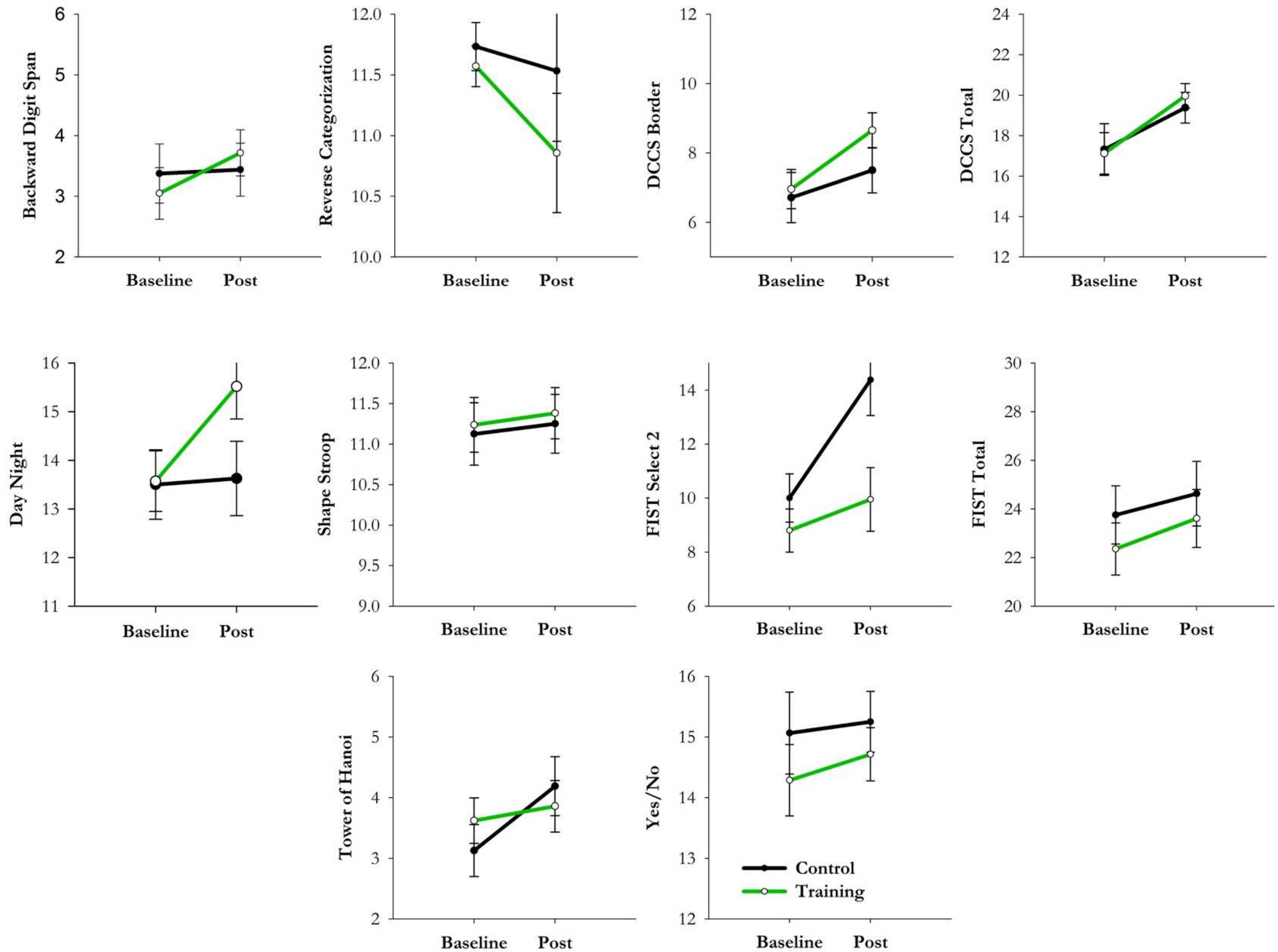


Figure B2. Graphical depiction of mean-level changes in behavioral tasks designed to measure attentional control skills from baseline to post-assessment in training (green) compared to control group (black). Standard error bars are shown. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Selection Task.

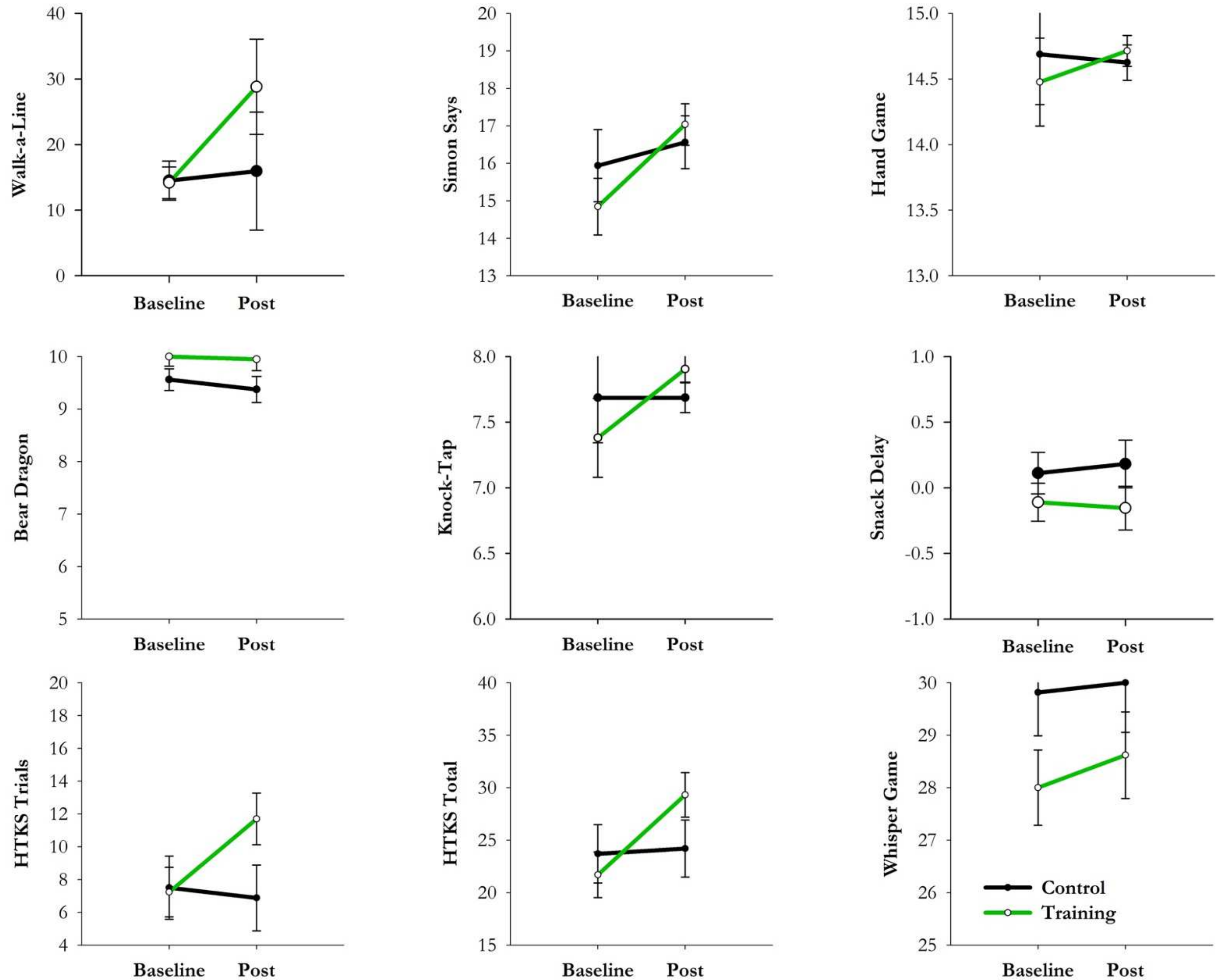


Figure B3. Graphical depiction of mean-level changes in behavioral tasks designed to measure inhibitory control skills from baseline to post-assessment in training (green) compared to control group (black). Standard error bars are shown. HTKS = Head-Toes-Knees-Shoulders Task.

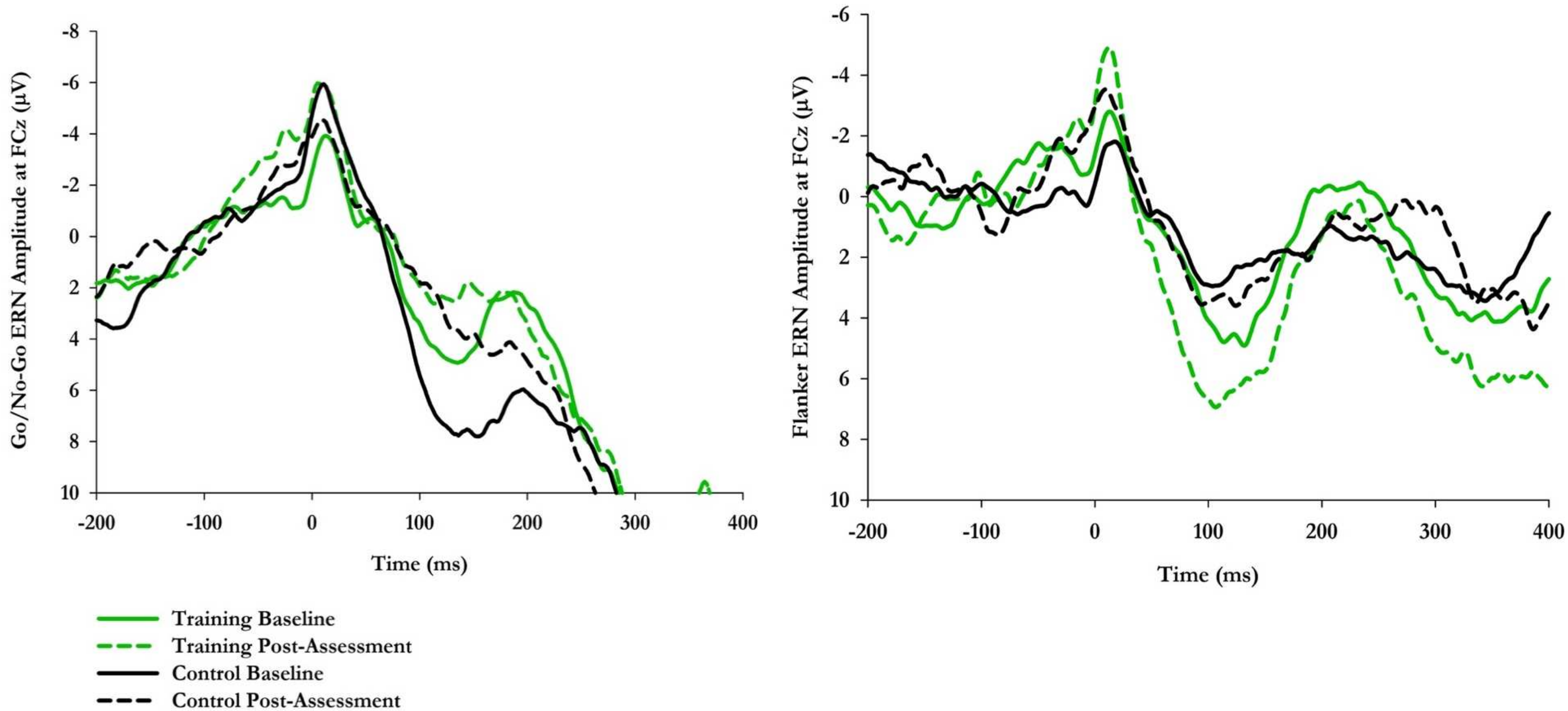


Figure B4. Error-related brain activity for Go/No-Go (left) and Flanker (right) tasks in training (green) compared to control (black) groups at baseline (solid line) and post-assessment (dashed line) time points. Time 0 represents response onset.

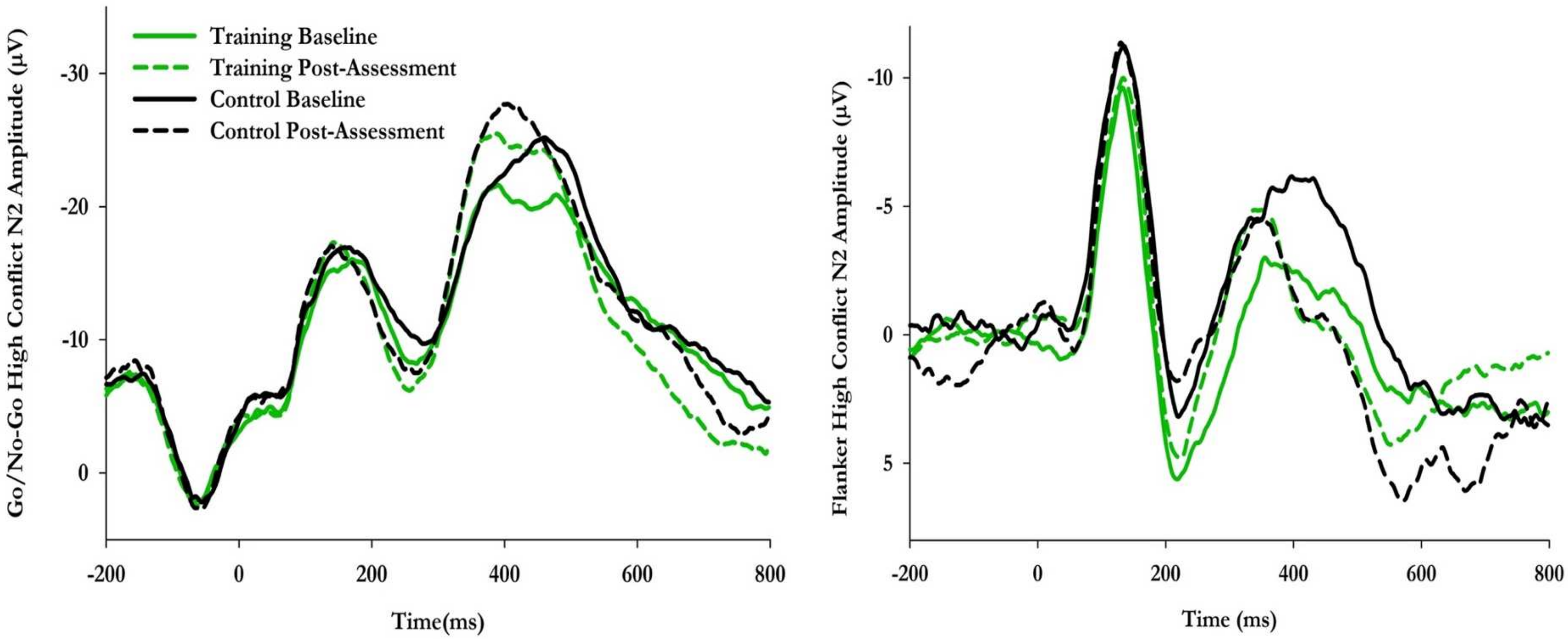


Figure B5. Stimulus-locked brain activity on high conflict trials for Go/No-Go (left) and Flanker (right) tasks in training (green) compared to control (black) groups at baseline (solid line) and post-assessment (dashed line) time points. Time 0 represents stimulus onset.

