MOMMY AND ME: HOW THE RELATIONSHIP BETWEEN MATERNAL AFFECT AND INFANT TEMPERAMENT PREDICTS EXECUTIVE FUNCTION DEVELOPMENT AND ACADEMIC ACHIEVEMENT

A thesis presented to the faculty of the Graduate School of
Western Carolina University in partial fulfillment of the
requirements for the degree of Specialist in School Psychology.

By

Nina Andre

Director: Dr. Alleyne Broomell
Assistant Professor of Psychology
Psychology Department

Committee Members: Dr. Lori Unruh, Psychology (co-director)
Dr. Ethan Schilling, Psychology

May 18, 2021

ACKNOWLEDGEMENTS

I would like to thank my committee members and director for their assistance and encouragement. In particular, I would like to thank Dr. Alleyne Broomell for her ongoing support and many hours spent assisting me with conducting analyses and providing feedback. Thank you for your willingness to take me on as a student and always expecting the most from me. Great thanks are also due to Dr. Martha Ann Bell for allowing me to use this data set, and all of the researchers at the Cognition, Affect, and Physiology Lab who collected that data.

I also extend sincere gratitude to my parents, without whom this thesis would not be possible. Thank you for investing in my education for so many years and supporting me through every step. Thank you for providing me with high expectations and unconditional love, and never failing to remind me what truly matters.

I would also like to thank the psychology professors at Anderson University for the strong foundation of knowledge and appreciation for the importance of the details. Specifically, I would like to thank Dr. Robert Franklin for his role in providing the opportunity to begin research and his encouragement to pursue it in the future. And to my original research partner, Madison, thank you for being the best research buddy and listening to me talk about this project for hours.

Lastly, my deepest gratitude to my fiancé, Chris. Thank you for encouraging me on the hard days and always being willing to get ice cream to cheer me up. I couldn’t imagine life with anyone else.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

MOMMY AND ME: HOW THE RELATIONSHIP BETWEEN MATERNAL AFFECT AND INFANT TEMPERAMENT PREDICTS EXECUTIVE FUNCTION DEVELOPMENT AND ACADEMIC ACHIEVEMENT.

Nina Andre

Western Carolina University (May 18, 2021)

Director: Dr. Alleyne Broomell

Infancy is a crucial period of development when children's experiences are shaped by both their internal traits and external circumstances. These influential factors shape the foundation of executive function abilities that can impact future outcomes, including academic achievement. The focus of this study was the interaction between infant temperament and maternal positive affect predicting later executive function (EF) skills and academic achievement. Data on the target variables were collected from 304 infants at 10-months, and subsequently at 48-months, and 9-years, and EF was assessed using both parent report and behavioral measures. Results indicated a significant relationship between infant negative affect and later EF, but not maternal positive affect nor an interaction between the variables. Measures of EF did not show a significant path from EF to math or reading achievement, but post-hoc analyses using separate models considering only task measures of EF did support a significant path from EF to academic achievement. Practical applications and future directions are discussed.

Understanding executive function (EF) development during early childhood can provide important insight into individuals' current functioning and future life outcomes, including academic achievement. Children's cognitive development during the first year of life is impacted by many factors, including infant temperament (Liu et al., 2018; Putnam et al., 2008) and parental interaction (Graziano et al., 2011; Hughes & Devine, 2019), and EF as early as infancy has been shown to predict later cognitive and academic success (e.g. Blankenship et al., 2019; Cheng et al., 2018; Cuevas et al., 2014; Kraybill & Bell, 2013). A sample of Kindergarten teachers noted that the factor most impacting school readiness was not academic skills, but self-regulation abilities (Rimm-Kuafman et al., 2000). Deficits in EF have been linked with many common pediatric disorders, including autism spectrum disorder (Schmitz et al., 2006), attention deficit hyperactivity disorder (Shallice et al., 2002), obsessive compulsive disorder, and Tourette's syndrome (Watkins et al., 2005). Even within the normal range, variations in EF abilities are predictive of later outcomes (Carlson, 2005), and is often more predictive than IQ alone (Duckworth & Seligman, 2005). Early identification of executive dysfunction is essential for early intervention opportunities that can have important benefits for later life outcomes, specifically regarding academic achievement. Research indicates that EF based interventions may be a salient approach in reducing socioeconomic gaps in achievement levels (Zelazo & Carlson, 2020). Therefore, further understanding of the measurement of EF abilities and the association with academic outcomes is essential for pursuing more equitable educational opportunities for all students.