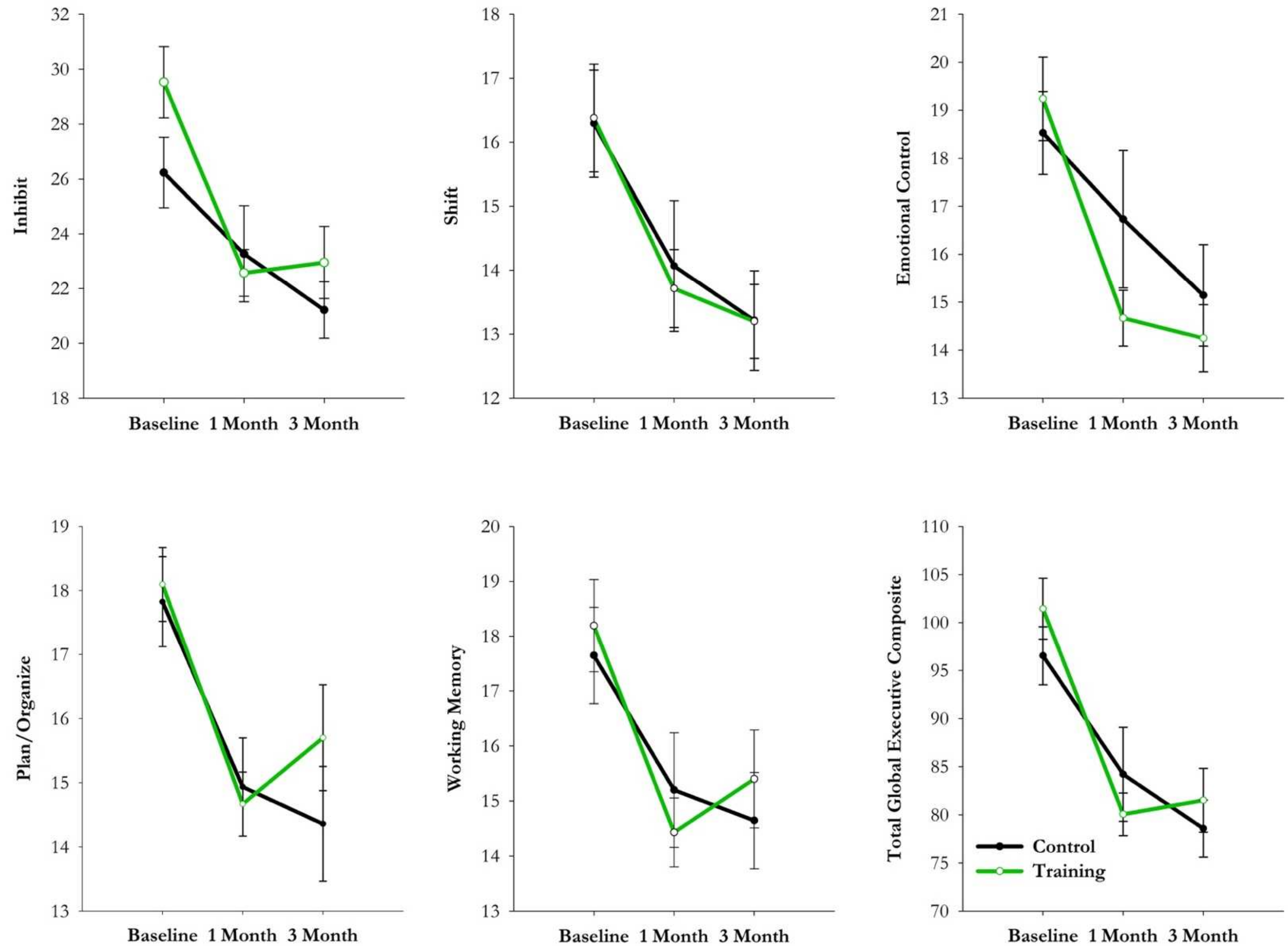


Figure B6. Graphical depiction of mean-level changes in parent-reported Behavior Rating Inventory of Executive Function (BRIEF) scores across baseline, 1-month follow-up, and 3-month follow-up time points in training (green) compared to control group (black). Standard error bars are shown.

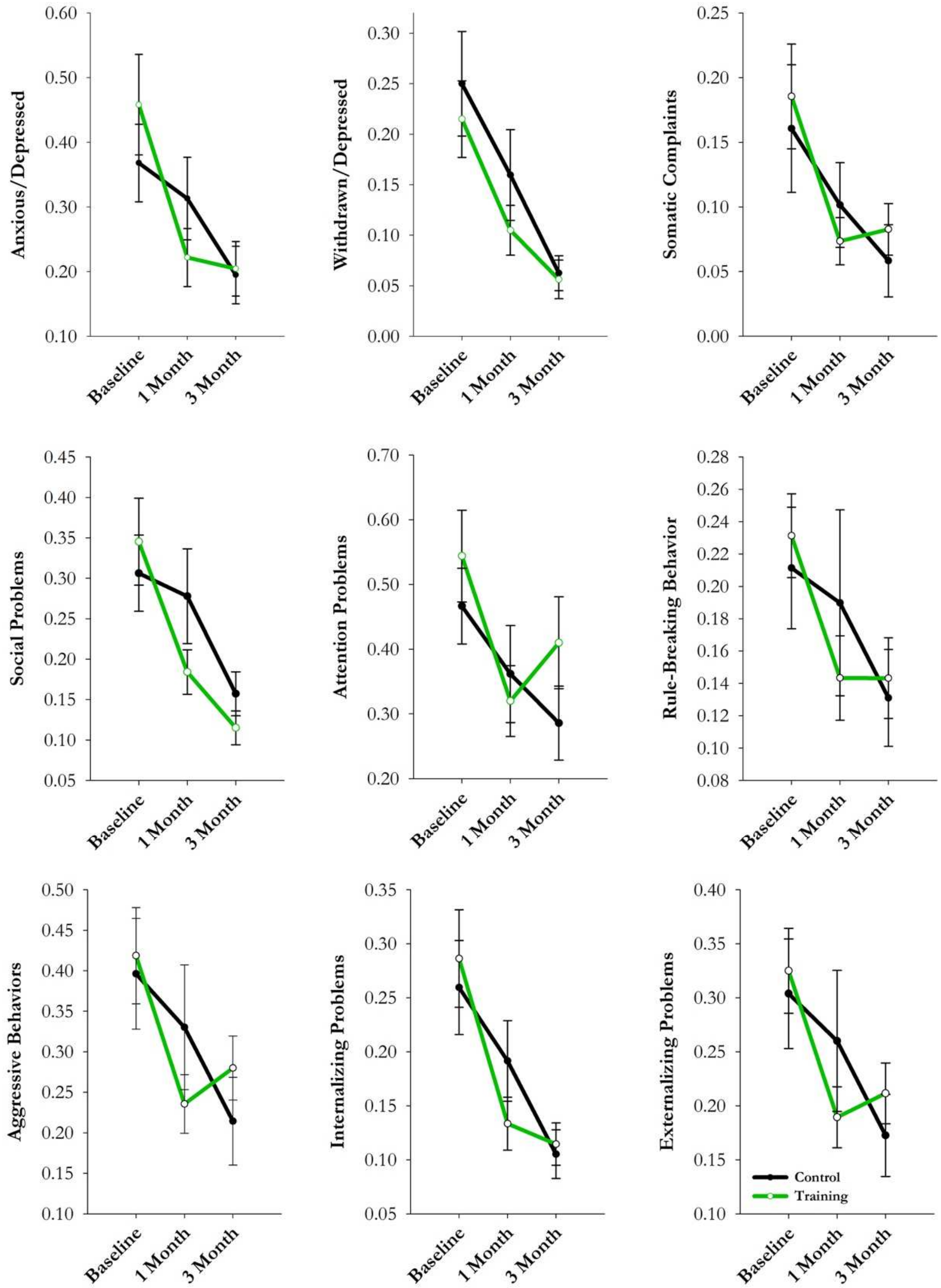


Figure B7. Graphical depiction of mean-level changes in parent-reported Child Behavior Checklist (CBCL) scores across baseline, 1-month follow-up, and 3-month follow-up time points in training (green) compared to control group (black). Standard error bars are shown.

APPENDIX C

A PRIORI PREDICTIONS AND STUDY RESULTS

Table C1

A Priori Predictions and Study Results

Method/Measure	A Priori Prediction[1]	Main Effect of Time[2]	Time By Group Interaction Effects[3]
Behavioral			
AC Near transfer	↑↑↑	--	--
AC Intermediate transfer	↑↑	↑	↑
AC Far-intermediate	↑	↑	↑†
IC Near transfer	↑↑↑	↑	--
IC Intermediate transfer	↑↑	↓	--
IC Far-intermediate	↑	↑	↑
Neurophysiological			
Go/No-Go bx perf	↑↑	--	--
Go/No-Go ERN	↑	--	--
Go/No-Go N2	↓	↑↑	--
Flanker bx perf	↑↑	↑↑	--
Flanker ERN	↑	--	--
Flanker N2	↓	↑	↓
Informant-report			
Parent BRIEF	↓↓	↓↓↓	↓
Parent CBCL	↓↓	↓↓↓	↓†
Teacher BRIEF	Lower	N/A	--
Teacher CBCL	Lower	N/A	--

Note. [1]A priori predictions are summarized as time by group interaction effects that favor training groups; control groups were anticipated to demonstrate no significant change across all methods and measures (e.g., ↑↑↑ on behavioral measures denotes a large effect of increased performance in training groups whereas control groups exhibit no change). Strength of predicted effect is indicated by number of arrows, with more arrows representing a larger effect or more consistent pattern of effects (e.g., ↑ = small effect, ↑↑ = moderate effect, ↑↑↑ = large effect). [2]Results depicted are summarized from mixed ANOVA analyses testing the main effect of time. More consistent pattern of results are indicated by more arrows. [3]Summarized results are from mixed ANOVA time by group interaction effects (with the exception of teacher-reported results, which were independent t-test results). Green arrows (↑) denote that effects are specific to the training group whereas red arrows (↑) denote that effects are specific to the control group. Unless otherwise noted, listed effects are significant (trend-level effects are denoted by the following symbol: "**†**")

APPENDIX D

MEASURES BY TRANSFER TYPE

Table D1

Near, Intermediate, Far-intermediate, and Far Transfer Measures Collected

	Reference	Overall Adjustment	Fluid Reas.	Recep. Vocab	Working Memory		Inhibitory Control	Attentional Control
Cognitive Abilities					Verbal	Vis-Spat.		
PPVT	Dunn & Dunn, 2007			✓				
WPPSI Block Design	Weschler, 2002		✓					
WPPSI Matrix Reasoning	Weschler, 2002		✓					
Stanford Binet Block Span	Roid, 2005					✓		
Stanford Binet Last-Word	Roid, 2005				✓			
Behavioral								
Walk-a-line	Kochanska, Murray, & Coy 2000						✓	
Simon Says	Strommen, 1972						✓	
Hand Game	Hughes 1998						✓	
Bear/Dragon	Reed et al., 1994						✓	
Knock-Tap	Klenberg et al., 2001						✓	
Snack Delay	Murray & Kochanska, 2002						✓	
HTKS	Ponitz et al., 2008						✓	
Whisper Game	Kochanska et al., 1996						✓	
Red Light/Purple Light	Tominey & McClelland, 2011						✓	
Freeze Game	Tominey & McClelland, 2011						✓	
Jumping Game	Tominey & McClelland, 2011				✓		✓	
Color Matching Freeze	Tominey & McClelland, 2011					✓	✓	
Drum Beats	Tominey & McClelland, 2011						✓	
Backward Digit Span	Davis & Pratt, 1996							✓
Reverse Categorization	Carlson, Mandell, & Williams, 2004							✓
DCCS	Zelazo, 2006							✓
Day/Night	Gerstadt et al., 1994							✓
Shape Stroop	Kochanska et al., 2000							✓
FIST	Jacques & Zelazo, 2001							✓
Tower of Hanoi	Welsh et al., 1991							✓
Yes/No Task	Wolfe & Bell, 2004							✓

Table D1 (cont'd)

Neurophsyiological								
Go/No-Go	Grammer et al., 2014						✓	
Flanker	Lo et al., 2015							✓
Informant-Reported								
CBCL	Achenbach, 1991	✓						
BRIEF-P	Gioia, Espy, & Isquith, 2002						✓	✓

✓ = near transfer tasks (practiced during training)

✓ = intermediate transfer tasks (alternative form used, or closely related procedure)

✓ = far-intermediate transfer (not practiced during training, but assessed at baseline and post-assessments)

✓ = far transfer (informant-report measures collected at baseline, 1-month, and 3 months or cognitive measures that were only assessed at baseline)

✓ = only practiced during training

APPENDIX E

SAMPLE SCHEDULE OF TRAINING DAY

Table E1
Sample Schedule of a Training Day

Schedule	Activity
Start @9AM: Group Welcome	Warm up; review schedule; review group rules
9:20AM: Small Group (Group A: n = 3; Group B: n = 3)	Group A: Hand Game Group B: Walk-a-Line
9:35AM: Large Group (Group A & B: n = 6)	Color Matching Freeze
9:45AM: Small Group (Group A: n = 3; Group B: n = 3)	Group A: DCCS Group B: Grass/Snow Stroop
10:00AM: Large Group (Group A & B: n = 6)	Drum Beats
10:10AM: Small Group (Group A: n = 3; Group B: n = 3)	Group A: Walk-a-Line Group B: Hand Game
10:25AM: Break	Snack Time
10:45AM: Free Play	Free Play
11:05AM: Small Group (Group A: n = 3; Group B: n = 3)	Group A: Grass/Snow Stroop Group B: DCCS
11:20AM: Large Group (Group A & B: n = 6)	Red Light/Purple Light
11:30AM: Large Group (Group A & B: n = 6)	Say Goodbye Review Marble Jar
11:40AM: End Group	

APPENDIX F

DESCRIPTION OF TIME-LOCKED AND STIMULUS-LOCKED EVENT-RELATED POTENTIAL ANALYSES

ERN. The location of maximal amplitude for the ERN was evaluated using a 2 (trial type: error; correct) X 5 (site: Fz; FCz; Cz; CPz; Pz) repeated measures ANOVA model and follow-up paired sample t-tests. The response-locked waveforms from the Go/No-Go task can be seen in Figure B4 and descriptive statistics in Tables F1 and F2. Analyses indicated a greater negativity on error No-Go trials compared to correct Go trials at pre- ($F(1, 43) = 52.88$, $p < .001$, $\eta^2_p = 0.55$) and post-assessment time points ($F(1, 43) = 28.00$, $p < .001$, $\eta^2_p = 0.39$). This effect varied by site X trial type interaction, indicating a larger error vs. correct difference at frontocentral recording sites at pre- ($F(4, 172) = 16.63$, $p < .001$, $\eta^2_p = 0.28$) and post-assessment time points ($F(4, 172) = 17.71$, $p < .001$, $\eta^2_p = 0.29$). ERN amplitudes were largest at FCz at post-assessment compared to Cz ($t(43) = 3.05$, $p = .004$), CPz ($t(43) = 3.70$, $p = .001$), and Pz ($t(43) = 3.07$, $p = .004$). Therefore, further analyses focused on FCz for post-assessments. While the ERN amplitude was numerically largest at Cz at baseline, it was not statistically different from the ERN at FCz ($t(43) = 1.02$, $p = 0.50$). Therefore, further analyses focused on FCz at baseline for consistency within task and across time points.

The response-locked waveforms from the Flanker task can be seen in Figure B5. A significant site X trial type interaction, indicating a larger error vs. correct difference at frontocentral recording sites was observed at baseline ($F(4, 164) = 8.26$, $p < .001$, $\eta^2_p = 0.17$). At post-assessment, a greater negativity on error trials compared to correct trials was observed ($F(1, 39) = 8.42$, $p = 0.01$, $\eta^2_p = 0.18$); the effect varied by site X trial type interaction with a larger error vs. correct difference at frontocentral recording sites ($F(4, 156) = 4.72$, $p = .001$, $\eta^2_p = 0.11$). At baseline, the ERN was largest in magnitude at FCz compared to Pz ($t(41) = 2.05$, $p = 0.05$); however there were no significant differences in magnitude between FCz and Fz ($t(41) = 1.05$, $p = 0.30$), Cz ($t(41) = 1.10$, $p = 0.28$), or CPz ($t(41) = 1.36$, $p = 0.18$). Since the ERN was numerically largest at FCz, further analyses focused on FCz for the baseline time point. At post-assessment, the ERN was significantly larger in magnitude at FCz compared to Fz ($t(39) = 2.60$, $p = 0.01$), CPz ($t(39) = 2.82$, $p = .008$), and Pz ($t(39) = 2.22$, $p = 0.03$). Even though the ERN was numerically largest at Cz, there were no significant differences in magnitude between FCz and Cz ($t(39) = 0.25$, $p = 0.80$). Therefore, further analyses focused on FCz for the post-assessment time point for consistency across and within task analyses.

N2. The location of maximal amplitude for the N2 was evaluated using a 2 (trial type: high conflict; low conflict) X 5 (site: Fz; FCz; Cz; CPz; Pz) repeated measures ANOVA model and follow-up paired sample t-tests. The stimulus-locked waveforms from the Go/No-Go task can be seen in Figure B5 and descriptive statistics in Tables F1 and F2. At baseline, a main effect of site ($F(4, 176) = 84.37$, $p < .001$, $\eta^2_p = 0.66$) indicated greater negativity at frontocentral recording sites. However, the effect did not vary by site X trial interaction ($F(4, 176) = 1.20$, $p = 0.31$, $\eta^2_p = 0.03$). In contrast, at post-assessment there was a main effect of site ($F(4, 172) = 64.92$, $p < .001$, $\eta^2_p = 0.60$) and trial type ($F(1, 43) = 11.02$, $p = .002$, $\eta^2_p = 0.20$). Further, a significant site X trial type interaction indicated a larger negativity on No-Go trials compared to Go trials at frontocentral sites ($F(4, 172) = 6.15$, $p < .001$, $\eta^2_p = 0.13$). The N2 amplitude on No-Go trials was largest in magnitude at FCz at both pre- and post-assessment time points compared to Fz ($t(44) = 4.81$, $p <.001$; $t(43) = 3.06$, $p = .004$), CPz ($t(44) = 5.87$, $p < .001$; $t(43) = 5.43$, $p <.001$), and Pz ($t(44) = 10.65$, $p < .001$; $t(43) = 10.25$, $p <.001$). At baseline, the N2 amplitude on No-Go trials at FCz was not significantly different from the amplitude at Cz ($t(45) = 0.02$, $p = 0.99$). However, at the post-assessment time point, the N2 amplitude on No-Go trials at FCz was more negative compared to Cz ($t(43) = 2.47$, $p = 0.02$). Therefore, further analyses focused on FCz for both pre- and post-assessment time points.

The stimulus-locked waveforms from the Flanker task can be seen in Figure B5 and descriptive statistics in Tables F1 and F2. At baseline and post-assessment time points, there was a main effect of site ($F(4, 164) = 42.89$, $p < .001$, $\eta^2_p = 0.51$; $F(4, 148) = 28.30$, $p < .001$, $\eta^2_p = 0.43$) such that frontocentral sites were more negative than posterior sites. At baseline, N2 amplitudes were also more negative on high conflict trials compared to low conflict trials ($F(1, 41) = 4.10$, $p = 0.05$, $\eta^2_p = 0.09$); and at post-assessment this effect approached trend-level significance ($F(1, 37) = 3.20$, $p = 0.08$, $\eta^2_p = 0.08$). There was not a significant site X trial interaction at baseline or post-assessments ($F(4, 164) = 1.20$, $p = 0.32$, $\eta^2_p = 0.03$; $F(4, 148) = 1.60$, $p = 0.18$, $\eta^2_p = 0.04$). However, the N2 amplitude on high conflict trials was significantly more negative at FCz compared to Fz ($t(41) = 2.59$), $p = 0.01$), CPz ($t(41) = 3.80$, $p < .001$), and Pz ($t(41) = 8.09$, $p < .001$). A trend-level effect of a more negative N2 amplitude on high conflict trials at FCz compared to Cz was also observed ($t(41) = 1.96$, $p = 0.06$). At post-assessment, a significantly more negative N2 amplitude on high conflict trials was observed at FCz compared to CPz ($t(39) = 2.85$, $p = .007$) and Pz ($t(38) = 6.79$, $p < .001$). Since the N2 amplitude on high conflict trials was numerically largest at FCz, further analyses focused on FCz for both pre- and post-assessment time points.

Table F1

Pre-assessment Means (*SD*) of Time-Locked and Stimulus-Locked Event-Related Potential Voltage Amplitudes (µV) for Go/No-Go and Flanker Tasks Across Five Midline Sites

	Fz	FCz	Cz	CPz	Pz
Go/No-Go					
ERN	-1.32 (5.09)	-2.27 (5.87)	-2.61 (5.33)	-1.53 (4.01)	-0.82 (4.20)
CRN	3.41 (2.20)	4.00 (2.57)	3.68 (2.83)	2.27 (2.50)	1.09 (2.40)
ΔERN	-4.73 (5.14)	-6.27 (6.02)	-6.29 (5.33)	-3.80 (3.56)	-1.90 (4.20)
High Conflict N2	-19.2 (6.23)	-20.5 (5.95)	-20.0 (5.27)	-16.1 (5.60)	-9.73 (4.92)
Low Conflict N2	-19.8 (4.76)	-20.8 (4.41)	-20.2 (4.37)	-16.5 (5.22)	-10.8 (4.51)
ΔN2	0.59 (3.46)	0.31 (3.45)	0.21 (3.02)	0.34 (3.11)	1.08 (2.94)
Fish Flanker					
ERN	-0.65 (3.15)	-0.84 (3.39)	-0.52 (3.40)	-0.15 (2.90)	0.26 (2.72)
CRN	0.82 (1.97)	0.64 (2.35)	0.38 (2.54)	0.05 (2.43)	-0.45 (2.10)
ΔERN	-1.46 (3.11)	-1.48 (3.46)	-0.90 (3.41)	-0.20 (3.48)	0.70 (3.18)
High Conflict N2	-1.02 (3.39)	-1.52 (3.84)	-1.02 (4.21)	-0.18 (4.07)	2.08 (3.46)
Low Conflict N2	-0.51 (3.68)	-0.62 (4.03)	0.06 (4.54)	1.18 (4.31)	3.31 (4.16)
ΔN2	-0.52 (3.45)	-0.90 (3.70)	-1.08 (3.65)	-1.37 (3.45)	-1.23 (3.34)

Note. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN.

Table F2

Post-assessment Means (*SD*) of Time-Locked and Stimulus-Locked Event-Related Potential Voltage Amplitudes (µV) for Go/No-Go and Flanker Tasks Across Five Midline Sites

	Fz	FCz	Cz	CPz	Pz
Go/No-Go					
ERN	-2.80 (4.95)	-3.08 (5.27)	-2.00 (5.09)	-0.74 (4.69)	-0.80 (4.78)
CRN	2.25 (2.65)	2.79 (2.94)	2.79 (2.89)	1.94 (2.37)	0.83 (2.19)
ΔERN	-5.04 (5.75)	-5.87 (6.24)	-4.79 (5.89)	-2.68 (5.06)	-1.63 (5.00)
High Conflict N2	-21.4 (6.91)	-22.3 (6.70)	-21.0 (6.01)	-16.8 (7.01)	-11.1 (6.00)
Low Conflict N2	-18.8 (4.91)	-19.7 (4.78)	-18.9 (4.66)	-15.5 (5.57)	-10.9 (4.83)
ΔN2	-2.58 (5.10)	-2.64 (4.90)	-2.05 (3.83)	-1.21 (3.64)	-0.25 (3.47)
Fish Flanker					
ERN	-0.99 (3.04)	-1.89 (3.57)	-1.81 (3.93)	-0.55 (3.15)	-0.30 (4.51)
CRN	0.78 (1.95)	0.56 (2.38)	0.59 (2.92)	0.59 (2.74)	0.28 (2.70)
ΔERN	-1.77 (3.15)	-2.45 (3.93)	-2.40 (4.55)	-1.14 (4.02)	-0.58 (5.09)
High Conflict N2	-0.54 (2.80)	-0.69 (3.29)	-0.29 (3.53)	0.60 (4.03)	2.34 (3.48)
Low Conflict N2	-0.44 (2.79)	-0.20 (2.74)	0.61 (3.41)	1.46 (3.92)	2.96 (3.63)
ΔN2	-0.10 (2.73)	-0.48 (2.50)	-0.90 (2.74)	-0.86 (3.44)	-0.62 (3.14)

Note. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN.

APPENDIX G

CORRELATED CHANGE ASSOCIATIONS BETWEEN EVENT-RELATED POTENTIALS AND BEHAVIORAL PERFORMANCE

Table G1

Associations Between Changes in Go/No-Go Event-Related Potentials and Changes in Go/No-Go Behavioral Performance

Go/No-Go	1	2	3	4	5	6	7	8	9
1. ERN	--								
2. ΔERN	0.96^{**}	--							
3. HC N2	0.16	0.11	--						
4. ΔN2	0.11	0.07	0.85^{**}	--					
5. Total accuracy (%)	0.10	0.19	$-0.26^{†}$	-0.22	--				
6. RT error (ms)	0.30^{*}	0.18	$0.27^{†}$	0.10	0.05	--			
7. RT correct (ms)	$0.29^{†}$	0.29^{*}	0.10	0.11	-0.01	0.22	--		
8. Post error accuracy (%)	0.25	0.21	-0.13	-0.09	0.60^{**}	$0.25^{†}$	-0.09	--	
9. RT post error (ms)	0.21	$0.26^{†}$	0.11	0.08	0.13	0.30^{*}	0.60^{**}	-0.16	--
10. RT post correct (ms)	$0.27^{†}$	0.06	0.04	-0.05	0.30^{*}	0.31^{*}	0.69^{**}	0.04	0.50^{**}

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{†}p < 0.10$. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN. RT = Reaction Time.

Table G2

Associations Between Changes in Flanker Event-Related Potentials and Changes in Flanker Behavioral Performance

Flanker	1	2	3	4	5	6	7	8	9
1. ERN	--								
2. ΔERN	0.91**	--							
3. HC N2	0.01	0.06	--						
4. ΔN2	0.19	0.19	0.72**	--					
5. Total accuracy (%)	-0.26	-0.30†	-0.04	-0.05	--				
6. RT error (ms)	0.35*	0.38*	-0.03	0.01	-0.33*	--			
7. RT correct (ms)	0.27†	0.18	0.21	0.23	0.02	0.34*	--		
8. Post error accuracy (%)	-0.23	-0.12	0.00	-0.08	0.54**	-0.19	-0.21	--	
9. RT post error (ms)	0.04	0.06	0.01	-0.11	0.03	-0.09	0.45**	0.09	--
10. RT post correct (ms)	0.26	0.20	0.02	0.09	0.30*	0.13	0.61**	0.01	0.32*

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN. RT = Reaction Time.