This study employs a developmental perspective to understand factors which impact EF development, while considering different measurements of EF, and the value of early EF in predicting later academic achievement. Most studies on this topic have looked at internal and environmental factors separately, and either behavioral or survey measures of EF, but relatively few have considered all variables concurrently. The aim of this study is to address this gap and consider the longitudinal impacts of EF measurement differences and the infant temperament-maternal care interaction on future academic success.

Executive Function Development

An essential part of cognitive development is the emergence of EF. These abilities are the cognitive skills which underlie the rapid social (Razza & Blair, 2009), and behavioral (Bathelt et al., 2018) development which occurs in early childhood. EF allows children to coordinate behavior and meet the expectations of their environments (Barkley, 2001), and can be thought of as an air traffic control system in the brain (Center on the Developing Child, 2011). EF has been defined in various ways throughout the literature, but essentially is the cognitive abilities that allow individuals to plan and carry out goal-directed behavior. Most conceptions of EF include the separate but related aspects of inhibition, working memory, and cognitive flexibility (Diamond, 2013; Miyake et al., 2000). Inhibitory control is the individual's ability to regulate their actions and reactions, and not act based solely on impulse and environmental cues. It is a difficult skill for young children and continues to develop through adolescence. Working memory is holding information in immediate consciousness and manipulating it as needed. Working memory and inhibitory control typically function simultaneously and both are needed to support each other. Basic working memory develops in infancy, but the ability to reorder conscious information is much slower to develop. Cognitive flexibility builds upon the other skills, but develops later in life. It is being able to consider different perspectives, adapt to changing circumstances, and develop new solutions when others have been unsuccessful (Diamond, 2013). Although each aspect of EF is distinct, factor analysis reveals a moderate correlation between the elements (Miyake et al., 2000). EF in early childhood is less diverse than in adults and strongly relates to one underlying factor (Wiebe et al., 2011).

According to the Cognitive Complexity and Control theory, EF develops naturally with age-related increases in children's ability to understand and apply complex rules in completing tasks and solving problems (Zelazo et al., 2003). Although much research has focused on the prefrontal cortex (PFC) as a region that develops later in adolescence, there is substantial support that PFC development is especially impacted by experiences in early life (Hodel, 2018; Carlson, 2005). PFC development, which underlies EF increases, begins in infancy (Hodel, 2018) and indicates a rapid increase in EF abilities from 2 to 5 years (Nelson et al., 2008). Neuroimaging also suggests that the neural architecture of EF is established in childhood and functions in the frontoparietal and cingulo-opercular networks similarly to those in adults (Engelhardt et al., 2019). Developing precursory skills of selective attention, evaluating information to make decisions, and resisting impulsive decisions is facilitated by experiences and opportunities to practice. Infants are born with genetic potential to develop EF skills, but it is through their interactions with the environment throughout development that this occurs and gene expression is determined (Grossman et al., 2003). Both internal and external factors impact the child’s experiences in their world.