APPENDIX H

SCATTERPLOTS OF SIGNIFICANT AND TREND-LEVEL CORRELATED CHANGE ASSOCIATIONS BETWEEN METHOD

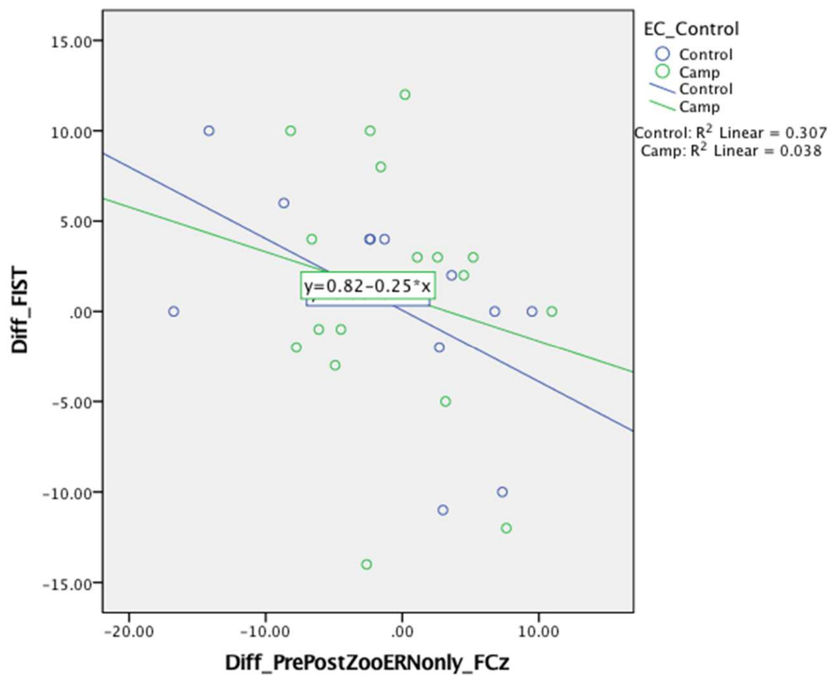


Figure H1. Increase in Go/No-Go ERN over time is associated with greater improvements in total FIST score.

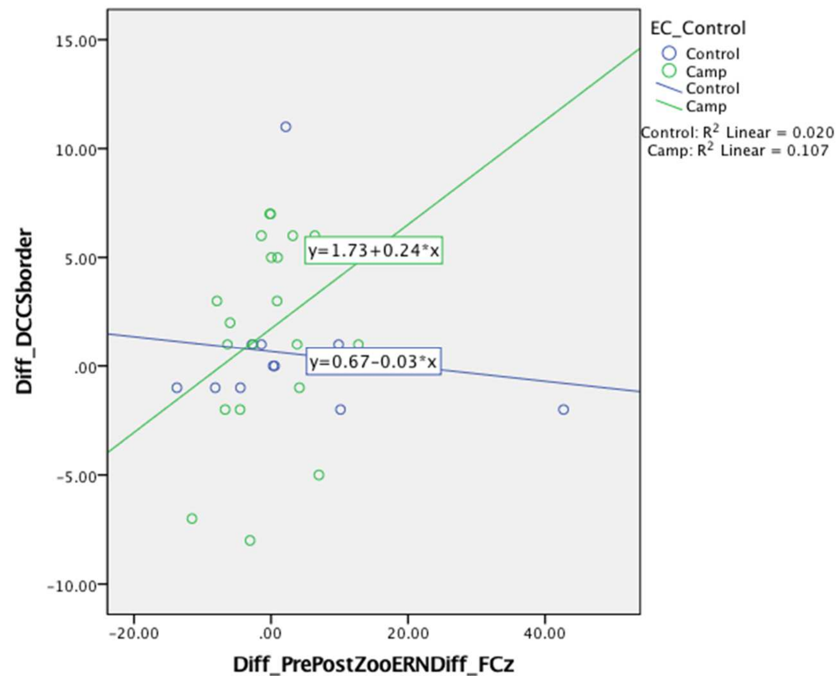


Figure H2. Reduction in Go/No-Go ERN over time is associated with greater improvements in DCCS border score.

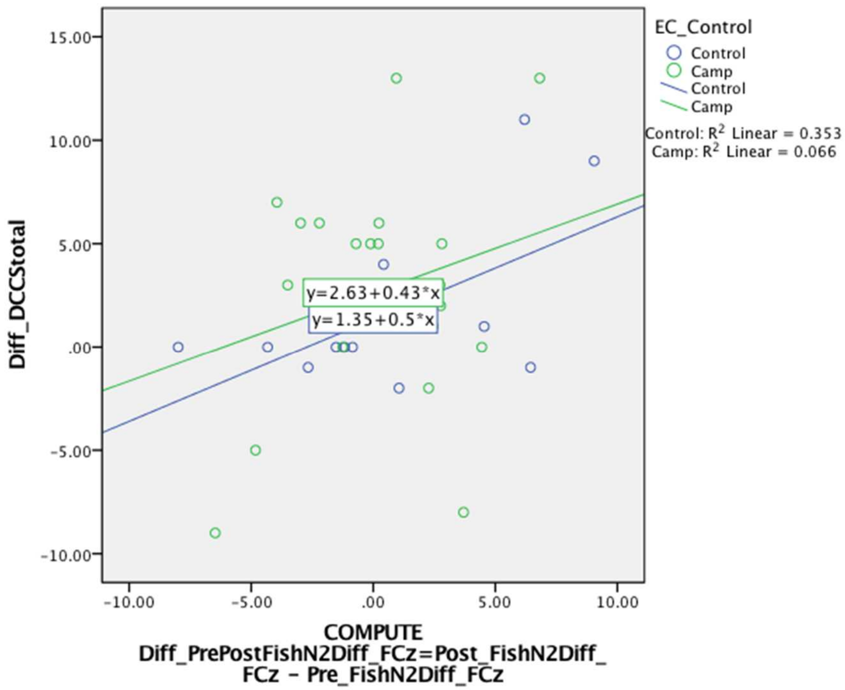


Figure H3. Reduction in Flanker ΔN2 over time is associated with greater improvements in DCCS border score.

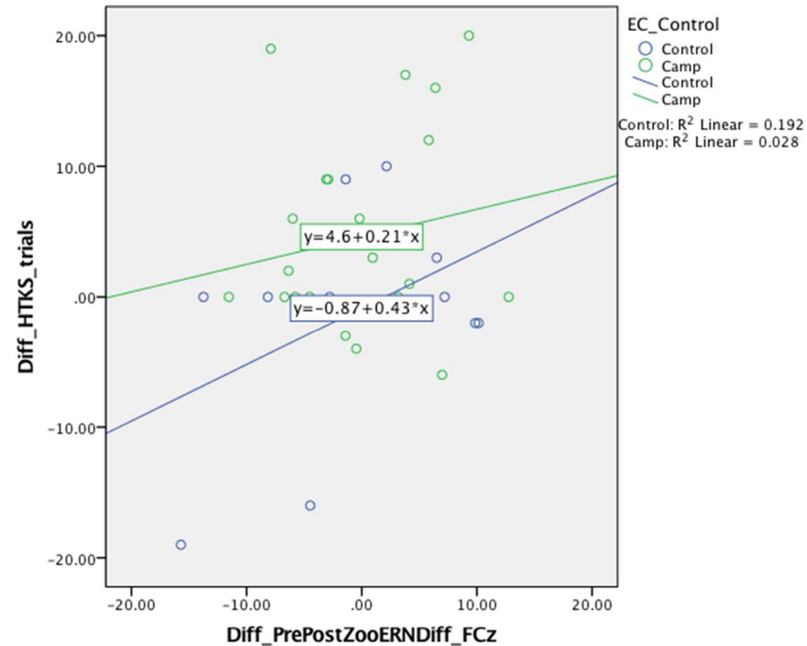


Figure H4. Reduction in Go/No-Go ΔERN over time is associated with greater improvements in HTKS task.

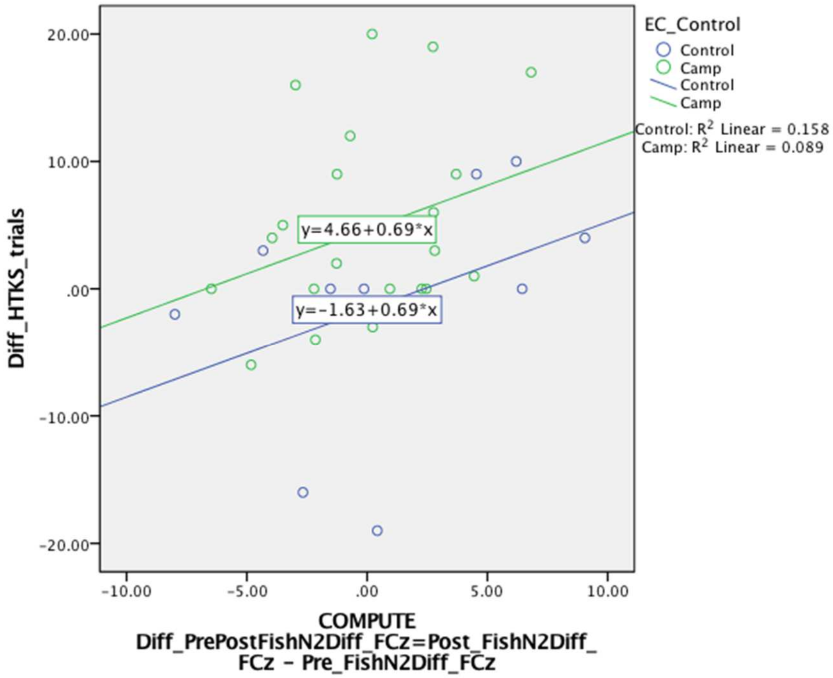


Figure H5. Reduction in Flanker ΔN2 over time is associated with greater improvements in HTKS task.

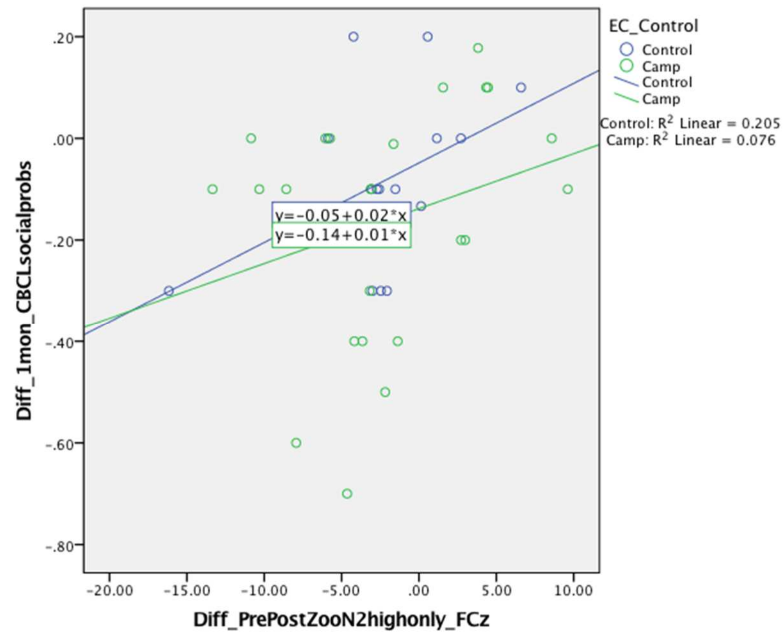


Figure H6. Increases in Go/No-Go N2 over time is associated with fewer parent-reported social problems.

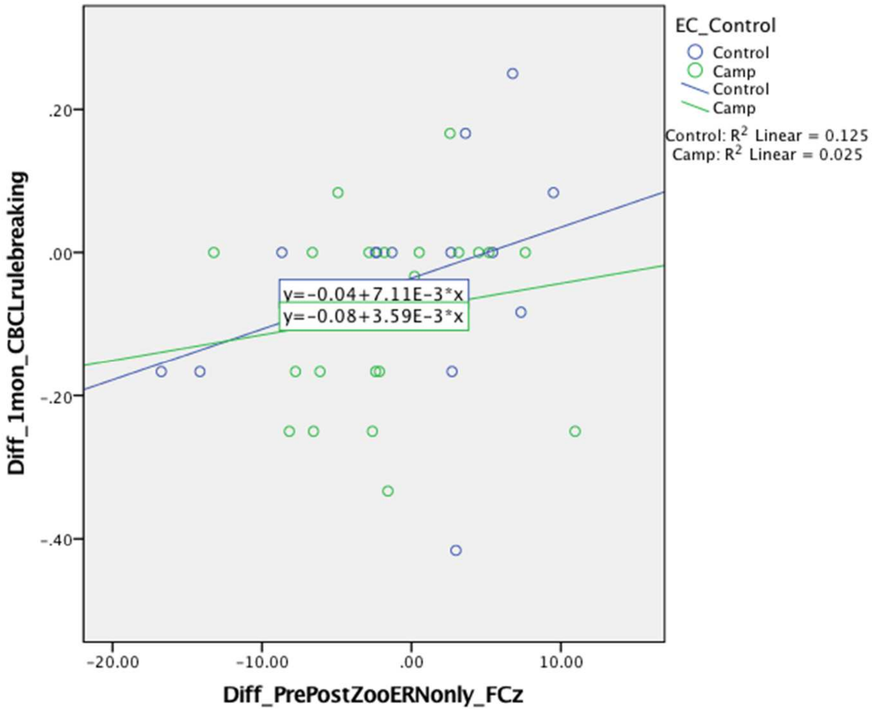


Figure H7. Increases in the Go/No-Go ERN over time is associated with fewer parent-reported problems with rule breaking.

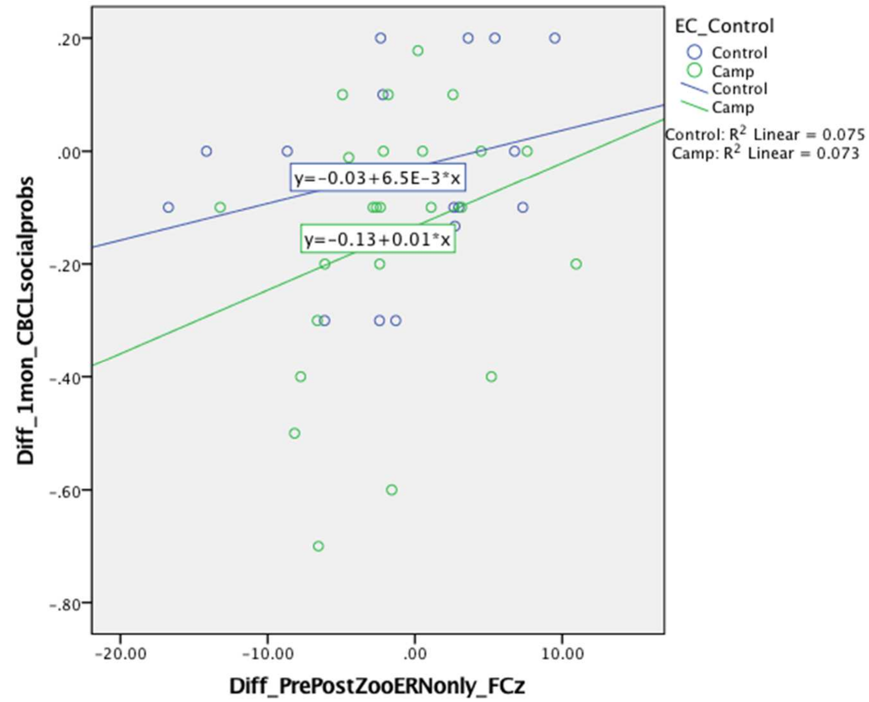


Figure H8. Increases in the Go/No-Go ERN over time is associated with fewer parent-reported social problems.

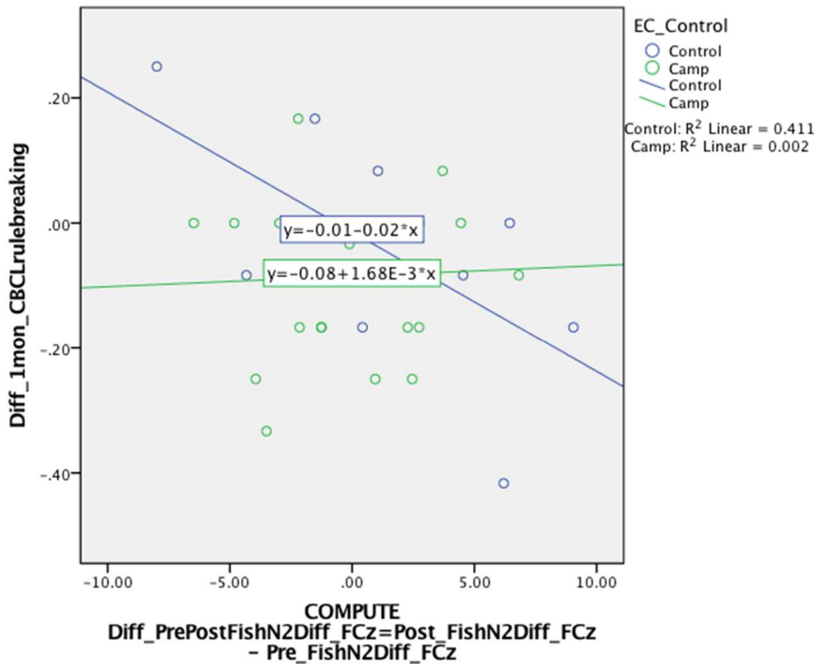


Figure H9. Reduction in Flanker ΔN2 is associated with fewer parent-reported problems with rule breaking.

APPENDIX I

BIVARIATE CORRELATIONS BETWEEN METHOD WITHIN TIME POINT

Table I1
Baseline Associations Between Go/No-Go Event-Related Potentials and Go/No-Go Behavioral Performance

Go/No-Go	1	2	3	4	5	6	7	8	9
1. ERN	--								
2. ΔERN	0.92**	--							
3. HC N2	0.35*	0.45**	--						
4. ΔN2	0.44**	0.47**	0.66**	--					
5. Total accuracy (%)	-0.16	-0.29†	-0.35*	-0.16	--				
6. RT error (ms)	0.56**	0.62**	0.16	0.07	-0.41**	--			
7. RT correct (ms)	0.43**	0.38*	0.19	0.21	-0.14	0.64**	--		
8. Post error accuracy (%)	0.14	0.14	-0.14	0.00	0.54**	-0.10	-0.21	--	
9. RT post error (ms)	0.50**	0.48**	0.16	0.14	0.09	0.63**	0.59**	0.08	--
10. RT post correct (ms)	0.29†	0.11	-0.09	0.05	0.41**	0.29†	0.69**	0.14	0.57**

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{†}p < 0.10$. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN. RT = Reaction Time.

Table I2
Post-Assessment Associations Between Go/No-Go Event-Related Potentials and Go/No-Go Behavioral Performance

Go/No-Go	1	2	3	4	5	6	7	8	9
1. ERN	--								
2. ΔERN	0.91**	--							
3. HC N2	0.39*	0.40**	--						
4. ΔN2	0.33*	0.27†	0.65**	--					
5. Total accuracy (%)	-0.33*	-0.39*	-0.49**	-0.29†	--				
6. RT error (ms)	0.58**	0.52**	0.52**	0.42**	-0.45**	--			
7. RT correct (ms)	0.48**	0.36*	0.13	0.20	-0.25	0.70**	--		
8. Post error accuracy (%)	-0.07	-0.07	-0.16	-0.08	0.69**	-0.20	-0.36*	--	
9. RT post error (ms)	0.67**	0.62**	0.49**	0.29	-0.26†	0.70**	0.47**	0.06	--
10. RT post correct (ms)	0.38*	0.22	0.00	0.10	0.26†	0.49**	0.78**	0.00	0.47**

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{†}p < 0.10$. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN. RT = Reaction Time.

Table I3
Baseline Associations Between Flanker Event-Related Potentials and Flanker Behavioral Performance

Flanker	1	2	3	4	5	6	7	8	9
1. ERN	--								
2. ΔERN	0.75**	--							
3. HC N2	0.25	0.22	--						
4. ΔN2	0.19	0.24	0.50**	--					
5. Total accuracy (%)	-0.39*	-0.39*	-0.13	0.22	--				
6. RT error (ms)	-0.19	0.15	-0.23	-0.13	-0.16	--			
7. RT correct (ms)	-0.19	0.10	-0.28†	-0.18	0.19	0.61**	--		
8. Post error accuracy (%)	-0.55**	-0.47**	-0.18	-0.08	0.76**	-0.01	0.17	--	
9. RT post error (ms)	-0.17	0.00	-0.19	0.03	0.60**	0.38*	0.57**	0.48**	--
10. RT post correct (ms)	-0.20	-0.07	-0.30†	-0.06	0.54**	0.43**	0.78**	0.47**	0.72**

Note. ***p < .001. **p < 0.01. *p < 0.05. †p < 0.10. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN. RT = Reaction Time.

Table I4
Post-Assessment Associations Between Flanker Event-Related Potentials and Flanker Behavioral Performance

Flanker	1	2	3	4	5	6	7	8	9
1. ERN	--								
2. ΔERN	0.80**	--							
3. HC N2	-0.02	0.07	--						
4. ΔN2	-0.08	0.10	0.59**	--					
5. Total accuracy (%)	-0.42**	-0.32*	-0.27	-0.13	--				
6. RT error (ms)	0.20	0.50**	0.00	0.21	-0.14	--			
7. RT correct (ms)	0.19	0.37*	-0.11	0.16	-0.06	0.70**	--		
8. Post error accuracy (%)	-0.32*	-0.11	-0.18	-0.13	0.79**	0.03	0.00	--	
9. RT post error (ms)	-0.25	-0.03	-0.20	0.01	0.53**	0.29	0.42**	0.71**	--
10. RT post correct (ms)	-0.12	0.12	-0.12	0.23	0.51**	0.45**	0.71**	0.51**	0.70**

Note. ***p < .001. **p < 0.01. *p < 0.05. †p < 0.10. ERN = Error-related Negativity. CRN = Correct-response Negativity, the voltage amplitude on correct trials identified in the same time window as the ERN. ΔERN = Difference between the ERN and CRN. RT = Reaction Time.

Table I5
Baseline Associations Between Neurophysiological Measures of Effortful Control and Behavioral Measures of Attentional Control

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.91^{**}	--						
3. HC N2	0.35^{*}	0.45^{**}	--					
4. ΔN2	0.44^{**}	0.47^{**}	0.68^{**}	--				
Fish Flanker								
5. ERN	0.13	0.05	0.14	-0.11	--			
6. ΔERN	-0.10	-0.06	0.13	-0.07	0.77^{**}	--		
7. HC N2	0.00	0.18	0.33^{*}	0.02	0.25	0.22	--	
8. ΔN2	-0.13	-0.05	-0.13	0.04	0.19	0.24	0.43^{**}	--
9. BDS	0.06	0.07	$0.27^{\dagger}$	0.20	-0.08	0.02	-0.10	0.15
10. RC	0.11	0.17	-0.01	0.04	-0.02	-0.04	-0.10	-0.17
11. DCCS Border	-0.02	-0.02	-0.11	-0.09	0.07	0.29	0.08	0.08
12. DCCS Total	-0.06	-0.05	0.03	0.09	-0.15	-0.01	-0.03	0.23
13. Day/Night	-0.03	-0.14	-0.06	0.10	0.06	-0.04	-0.22	$-0.31^{\dagger}$
14. Shape Stroop	-0.10	-0.09	0.10	0.17	0.05	0.34^{*}	-0.04	-0.15
15. FIST Select 2	-0.21	-0.18	0.08	0.12	-0.14	0.06	0.01	0.06
16. FIST Total	-0.19	-0.16	0.13	0.18	-0.19	0.04	-0.08	0.02
17. Tower of Hanoi	0.26	0.18	0.18	0.34^{*}	-0.03	-0.18	-0.01	0.08
18. Yes/No	-0.09	-0.13	0.02	-0.01	0.11	0.17	0.07	0.02

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. BDS = Backward Digit Span. RC = Reverse Categorization. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Shift Task.

Table I6

Post-Assessment Associations Between Neurophysiological Measures of Effortful Control and Behavioral Measures of Attentional Control

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.88**	--						
3. HC N2	0.38*	0.40*	--					
4. ΔN2	0.33*	0.27†	0.70**	--				
Fish Flanker								
5. ERN	0.35*	0.25	0.17	0.34*	--			
6. ΔERN	0.35*	0.37*	0.13	0.28†	0.80**	--		
7. HC N2	0.01	-0.02	0.28†	0.05	-0.02	0.07	--	
8. ΔN2	0.07	0.06	-0.08	-0.17	-0.08	0.10	0.58**	--
9. BDS	0.09	0.03	0.07	-0.02	0.18	0.19	-0.27†	-0.28†
10. RC	0.22	0.07	0.02	0.17	0.00	-0.09	0.00	-0.09
11. DCCS Border	0.09	0.11	0.05	-0.20	0.15	0.29	-0.16	-0.08
12. DCCS Total	0.09	0.06	-0.11	-0.16	0.13	0.30	-0.29†	0.01
13. Day/Night	0.01	-0.20	-0.13	-0.06	0.13	0.08	-0.17	-0.23
14. Shape Stroop	0.10	0.11	-0.10	-0.11	-0.24	-0.07	-0.08	0.17
15. FIST Select 2	-0.07	-0.02	0.01	-0.10	0.07	0.04	0.03	-0.16
16. FIST Total	-0.24	-0.38*	-0.39*	-0.19	0.04	0.00	-0.24	-0.24
17. Tower of Hanoi	0.00	0.10	0.08	0.04	-0.05	-0.01	-0.19	0.07
18. Yes/No	-0.20	-0.17	-0.12	0.06	-0.09	-0.14	0.16	-0.04

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. BDS = Backward Digit Span. RC = Reverse Categorization. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Shift Task.

Table I7

Baseline Associations Between Neurophysiological Measures of Effortful Control and Behavioral Measures of Inhibitory Control

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.91**	--						
3. HC N2	0.35*	0.45**	--					
4. ΔN2	0.44**	0.47**	0.68**	--				
Fish Flanker								
5. ERN	0.13	0.05	0.14	-0.11	--			
6. ΔERN	-0.10	-0.06	0.13	-0.07	0.77**	--		
7. HC N2	0.00	0.18	0.33*	0.02	0.25	0.22	--	
8. ΔN2	-0.13	-0.05	-0.13	0.04	0.19	0.24	0.43**	--
9. Walk-a-line	-0.11	-0.12	-0.14	-0.08	0.02	-0.03	0.11	0.21
10. Simon Says	-0.30	-0.36*	-0.31*	-0.12	-0.11	-0.12	-0.40*	-0.08
11. Hand Game	0.09	-0.07	-0.16	0.08	0.05	0.02	-0.20	-0.14
12. *Bear Dragon*	-0.36*	-0.47**	-0.17	-0.11	0.09	0.17	-0.04	0.02
13. *Knock-Tap*	-0.23	-0.23	-0.11	-0.06	-0.11	-0.28	-0.12	0.26
14. *Snack Delay*	-0.02	-0.08	-0.10	0.06	0.13	0.12	-0.08	0.27
15. HTKS trials	-0.36*	-0.45**	0.00	0.04	-0.23	-0.11	-0.12	0.02
16. HTKS total	-0.40*	-0.44**	-0.19	-0.11	-0.12	0.00	-0.10	0.06
17. Whisper Game	-0.10	0.02	0.08	0.08	-0.18	-0.11	-0.32	0.02

Note. ***p < .001. **p < 0.01. *p < 0.05. †p < 0.10. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. HTKS = Head-Toes-Knees-Shoulders task.