Factors That Impact Executive Function Development

Infant temperament

One internal factor which impacts a child’s development of EF skills is infant temperament (Wolfe & Bell, 2007). Infant’s temperament both impacts and is impacted by daily interactions with the environment (Rothbart et al., 2000). Temperament is a child’s typical style of reaction to stimuli in the environment. The trait is largely biologically based and shows stability through infancy (Bornstein et al., 2015). Thus, it is a useful intrinsic factor to consider in the development of EF in early life. Infant temperament has been measured in various ways throughout the literature. Early conceptions included dimensions such as activity level, intensity, and adaptability (Thomas & Chess, 1977), but these original structures have now been updated using new psychometric approaches (Rothbart, 2007). Perhaps the most widely used modern assessment is the Infant Behavior Questionnaire (IBQ) developed by Rothbart (1981). The IBQ is a parent report form in which parents are asked to report the typical responses of their infant on 94 questions. These questions are then split into six scales: Activity Level, Smiling and Laughter, Fear, Distress to Limitations, Duration of Orienting, and Soothability. These scales have been further broken down into a three-factor model of infant temperament, which includes two reactive dimensions and one regulatory dimension (Rothbart et al., 2013). Each of these factors may be differentially associated with EF, although few studies have been completed in the area and results are varied (Conway & Stifter, 2012; Suor et al., 2019; Wolfe & Bell, 2007).

The first of the reactive dimensions, negative affectivity, includes measures of fear, discomfort, and frustration. Studies have consistently supported the negative association between negative affectivity and measures of cognitive functioning (Liu et al., 2018; Putnam et al., 2008). Extraversion/surgency, which includes positive emotionality, high levels of approach and intensity pleasure, and low levels of shyness, impulsivity, and affiliation, has been less consistently associated with cognitive outcomes. Frick and colleagues (2018) found that surgency in infancy was associated with global EF at 18 months, but had no unique contribution to the global EF measure when also considering sustained attention. Although there was an association, it was likely due to an overlap in the measurement of surgency and attention, which is related to early EF. There is a positive relationship between infant surgency and toddler effortful control, but a negative relationship between toddler surgency and effortful control (Putnam et al., 2008). The complexity and at times inconsistency in the research on the relationship between temperament and cognition further strengthens the likelihood of other factors that may impact the relationship between infant temperament and EF. Therefore it is essential to consider multiple predictors of EF development, especially external factors which could serve as an intervention point for improving EF abilities, such as early parental interaction.

Maternal positive affect

For many children, maternal relationships are the primary source of interaction during infancy and early childhood and have an impact on later EF (Valcan et al., 2018). Many parenting behaviors could have potential impacts on later EF, including scaffolding, stimulation, control, and sensitivity/responsiveness (Fay-Stammbach et al., 2014). The parent behavior considered in this study was maternal positive affect, measured by observing interactions between the mother and infant, specifically focusing on her vocal affect and facial expressions. Defined in this way, maternal positive affect has been found to be related to later EF. Maternal positive affect during infancy was shown as a predictor of EF at 4 and 6 years by Kraybill and Bell in 2013. Another 2017 study by Swingler, Perry, Calkins, and Bell further supported the association. The study found a direct and positive relationship between maternal affect at 5-months and EF at 10-months, and an inverse relationship between maternal intrusiveness and EF. The connection between parental responsiveness and positivity with EF may be due to maternal sensitivity's role in a child developing secure attachment and is thought to improve regulatory control (Fay-Stammbach et al., 2014). The current work will extend these findings by considering the interaction between infant temperament and maternal positive affect as a predictor of later EF.

Infant temperament and parental behavior

Recent research supports the interaction of multiple variables in development, and it is essential to apply this approach in understanding infant predictors of EF. One such interaction occurs between infant temperament and parenting, which then impacts the formation of self-regulation a concept related to EF. Specifically, researchers have found that children high in negative affectivity were low in self-regulation when they were in relationships with low responsiveness, but high in self-regulation in highly responsive relationships (Kim & Kochanska, 2012). Results of another study indicated that children with more negative temperaments are more impacted by maternal behaviors than children with less difficult temperaments (Rochette & Bernier, 2016). Those with difficult temperament at age 1, who experienced negative or hostile maternal interactions, did worse on measures of impulse control at age 3, but those who experienced positive maternal interaction performed best (Rochette & Bernier, 2016). This finding aligns with the differential susceptibility theory, which proposes that some individuals are not only more susceptible to the effects of negative experiences, but also to the effects of positive experiences (Belsky & Pleuss, 2009). The differences that increase an individual's susceptibility to outside influences include temperamental characteristics, physiological differences, and genetic differences (Belsky & Pleuss, 2009). Rochette and Bernier included many measures of EF in their investigation into the differential impact of maternal interactions on children with difficult temperaments, and found these children were more susceptible to the impact of positive caregiving in developing the impulse control aspect of EF, but not conflict EF. These results highlight the nuance of measuring EF and potential differences in the developmental trajectory of different aspects of EF. In order to understand the impacts of other factors on the development of EF, there must be valid and consistent ways to measure EF.