Table I8
Post-Assessment Baseline Associations Between Neurophysiological Measures of Effortful Control and Behavioral Measures of Inhibitory Control

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.88**	--						
3. HC N2	0.38*	0.40*	--					
4. ΔN2	0.33*	0.27†	0.70**	--				
Fish Flanker								
5. ERN	0.35*	0.25	0.17	0.34*	--			
6. ΔERN	0.35*	0.37*	0.13	0.28†	0.80**	--		
7. HC N2	0.01	-0.02	0.28†	0.05	-0.02	0.07	--	
8. ΔN2	0.07	0.06	-0.08	-0.17	-0.08	0.10	0.58**	--
9. Walk-a-line	-0.22	-0.19	-0.21	-0.21	-0.10	0.01	0.06	0.20
10. Simon Says	-0.15	-0.19	-0.11	-0.05	-0.06	-0.17	-0.12	-0.21
11. Hand Game	0.04	-0.14	-0.09	-0.22	0.08	0.06	-0.19	0.16
12. *Bear Dragon*	0.06	-0.01	-0.18	-0.08	0.16	0.16	-0.24	-0.48**
13. *Knock-Tap*	-0.11	-0.09	-0.37*	-0.27†	-0.23	-0.08	-0.37*	0.06
14. *Snack Delay*	-0.02	-0.04	-0.10	0.08	0.20	0.21	0.05	-0.15
15. HTKS trials	-0.26	-0.16	-0.14	-0.27†	-0.21	-0.06	-0.20	0.21
16. HTKS total	-0.16	-0.09	-0.17	-0.27†	-0.19	-0.05	-0.28†	0.06
17. Whisper Game	0.17	0.24	0.18	-0.01	-0.07	-0.05	-0.08	0.04

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. HTKS = Head-Toes-Knees-Shoulders task.

Table I9

Baseline Associations Between Neurophysiological Measures of Effortful Control and Parent-Reported Behavior Rating Inventory of Executive Function

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.91^{**}	--						
3. HC N2	0.35^{*}	0.45^{**}	--					
4. ΔN2	0.44^{**}	0.47^{**}	0.68^{**}	--				
Fish Flanker								
5. ERN	0.13	0.05	0.14	-0.11	--			
6. ΔERN	-0.10	-0.06	0.13	-0.07	0.77^{**}	--		
7. HC N2	0.00	0.18	0.33^{*}	0.02	0.25	0.22	--	
8. ΔN2	-0.13	-0.05	-0.13	0.04	0.19	0.24	0.43^{**}	--
9. BRIEF Inhibit	$-0.32^{\dagger}$	-0.36^{*}	-0.13	0.06	-0.27	-0.29	-0.13	0.11
10. BRIEF Shift	$-0.32^{\dagger}$	$-0.30^{\dagger}$	-0.11	-0.10	0.07	0.15	0.09	0.26
11. BRIEF EmoC	-0.37^{*}	-0.44^{**}	-0.07	-0.07	0.14	0.14	-0.06	0.22
12. BRIEF WM	-0.21	-0.14	-0.15	-0.05	-0.27	-0.32	-0.03	0.12
13. BRIEF PlanO	-0.15	-0.26	-0.20	-0.07	-0.22	-0.27	-0.25	0.00
14. BRIEF GEC	-0.41^{*}	-0.44^{**}	-0.17	-0.05	-0.17	-0.18	-0.11	0.21

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Function, Preschool Version. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organization. GEC = Global Executive Composite.

Table I10

One-Month Follow-Up Associations Between Neurophysiological Measures of Effortful Control and Parent-Reported Behavior Rating Inventory of Executive Function

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.88**	--						
3. HC N2	0.38*	0.40*	--					
4. ΔN2	0.33*	0.27†	0.70**	--				
Fish Flanker								
5. ERN	0.35*	0.25	0.17	0.34*	--			
6. ΔERN	0.35*	0.37*	0.13	0.28†	0.80**	--		
7. HC N2	0.01	-0.02	0.28†	0.05	-0.02	0.07	--	
8. ΔN2	0.07	0.06	-0.08	-0.17	-0.08	0.10	0.58**	--
9. BRIEF Inhibit	-0.05	0.04	-0.13	0.06	0.20	0.11	-0.35†	-0.19
10. BRIEF Shift	-0.07	-0.04	-0.11	-0.10	-0.12	-0.08	0.10	0.03
11. BRIEF EmoC	0.09	0.01	-0.07	-0.07	0.01	-0.20	-0.24	-0.21
12. BRIEF WM	-0.19	0.01	-0.15	-0.05	0.13	0.12	-0.22	-0.14
13. BRIEF PlanO	-0.18	-0.04	-0.20	-0.07	0.04	-0.05	-0.33†	-0.15
14. BRIEF GEC	-0.09	0.00	-0.17	-0.05	0.09	-0.02	-0.29†	-0.19

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{†}p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Function, Preschool Version. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organization. GEC = Global Executive Composite.

Table I11
Baseline Associations Between Neurophysiological Measures of Effortful Control and Parent-Reported Child Behavior Checklist

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.91**	--						
3. HC N2	0.35*	0.45**	--					
4. ΔN2	0.44**	0.47**	0.68**	--				
Fish Flanker								
5. ERN	0.13	0.05	0.14	-0.11	--			
6. ΔERN	-0.10	-0.06	0.13	-0.07	0.77**	--		
7. HC N2	0.00	0.18	0.33*	0.02	0.25	0.22	--	
8. ΔN2	-0.13	-0.05	-0.13	0.04	0.19	0.24	0.43**	--
9. CBCL Anxdep	-0.18	-0.24	-0.26	-0.31*	0.31*	0.29	0.31*	0.29
10. CBCL Withdep	-0.16	-0.22	-0.44**	-0.37*	0.39*	0.35*	0.39*	0.35*
11. CBCL Somatic	0.10	0.02	-0.27	-0.11	0.20	0.11	0.20	0.11
12. CBCL Soc P	0.13	0.08	-0.07	0.03	0.20	0.16	0.20	0.16
13. CBCL Attn	0.03	0.05	-0.01	-0.07	0.07	0.01	0.07	0.01
14. CBCL Ruleb	-0.04	-0.02	-0.09	-0.14	0.15	0.15	0.15	0.15
15. CBCL Agg	-0.09	-0.14	-0.11	-0.09	0.12	0.06	0.12	0.06
16. CBCL Int	-0.11	-0.19	-0.38*	-0.32*	0.36*	0.31	0.36*	0.31
17. CBCL Ext	-0.08	-0.11	-0.11	-0.11	0.14	0.09	0.14	0.09

Note. $^{***}p$ < .001. $^{**}p$ < 0.01. $^{*}p$ < 0.05. $^{†}p$ < 0.10. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale.

Table I12
One-Month Follow-Up Associations Between Neurophysiological Measures of Effortful Control and Parent-Reported Child Behavior Checklist

Go/No-Go	1	2	3	4	5	6	7	8
1. ERN	--							
2. ΔERN	0.88**	--						
3. HC N2	0.38*	0.40*	--					
4. ΔN2	0.33*	0.27†	0.70**	--				
Fish Flanker								
5. ERN	0.35*	0.25	0.17	0.34*	--			
6. ΔERN	0.35*	0.37*	0.13	0.28†	0.80**	--		
7. HC N2	0.01	-0.02	0.28†	0.05	-0.02	0.07	--	
8. ΔN2	0.07	0.06	-0.08	-0.17	-0.08	0.10	0.58**	--
9. CBCL Anxdep	-0.15	0.00	-0.23	-0.04	0.01	-0.03	-0.25	-0.08
10. CBCL Withdep	-0.10	-0.08	-0.27	0.02	0.18	0.07	-0.09	-0.10
11. CBCL Somatic	0.00	0.02	-0.08	0.20	0.15	0.13	-0.29	-0.04
12. CBCL Soc P	0.13	0.09	0.06	0.24	0.09	-0.06	-0.29	0.00
13. CBCL Attn	0.03	0.16	0.00	0.14	0.36*	0.33	-0.22	-0.18
14. CBCL Ruleb	0.13	0.20	-0.13	0.14	0.16	0.18	-0.44**	-0.26
15. CBCL Agg	0.16	0.14	-0.13	0.10	0.04	-0.02	-0.39*	-0.17
16. CBCL Int	-0.13	-0.02	-0.24	0.04	0.11	0.04	-0.26	-0.10
17. CBCL Ext	0.16	0.18	-0.14	0.12	0.09	0.07	-0.43**	-0.22

Note. ***$p < .001$. **$p < 0.01$. *$p < 0.05$. †$p < 0.10$. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale.

Table I13

Baseline Associations Between Behavioral Measures of Attentional Control and Inhibitory Control

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. BDS	--																	
2. RC	0.07	--																
3. DCCS Border	-0.01	-0.15	--															
4. DCCS Total	0.29	-0.22	0.83**	--														
5. Day/Night	0.29	0.03	0.14	0.12	--													
6. Shape Stroop	0.17	0.00	0.08	0.04	0.28	--												
7. FIST Select 2	0.48**	-0.10	0.26	0.44**	0.20	0.28	--											
8. FIST Total	0.43**	-0.09	0.21	0.41*	0.13	0.29	0.95**	--										
9. Tower of Hanoi	0.36*	-0.03	0.03	0.27	0.25	0.08	0.12	0.14	--									
10. Yes/No	0.21	-0.01	0.16	-0.02	0.31	0.36*	0.36*	0.32	-0.04	--								
11. Walk-a-line	0.33*	0.19	0.19	0.09	0.04	-0.18	0.11	0.06	0.10	0.00	--							
12. Simon Says	0.30	0.05	-0.06	0.11	0.35*	-0.10	0.35*	0.31	0.16	0.02	0.32*	--						
13. Hand Game	-0.14	0.08	0.10	0.07	0.14	-0.05	0.02	-0.02	0.02	-0.03	-0.40*	-0.15	--					
14. Bear Dragon	-0.10	-0.11	-0.06	-0.03	0.20	-0.07	0.12	0.17	-0.08	-0.06	0.04	0.13	0.11	--				
15. Knock-Tap	0.14	-0.14	-0.12	0.35*	0.17	-0.15	0.11	0.07	0.16	-0.11	0.08	0.24	0.07	0.06	--			
*16. Snack Delay**	0.06	-0.25	0.14	0.14	-0.11	0.03	0.30	0.30	0.16	0.06	-0.17	0.18	0.09	0.07	0.09	--		
17. HTKS trials	0.29	0.19	0.17	0.33*	0.18	0.13	0.40*	0.36*	0.25	0.22	0.21	0.35*	0.23	0.19	0.28	0.05	--	
18. HTKS total	0.12	0.08	0.04	0.20	0.13	0.23	0.40*	0.41*	0.19	0.11	0.24	0.43**	0.09	0.33*	0.46**	0.20	0.78**	--
19. Whisper Game	0.44**	0.28	-0.34	-0.23	-0.15	0.23	0.09	0.20	-0.01	0.00	0.25	0.10	-0.11	-0.09	0.06	-0.16	0.16	0.19

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. First 10 tasks listed are behavioral tasks designed to elicit attentional control skills. Variables 11-19 represent behavioral tasks designed to elicit inhibitory control skills. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. Bolded text represents far-intermediate transfer tasks. *Snack delay scores are reverse coded so that higher scores mean better effortful control skills. BDS = Backward Digit Span. RC = Reverse Categorization. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Selection Task. HTKS = Head-Toes-Knees-Shoulders task.

Table I14

Post-Assessment Associations Between Behavioral Measures of Attentional Control and Inhibitory Control

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. BDS	--																	
2. RC	-0.01	--																
3. DCCS Border	0.22	-0.22	--															
4. DCCS Total	0.20	-0.20	0.98**	--														
5. Day/Night	0.16	0.20	0.04	0.09	--													
6. Shape Stroop	0.28	-0.10	-0.10	-0.12	0.05	--												
7. FIST Select 2	0.13	-0.13	0.04	0.08	-0.35*	-0.45**	--											
8. FIST Total	0.22	-0.06	-0.23	-0.18	0.36*	0.19	0.21	--										
9. Tower of Hanoi	0.12	0.35*	0.07	0.09	-0.01	0.04	0.04	0.10	--									
10. Yes/No	-0.11	-0.02	-0.27	-0.24	0.18	0.30	0.06	0.22	-0.07	--								
11. Walk-a-line	0.00	-0.19	-0.01	-0.01	0.17	0.29	0.21	0.33	-0.06	0.56**	--							
12. Simon Says	0.35*	0.09	0.18	0.15	0.17	0.04	0.00	-0.01	0.12	0.11	-0.05	--						
13. Hand Game	-0.09	0.25	0.22	0.21	0.21	-0.19	-0.13	-0.07	0.27	-0.23	-0.05	-0.17	--					
14. Bear Dragon	0.04	-0.06	0.20	0.19	0.14	-0.10	-0.04	-0.03	-0.20	-0.09	0.11	0.17	0.18	--				
15. Knock-Tap	-0.13	-0.03	0.07	0.05	-0.10	0.35*	-0.44**	-0.01	0.25	-0.16	0.05	-0.11	0.31	0.23	--			
*16. Snack Delay**	0.08	-0.04	-0.14	-0.15	-0.23	0.13	0.21	0.24	-0.10	0.15	0.07	0.14	-0.28	0.07	-0.16	--		
17. HTKS trials	0.10	0.03	0.38*	0.36*	0.21	0.07	-0.04	-0.17	0.17	-0.02	0.37*	0.35*	0.29	0.19	0.23	-0.37*	--	
18. HTKS total	0.12	0.07	0.33*	0.32*	0.15	0.12	0.00	-0.17	0.12	-0.05	0.25	0.33*	0.32	0.34*	0.31	-0.20	0.91**	--
19. Whisper Game	0.03	0.00	0.14	0.13	-0.09	0.40*	-0.06	-0.17	-0.14	0.12	-0.20	0.06	-0.08	-0.06	-0.06	0.03	0.12	0.34*

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. First 10 tasks listed are behavioral tasks designed to elicit attentional control skills. Variables 11-19 represent behavioral tasks designed to elicit inhibitory control skills. Text styles of behavioral tasks denote type of transfer level. Regular text represents near transfer tasks. *Italicized text* represents intermediate transfer tasks. **Bolded text** represents far-intermediate transfer tasks. *Snack delay scores are reverse coded so that higher scores mean better effortful control skills. BDS = Backward Digit Span. RC = Reverse Categorization. DCCS = Dimensional Card Change Sort. FIST = Flexible Item Selection Task. HTKS = Head-Toes-Knees-Shoulders task.

Table I15

Baseline Associations Between Parent-Reported Behavior Rating Inventory of Executive Functions and Child Behavior Checklist Scores

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. BRIEF Inhibit	--													
2. BRIEF Shift	0.01	--												
3. BRIEF EmoC	0.47**	0.51**	--											
4. BRIEF WM	0.62**	0.04	0.23	--										
5. BRIEF PlanO	0.46**	0.11	0.29†	0.64**	--									
6. BRIEF GEC	0.82**	0.44**	0.73**	0.74**	0.68**	--								
7. CBCL Anxdep	0.10	0.26	0.34*	0.17	0.25	0.30	--							
8. CBCL Withdep	0.04	0.42**	0.31	0.06	0.17	0.26	0.62**	--						
9. CBCL Somatic	0.31†	0.29†	0.50**	0.19	0.31†	0.46**	0.47**	0.47**	--					
10. CBCL Soc P	0.18	0.04	0.22	0.33*	0.31†	0.30†	0.60**	0.36*	0.67**	--				
11. CBCL Attn	0.48**	0.08	0.27	0.60**	0.39*	0.54**	0.42**	0.29	0.61**	0.66**	--			
12. CBCL Ruleb	0.33*	-0.25	0.16	0.29†	0.22	0.24	0.34*	0.21	0.52**	0.54**	0.52**	--		
13. CBCL Agg	0.60**	0.04	0.45**	0.29†	0.27†	0.52**	0.30*	0.22	0.51**	0.42**	0.58**	0.73**	--	
14. CBCL Int	0.16	0.37*	0.44**	0.17	0.29†	0.39*	0.90**	0.82**	0.74**	0.66**	0.52**	0.42**	0.41**	--
15. CBCL Ext	0.54**	-0.06	0.38*	0.31†	0.27†	0.46**	0.34*	0.23	0.55**	0.49**	0.60**	0.87**	0.97**	0.44**

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Functions. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organize. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale.

Table I16

One-Month Follow-Up Associations Between Parent-Reported Behavior Rating Inventory of Executive Functions and Child Behavior Checklist Scores

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. BRIEF Inhibit	--													
2. BRIEF Shift	0.34*	--												
3. BRIEF EmoC	0.54**	0.57**	--											
4. BRIEF WM	0.67**	0.25	0.30	--										
5. BRIEF PlanO	0.68**	0.23	0.37*	0.89**	--									
6. BRIEF GEC	0.87**	0.62**	0.75**	0.78**	0.79**	--								
7. CBCL Anxdep	0.37*	0.58**	0.37*	0.44**	0.34*	0.53**	--							
8. CBCL Withdep	0.54**	0.71**	0.44**	0.45**	0.38*	0.66**	0.54**	--						
9. CBCL Somatic	0.14	0.23	0.17	0.16	0.17	0.22	0.58**	0.17	--					
10. CBCL Soc P	0.39*	0.30	0.46**	0.18	0.12	0.41*	0.64**	0.38*	0.59**	--				
11. CBCL Attn	0.77**	0.21	0.29	0.65**	0.52**	0.66**	0.50**	0.51**	0.35*	0.52**	--			
12. CBCL Ruleb	0.71**	0.26	0.49**	0.55**	0.55**	0.68**	0.43**	0.52**	0.24	0.43**	0.56**	--		
13. CBCL Agg	0.77**	0.34*	0.69**	0.40*	0.48**	0.73**	0.38*	0.43**	0.14	0.44**	0.52**	0.81**	--	
14. CBCL Int	0.48**	0.70**	0.45**	0.49**	0.40*	0.65**	0.94**	0.73**	0.67**	0.67**	0.57**	0.51**	0.42**	--
15. CBCL Ext	0.78**	0.32	0.63**	0.48**	0.53**	0.75**	0.42**	0.49**	0.19	0.46**	0.56**	0.94**	0.97**	0.48**

Note. $^{***}p < .001$. $^{**}p < 0.01$. $^{*}p < 0.05$. $^{\dagger}p < 0.10$. BRIEF = Behavior Rating Inventory of Executive Functions. EmoC = Emotional Control. WM = Working Memory. PlanO = Plan/Organize. GEC = Global Executive Composite. CBCL = Child Behavior Checklist. AnxDep = Anxious/Depressed. WithDep = Withdrawn/Depressed. Somatic = Somatic Problems. Soc = Social Problems. Attn = Attention Problems. Ruleb = Rule-breaking Behaviors. Agg = Aggressive Behaviors. Int = Internalizing Scale. Ext = Externalizing Scale.

REFERENCES

REFERENCES

Achenbach, T. M. (1991). Child behavior checklist/4--18. *Burlington: University of Vermont*.

Albrecht, B., Brandeis, D., Uebel, H., Heinrich, H., Heise, A., Hasselhorn, M., … Banaschewski, T. (2010). Action monitoring in children with or without a family history of ADHD--effects of gender on an endophenotype parameter. *Neuropsychologia*, *48*(4), 1171–7. http://doi.org/10.1016/j.neuropsychologia.2009.12.018

Barry, R., Johnstone, S., & Clarke, A. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: II. Event-related potentials. *Clinical Neurophysiology*, *114*, 184–198. http://doi.org/10.1016/S

Bergman-Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: a controlled, randomized study. *Developmental Science*, *14*(3), 591–601. http://doi.org/10.1111/j.1467-7687.2010.01022.x

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), 647–63. http://doi.org/10.1111/j.1467-8624.2007.01019.x

Brandeis, D., Van Leeuwen, T. H., Rubia, K., Vitacco, D., Steger, J., Pascual-Marqui, R. D., & Steinhausen, H. C. (1998). Neuroelectric mapping reveals precursor of stop failures in children with attention deficits. *Behavioural Brain Research*, *94*, 111–125. http://doi.org/10.1016/S0166-4328(97)00174-5

Bress, J. N., Meyer, A., & Hajcak, G. (2013). Differentiating Anxiety and Depression in Children and Adolescents: Evidence From Event-Related Brain Potentials. *Journal of Clinical Child and Adolescent Psychology : The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, 1–12. http://doi.org/10.1080/15374416.2013.814544

Broyd, S. J., Johnstone, S. J., Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M., & Lawrence, C. a. (2005). The effect of methylphenidate on response inhibition and the event-related potential of children with attention deficit/hyperactivity disorder. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *58*(1), 47–58. http://doi.org/10.1016/j.ijpsycho.2005.03.008

Burgio-Murphy, A., Klorman, R., Shaywitz, S. E., Fletcher, J. M., Marchione, K. E., Holahan, J., … Shaywitz, B. a. (2007). Error-related event-related potentials in children with attention-deficit hyperactivity disorder, oppositional defiant disorder, reading disorder, and math disorder. *Biological Psychology*, *75*(1), 75–86. http://doi.org/10.1016/j.biopsycho.2006.12.003

Cao, J., Wang, S., Ren, Y., Zhang, Y., Cai, J., Tu, W., … Xia, Y. (2013). Interference control in 6-11 year-old children with and without ADHD: behavioral and ERP study. *International Journal of Developmental Neuroscience : The Official Journal of the International Society for Developmental Neuroscience*, *31*(5), 342–9. http://doi.org/10.1016/j.ijdevneu.2013.04.005

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental neuropsychology*, *28*(2), 595-616.

Carlson, S. M., Mandell, D. J.,& Williams, L. (2004). Executive function and theory of mind: Stability and prediction from age 2 to 3. *Developmental Psychology*, 40, 1105–1122.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. http://doi.org/10.1016/j.jecp.2004.01.002

Casey, B., & Somerville, L. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the …*, *108*(36). http://doi.org/10.1073/pnas.1108561108

Casey, B., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3), 104–110. http://doi.org/10.1016/j.tics.2005.01.011

Caspi, A. (1996). Behavioral Observations at Age 3 Years Predict Adult Psychiatric Disorders. *Archives of General Psychiatry*, *53*(11), 1033. http://doi.org/10.1001/archpsyc.1996.01830110071009

Clark, C. A. C., Chevalier, N., Nelson, J. M., James, T. D., Garza, J. P., Choi, H. J., & Espy, K. A. (2016). I. EXECUTIVE CONTROL IN EARLY CHILDHOOD. *Monographs of the Society for Research in Child Development*, *81*(4), 7-29.

Cortese S., Ferrin M., Brandeis D., Buitelaar J., Daley D., Dittmann R. W., . . . Sonuga-Barke E. (2015). Cognitive training for attention-deficit/hyperactivity disorder: Meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. Journal of the American Academy of Child & Adolescent Psychiatry, 54, 164–174. doi:10.1016/j.jaac.2014.12.010

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology*, *25*, 355–376. http://doi.org/10.1207/s15326942dn2503_6

Davies, P., Segalowitz, S., & Gavin, W. (2004). Development of Response- Monitoring ERPs in 7- to 25- Year-Olds. *Developmental …*, (February 2015), 37–41. http://doi.org/10.1207/s15326942dn2503

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science (New York, N.Y.)*, *333*(6045), 959–64. http://doi.org/10.1126/science.1204529

Drollette, E. S., Scudder, M. R., Raine, L. B., Moore, R. D., Saliba, B. J., Pontifex, M. B., & Hillman, C. H. (2014). Acute exercise facilitates brain function and cognition in children who need it most: an ERP study of individual differences in inhibitory control capacity. *Developmental Cognitive Neuroscience*, *7*, 53–64. http://doi.org/10.1016/j.dcn.2013.11.001

Durbin, C. E., & Wilson, S. (2012). Convergent validity of and bias in maternal reports of child emotion. *Psychological Assessment*, *24*(3), 647–60. http://doi.org/10.1037/a0026607

Eisenberg, N., Fabes, R., Nyman, M., Bernzweig, J., & Pinuelas, A. (1994). The Relations of Emotionality and Regulation to Children' s Anger-related Reactions. *Child …*, *65*(1), 109–128. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8624.1994.tb00738.x/full

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*(5), 1043-1053.

Espinet, S., Anderson, J., & Zelazo, P. (2012). N2 amplitude as a neural marker of executive function in young children: an ERP study of children who switch versus perseverate on the dimensional change card sort. *Developmental Cognitive …*, *25*, 549–558. Retrieved from http://www.sciencedirect.com/science/article/pii/S1878929311001356

Espinet, S. D., Anderson, J. E., & Zelazo, P. D. (2013). Reflection training improves executive function in preschool-age children: behavioral and neural effects. *Developmental Cognitive Neuroscience*, *4*, 3–15. http://doi.org/10.1016/j.dcn.2012.11.009

Espy, K. A., Clark, C. A. C., Garza, J. P., Nelson, J. M., James, T. D., & Choi, H. J. (2016). VI. EXECUTIVE CONTROL IN PRESCHOOLERS: NEW MODELS, NEW RESULTS, NEW IMPLICATIONS. *Monographs of the Society for Research in Child Development*, *81*(4), 111-128.

Fallgatter, A. J., Ehlis, A.-C., Seifert, J., Strik, W. K., Scheuerpflug, P., Zillessen, K. E., … Warnke, A. (2004). Altered response control and anterior cingulate function in attention-deficit/hyperactivity disorder boys. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *115*(4), 973–81. http://doi.org/10.1016/j.clinph.2003.11.036

Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., Locke, J., … Kasari, C. (2010). Effects of Mindful Awareness Practices on Executive Functions in Elementary School Children. *Journal of Applied School Psychology*, *26*(1), 70–95. http://doi.org/10.1080/15377900903379125

Franco, A., Malhotra, N., & Simonovits, G. (2016). Underreporting in psychology experiments: Evidence from a study registry. *Social Psychological and Personality Science*, *7*(1), 8-12.

Gehring, W., Liu, Y., Orr, J., & Carp, J. (2012). The error-related negativity (ERN/Ne). In S. Luck & E. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components* (pp. 231–291). New York: Oxford University Press.

Gerstadt, C. L., Hong, Y. J., & Diamond, a. (1994). The relationship between cognition and action: performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition*, *53*(2), 129–

153. http://doi.org/10.1016/0010-0277(94)90068-X

Gioia, G. A., Espy, K. A., & Isquith, P. K. (2002). *Behavior Rating Inventory of Executive Function, Preschool Version (BRIEF-P).* Odessa, FL: Psychological Assessment Resources.

Goldsmith, H. H., Reilly, J., Lemery, K. S., Longley, S., & Prescott, A. (1995). Laboratory Temperament Assessment Battery: Preschool version. Unpublished manuscript.

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. http://doi.org/10.1016/j.dcn.2014.02.001

Groen, Y., Wijers, A. a, Mulder, L. J. M., Waggeveld, B., Minderaa, R. B., & Althaus, M. (2008). Error and feedback processing in children with ADHD and children with Autistic Spectrum Disorder: an EEG event-related potential study. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *119*(11), 2476–93. http://doi.org/10.1016/j.clinph.2008.08.004

Groom, M. J., Liddle, E. B., Scerif, G., Liddle, P. F., Batty, M. J., Liotti, M., & Hollis, C. P. (2013). Motivational incentives and methylphenidate enhance electrophysiological correlates of error monitoring in children with attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*(8), 836–45. http://doi.org/10.1111/jcpp.12069

Hajcak, G., Franklin, M. E., Foa, E. B., & Simons, R. F. (2008). Increased error-related brain activity in pediatric obsessive-compulsive disorder before and after treatment. *American Journal of* …, 116–123. Retrieved from http://journals.psychiatryonline.org/article.aspx?articleid=99415

Henderson, H. a. (2010). Electrophysiological correlates of cognitive control and the regulation of shyness in children. *Developmental Neuropsychology*, *35*(2), 177–93. http://doi.org/10.1080/87565640903526538

Hillman, C. H., Buck, S. M., Themanson, J. R., Pontifex, M. B., & Castelli, D. M. (2009). Aerobic fitness and cognitive development: Event-related brain potential and task performance indices of executive control in preadolescent children. *Developmental Psychology*, *45*(1), 114–29. http://doi.org/10.1037/a0014437

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nature Reviews. Neuroscience*, *9*, 58–65. http://doi.org/10.1038/nrn2298

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), F9-15. http://doi.org/10.1111/j.1467-7687.2009.00848.x

Hughes, C. (1998). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental psychology*, *34*(6), 1326 - 1339.

Hum, K. M., Manassis, K., & Lewis, M. D. (2013). Neural mechanisms of emotion regulation in childhood anxiety. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*, 552–564. http://doi.org/10.1111/j.1469-7610.2012.02609.x

Imhoff, R., Schmidt, A. F., & Gerstenberg, F. (2014). Exploring the interplay of trait self-control and ego depletion: Empirical evidence for ironic effects. *European Journal of Personality*, *28*(5), 413-424.