Measurement of Executive Function

In considering EF as a potential area for intervention to improve children's life outcomes, it is imperative to verify a reliable measurement of the construct. Since EF is a complicated system of interrelated skills, one primary challenge is in measuring these abilities in children. Measures of EF have to be adapted according to the age of the child to match their developmental level. For example, a preschool child may not be able to correctly count groups of objects, an ability required for the *Stroop Number* task, but most are familiar with the concepts of day and night. Researchers are then able to evaluate the same underlying ability of inhibiting an automatic response by asking children to say day when shown a card depicting a night sky (Montgomery & Koeltzow, 2010). When conducting longitudinal research with EF it is important to consider the impact of different measurement approaches being employed (Carlson, 2005). The two primary approaches used in measuring EF is through behavioral tasks and rating scales. These differential measurement tactics can have both shared and unique influences when predicting later outcomes (Dekker et al., 2017).

Behavioral Measures

Behavioral EF measures are typically neurocognitive tasks originally designed for adults that have been adapted for use in childhood, although some have been specifically designed for the developmental level of children (Carlson, 2005). One measure developed for infants is the *A-not-B* task. In this task a child must attend to the cues of the researcher when their name is called, remember which cup the ball was placed underneath, and suppress competing stimulus from past trials and the environment (Piaget, 1954). To correctly complete these steps infants must have developed some degree of working memory, inhibition, and attention. Children at 7 months have difficulty inhibiting their response towards a previously correct location, while at 12-months old most children are successful (Diamond, 1990; Diamond et al., 1997).

Although young children are not able to complete the original *Stroop* tests which require inhibiting automatic responses to numerical or lexical stimuli, several Stroop like tasks have been developed for use with children. One of these is the *Day/Night* task during which a child responds to a card with a picture of sun by saying night. While less than half of all three-year-olds can complete the task successfully, 80% of 5-year olds are successful, illustrating the development of EF skills which occurs during this period (Montgomery & Koeltzow, 2010). Another well-known child adaptation is the *Dimension Change Card Sort* (DCCS) task, where children are asked to sort cards according to a specific rule, then another rule, then sort some cards according to one rule and other cards according to another (Zelazo, 2006). These tests primarily measure inhibitory aspects of EF similarly to their counterparts, the Number Stroop and the *Wisconsin Card Sort Task* (WCST). The Number Stroop task requires an individual to inhibit the automatic response to respond to switching rules in counting letters and numbers, while the WCST is a card sort in which the rules of sorting change. Like the DCCS, the WCST requires the individual to flexibly switch between rules when sorting cards.

Behavioral measures like those discussed above are commonly used measures of EF when considering its impact on other outcomes, including academic achievement (Cortés Pascual et al., 2019). Clearly this approach has many benefits, including the ability to evaluate specific aspects of EF and be completed in controlled laboratory settings. The laboratory setting allows researchers to minimize experimenter bias and evaluate performance using clearly defined procedures. As with many tasks completed in laboratories, there are downfalls regarding the ecological validity of these measures (Miranda et al., 2015). Children may be reluctant to engage in an unfamiliar setting and the highly structured environment may not match their typical life.

Survey Measures

Another approach to measuring EF in childhood is through rating scales completed by the caregiver. A common rating scale used with children is the Behavior Rating Inventory of Executive Function (BRIEF) and the preschool version the Behavior Rating Inventory of Executive Function (BRIEF-P). The BRIEF provides two specific indexes, Behavior Regulation and Metacognition, as well as a Global Executive Composite (Gioia et al., 2000b). The BRIEF-P also provides three specific indexes, Inhibitory Self-Control, Flexibility, and Emergent Metacognition, and the Global Executive Composite (Gioia et al., 2000a).