Jacques, S., & Zelazo, P. (2001). The Flexible Item Selection Task (FIST): A Measure of Executive Function in Preschoolers. *Developmental Neuropsychology*, *20*, 37–41. http://doi.org/10.1207/S15326942DN2003

Johnstone, S., Barry, R., & Clarke, A. (2013). Ten years on : A follow-up review of ERP research in attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, *124*, 644–657. Retrieved from http://www.sciencedirect.com/science/article/pii/S1388245712006281

Johnstone, S. J., & Galletta, D. (2013). Event-rate effects in the flanker task: ERPs and task performance in children with and without AD/HD. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *87*(3), 340–8. http://doi.org/10.1016/j.ijpsycho.2012.07.170

Johnstone, S. J., Watt, A. J., & Dimoska, A. (2010). Varying required effort during interference control in children with AD/HD: task performance and ERPs. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *76*(3), 174–85. http://doi.org/10.1016/j.ijpsycho.2010.03.010

Jonkman, L. M., van Melis, J. J. M., Kemner, C., & Markus, C. R. (2007). Methylphenidate improves deficient error evaluation in children with ADHD: an event-related brain potential study. *Biological Psychology*, *76*(3), 217–29. http://doi.org/10.1016/j.biopsycho.2007.08.004

Klauer, K. J. (2001). Handbook of cognitive training [Handbuch Kognitives Training]. Göttingen: Hogrefe.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential Development of Attention and Executive Functions in 3- to 12-Year-Old Finnish Children. *Developmental Neuropsychology*, *20*(1), 407–428. http://doi.org/10.1207/S15326942DN2001

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., … Westerberg, H. (2005). Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(2), 177–86. http://doi.org/10.1097/00004583-200502000-00010

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. Child Development, 67, 490–507.

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood:

continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*, 220–232. http://doi.org/10.1037/0012-1649.36.2.220

Kopp, C. (1982). Antecedents of Self-Regulation : A Developmental Perspective. *Developmental Psychology*, *18*(2), 199–214. Retrieved from http://psycnet.apa.org/journals/dev/18/2/199/

Ladouceur, C. D., Conway, A., & Dahl, R. E. (2010). Attentional control moderates relations between negative affect and neural correlates of action monitoring in adolescence. *Developmental Neuropsychology*, *35*(2), 194–211. http://doi.org/10.1080/87565640903526553

Ladouceur, C. D., Dahl, R. E., Birmaher, B., Axelson, D. A., & Ryan, N. D. (2006). Increased error-related negativity (ERN) in childhood anxiety disorders: ERP and source localization. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*, 1073–1082. http://doi.org/10.1111/j.1469-7610.2006.01654.x

Lakes, K. D., & Hoyt, W. T. (2004). Promoting self-regulation through school-based martial arts training. *Journal of Applied Developmental Psychology*, *25*(3), 283–302. http://doi.org/10.1016/j.appdev.2004.04.002

Lin, M., Gavin, W. J., & Davies, P. L. (2015). Developmental trend of error-related negativity (ERN) in 7- to 25-year-olds after adjusting for trial-to-trial variability. Poster presented at annual meeting of Society of Psychophysiological Research, Seattle, WA.

Lo, S. L., Schroder, H. S., Moran, T. P., Durbin, C. E., & Moser, J. S. (2015). Neurophysiological evidence of an association between cognitive control and defensive reactivity processes in young children. *Developmental Cognitive Neuroscience*, *15*, 35–47. http://doi.org/10.1016/j.dcn.2015.09.001

Lo, S. L., Schroder, H. S., Fisher, M. E., Durbin, C. E., Fitzgerald, K. D., Danovitch, J. H., & Moser, J. S. (2016). Associations between Disorder-Specific Symptoms of Anxiety and Error-Monitoring Brain Activity in Young Children. *Journal of Abnormal Child Psychology*, 1-10.

Mathalon, D. H., Whitfield, S. L., & Ford, J. M. (2003). Anatomy of an error: ERP and fMRI. *Biological Psychology*, *64*, 119–141. http://doi.org/10.1016/S0301-0511(03)00105-4

McClelland, M. M., & Cameron, C. (2012). Self-regulation in early child- hood: Improving conceptual clarity and developing ecologically-valid measures. *Child Development* Perspectives, 6, 136–142. http://dx.doi.org/10.1111/ j.1750-8606.2011.00191.x

McDermott, J. M., Perez-Edgar, K., Henderson, H. a, Chronis-Tuscano, A., Pine, D. S., & Fox, N. a. (2009). A history of childhood behavioral inhibition and enhanced response monitoring in adolescence are linked to clinical anxiety. *Biological Psychiatry*, *65*(5), 445–8. http://doi.org/10.1016/j.biopsych.2008.10.043

Melby-Lervåg M., Hulme C. (2013). Is working memory training effective? A meta-analytic review. Developmental Psychology, 49, 270–291. doi:10.1037/a0028228

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer" evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*(4), 512-534.

Meyer, A., Weinberg, A., Klein, D. N., & Hajcak, G. (2012). The development of the error-related negativity (ERN) and its relationship with anxiety: evidence from 8 to 13 year-olds. *Developmental Cognitive Neuroscience*, *2*(1), 152–61. http://doi.org/10.1016/j.dcn.2011.09.005

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2693–8. http://doi.org/10.1073/pnas.1010076108

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*, 466. http://doi.org/10.3389/fnhum.2013.00466

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2014). The case for compensatory processes in the relationship between anxiety and error monitoring: a reply to Proudfit, Inzlicht, and Mennin. *Frontiers in Human Neuroscience*, *8*(February), 64. http://doi.org/10.3389/fnhum.2014.00064

Murray, K., & Kochanska, G. (2002). Effortful control: Factor structure and relation to externalizing and internalizing behaviors. *Journal of Abnormal Child Psychology*, *30*(5), 503–514. Retrieved from http://link.springer.com/article/10.1023/A:1019821031523

Pliszka, S. R., Liotti, M., & Woldorff, M. G. (2000). Inhibitory control in children with attention-deficit/hyperactivity disorder: Event-related potentials identify the processing component and timing of an impaired right-frontal response-inhibition mechanism. *Biological Psychiatry*, *48*, 238–246. http://doi.org/10.1016/S0006-3223(00)00890-8

Pontifex, M. B., Raine, L. B., Johnson, C. R., Chaddock, L., Voss, M. W., Cohen, N. J., … Hillman, C. H. (2011). Cardiorespiratory fitness and the flexible modulation of cognitive control in preadolescent children. *Journal of Cognitive Neuroscience*, *23*(6), 1332–45. http://doi.org/10.1162/jocn.2010.21528

Pontifex, M. B., Saliba, B. J., Raine, L., Picchietti, D., & Hillman, C. (2013). Exercise Improves Behavioral, Neurocognitive, and Scholastic Performance in Children with ADHD. *Journal of Pediatrics*, *162*(3), 543–551. http://doi.org/10.1016/j.jpeds.2012.08.036.Exercise

Proudfit, G. H., Inzlicht, M., & Mennin, D. S. (2013). Anxiety and error monitoring: the importance of motivation and emotion. *Frontiers in Human Neuroscience*, *7*(October), 1–4. http://doi.org/10.3389/fnhum.2013.00636

Rapport M. D., Orban S. A., Kofler M. J., Friedman L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. Clinical Psychology

Review, 33, 1237–1252. doi:10.1016/j.cpr.2013.08.005

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., ... & Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*(2), 359.

Riesel, A., Weinberg, A., Endrass, T., Meyer, A., & Hajcak, G. (2013). The ERN is the ERN is the ERN? Convergent validity of error-related brain activity across different tasks. *Biological Psychology*, *93*(3), 377–85. http://doi.org/10.1016/j.biopsycho.2013.04.007

Rosch, K. S., & Hawk, L. W. (2013). The effects of performance-based rewards on neurophysiological correlates of stimulus, error, and feedback processing in children with ADHD. *Psychophysiology*, *50*(11), 1157–73. http://doi.org/10.1111/psyp.12127

Rothbart, M. (1989). Temperament in childhood: A framework. In *Temperament in childhood* (pp. 59–73).

Rothbart, M., Ellis, L. K., Rueda, M. R., & Posner, M. I. (2003). Developing Mechanisms of Temperamental Effortful Control. *Journal of Personality*, *71*, 1113–1143. http://doi.org/10.1111/1467-6494.7106009

Rothbart, M., & Rueda, M. R. (2005). The development of effortful control. In *Developing individuality in the human brain: A tribute to Michael I. Posner* (pp. 167–188). http://doi.org/10.1037/11108-009

Rothbart, M., Sheese, B. E., & Posner, M. I. (2007). Executive Attention and Effortful Control: Linking Temperament, Brain Networks, and Genes. *Child Development Perspectives*, *1*(1), 2–7. http://doi.org/10.1111/j.1750-8606.2007.00002.x

Rueda, M. R., Checa, P., & Cómbita, L. M. (2012). Enhanced efficiency of the executive attention network after training in preschool children: immediate changes and effects after two months. *Developmental Cognitive Neuroscience*, *2 Suppl 1*, S192-204. http://doi.org/10.1016/j.dcn.2011.09.004

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(41), 14931–6. http://doi.org/10.1073/pnas.0506897102

Sala, G., & Gobet, F. (2017). Working Memory Training in Typically Developing Children: A Meta-Analysis of the Available Evidence. *Developmental Psychology*, 37–41.

Santesso, D., Segalowitz, S., & Schmidt, L. (2006). Error-Related Electrocortical Responses Are Enhanced in Children With Obsessive – Compulsive Behaviors. *Developmental …*, *29*(3), 37–41. http://doi.org/10.1207/s15326942dn2903

Schmitt, S. a., McClelland, M. M., Tominey, S. L., & Acock, A. C. (2014). Strengthening school readiness for Head Start children: Evaluation of a self-regulation intervention. *Early Childhood*

Research Quarterly, 1–12. http://doi.org/10.1016/j.ecresq.2014.08.001

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 : Official Journal of the International Organization of Psychophysiology*, *85*(1), 93–105. http://doi.org/10.1016/j.ijpsycho.2011.05.007

Shiels, K., & Hawk, L. W. (2010). Self-regulation in ADHD: the role of error processing. *Clinical Psychology Review*, *30*(8), 951–61. http://doi.org/10.1016/j.cpr.2010.06.010

Sibley, B., & Etnier, J. (2003). The relationship between physical activity and cognition in children: a meta-analysis. *Pediatric Exercise Science*, *2000*(15), 243–256. Retrieved from http://www.humankinetics.com/acucustom/sitename/Documents/DocumentItem/2196.pdf

Smith, J. L., Johnstone, S. J., & Barry, R. J. (2004). Inhibitory processing during the Go/NoGo task: an ERP analysis of children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *115*(6), 1320–31. http://doi.org/10.1016/j.clinph.2003.12.027

Spronk, M., Jonkman, L. M., & Kemner, C. (2008). Response inhibition and attention processing in 5- to 7-year-old children with and without symptoms of ADHD: An ERP study. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *119*(12), 2738–52. http://doi.org/10.1016/j.clinph.2008.09.010

Strommen, E. A. (1973). Verbal self-regulation in a children's game: Impulsive errors on "Simon says." Child Development, 44, 849–853.

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106–13. http://doi.org/10.1111/j.1467-7687.2008.00745.x

Tominey, S. L., & McClelland, M. M. (2011). Red Light, Purple Light: Findings From a Randomized Trial Using Circle Time Games to Improve Behavioral Self-Regulation in Preschool. *Early Education & Development*, *22*(3), 489–519. http://doi.org/10.1080/10409289.2011.574258

Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A. J., Dyson, M. W., Olino, T. M., & Klein, D. N. (2013). Error-related brain activity in young children: associations with parental anxiety and child temperamental negative emotionality. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*(8), 854–62. http://doi.org/10.1111/jcpp.12041

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.

Tuckman, B. W., & Hinkle, J. S. (1986). An experimental study of the physical and psychological effects of aerobic exercise on schoolchildren. *Health Psychology*, *5*, 197–207.

Tye, C., Asherson, P., Ashwood, K. L., Azadi, B., Bolton, P., & McLoughlin, G. (2014). Attention and inhibition in children with ASD, ADHD and co-morbid ASD + ADHD: an event-related potential study. *Psychological Medicine*, *44*(5), 1101–16. http://doi.org/10.1017/S0033291713001049

van Meel, C. S., Heslenfeld, D. J., Oosterlaan, J., & Sergeant, J. a. (2007). Adaptive control deficits in attention-deficit/hyperactivity disorder (ADHD): the role of error processing. *Psychiatry Research*, *151*(3), 211–20. http://doi.org/10.1016/j.psychres.2006.05.011

van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*(4), 593–602. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/08989290260045837

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. a, & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage*, *14*(6), 1302–8. http://doi.org/10.1006/nimg.2001.0923

Weinberg, A., & Hajcak, G. (2011). Longer term test-retest reliability of error-related brain activity. *Psychophysiology*, *48*, 1420–1425. http://doi.org/10.1111/j.1469-8986.2011.01206.x

Wechsler, D. (2003). Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). San Antonio, TX: The Psychological Corporation.

Whittle, S., Allen, N. B., Lubman, D. I., & Yücel, M. (2006). The neurobiological basis of temperament: towards a better understanding of psychopathology. *Neuroscience and Biobehavioral Reviews*, *30*(4), 511–25. http://doi.org/10.1016/j.neubiorev.2005.09.003

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental psychology*, *44*(2), 575.

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: psychometric properties and relationship to academic achievement. *Psychological assessment*, *24*(1), 226.

Wolfe, C. D., & Bell, M. A. (2004). Working memory and inhibitory control in early childhood: Contributions from physiology, temperament, and language. *Developmental Psychobiology*, *44*(1), 68–83. http://doi.org/10.1002/dev.10152

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and biological engineering*, *5*(6), 539-554.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). *The neural basis of error detection: conflict monitoring and the error-related negativity. Psychological review* (Vol. 111).

Yeung, N., & Cohen, J. D. (2006). The Impact of Cognitive Deficits on Conflict Monitoring. *Psychological Science*, *17*(2), 164–171.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*. http://doi.org/10.1098/rstb.2011.0416

Zelazo, P.D., 2006. The dimensional change card sort (DCCS): a method of assessing executive function in children. Nature Protocols 1, 297–301.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. http://doi.org/10.1111/mono.12032