Survey measures provide the unique benefit of allowing individuals who commonly interact with the child and know them well to report their abilities. Weaknesses of the approach include concerns with construct validity, with some studies showing that survey measures of EF may actually measure goal pursuit (Toplak et al., 2013) or externalizing behaviors (Spiegel et al., 2017). Parents may also have a different perspective of their child's EF abilities than a teacher, due to the differential expectations of the environments. Although these scales are commonly used and strongly related to measures of behavioral disruption, scores on these behavior rating inventories are not always correlated with task measures of EF (Mcauley et al., 2010; Toplak et al., 2013).

Consistency Between Measurement Approaches

Both performance-based tests and EF rating scales are accepted approaches for measuring EF, but as mentioned above, correlations between these approaches are often low. Mcauley and colleagues (2010), found that in a population of 6-15 year olds, scores on the

BRIEF Behavior Regulation Index and Metacognition Index were not related to task measures of inhibition, monitoring, or working memory. In contrast, the BRIEF scores were related to ratings of impairment and disruptive behavior from parents and teachers. This raises the question of if the BRIEF is truly measuring EF (Mcauley et al., 2010). Another study which compared performance-based tests and rating scales analyzed the associations of the different measures with ADHD symptoms and reading achievement (Miranda et al., 2015). Overall they found moderate correlations between the measures of EF, dependent on the specific domain, and both measures were predictive of ADHD symptoms and reading achievement. Specifically, the BRIEF explained more variance in ADHD ratings, while task measures were better predictors of reading achievement (Miranda et al., 2015). Yet another study found that when considering performance measures of EF in predicting math outcomes, parent and teacher ratings did not explain any additional variance beyond the performance measures, and only the performance measure and teacher report were uniquely predictive of spelling ability (Dekker et al., 2017). A 2013 meta-analysis considered 20 studies which used both performance-based and rating scale measures of EF. They found that only 24% of the correlations between measures in these studies were significant, and of those that were significant the median correlation was .19 (Toplak et al., 2013). Based on these findings, the authors conclude that performance measures may be more related to cognitive efficiency, whereas rating scales provide a more nuanced perspective on an individual's ability to apply EF skills to meet goals (Toplak et al., 2013). A recent study considered both performance test and rating scales measures of EF and academic achievement, completed by multiple informants. Researchers found a modest correlation between the EF measurements approaches ($r = .30$). Test measures of EF showed significantly higher predictive validity with both rating scale and test measures of academic achievement, although survey measures also uniquely predicted academic achievement (Soto et al., 2020). The current study will use a data-driven approach to determine which behavioral and rating scale measures load onto a latent factor of EF. Using confirmatory factor analysis provides a more complete context to understand the shared and unique variances among constructs with the goal to better predict academic achievement (Karr et al., 2018).

Executive Function and Academic Achievement

EF skills are essential for academic success, these abilities allow students to attend to academic tasks, follow teacher directions, and behave appropriately in the classroom. Since EF begins developing in infancy and can be measured starting at an early age, it serves as an early indicator of future academic competence and target for interventions. The potential influence of EF on academic achievement is a current focus in the fields of education and psychology alike (Best et al., 2011; Monette et al., 2011; Willoughby et al. 2017). Recent meta-analyses reported moderate associations between EF and academic achievement (Allan et al., 2014; Cortes Pascual et al., 2019; Jacob & Parkinson, 2015). Although there appears to be some relationship between EF and academic success, the relationship between EF in infancy and early childhood and academic performance in middle childhood is less clear. An analysis of a large representative sample of kindergarteners found EF significantly predicted reading, math, and science achievement in second grade (Morgan et al., 2019). A recent review article highlights the importance of EF skills for academic achievement, especially in math, and that EF skill training interventions have the potential to prepare students for academic success, especially those from a disadvantaged background (Zelazo & Carlson, 2020). Prior research on the dataset used for the current project has established the relationship between negative affectivity at age 2 and reading and math achievement at age six, moderated by EF at age 4 (Liu et al., 2018), as well as the relationship between EF and reading abilities at age six (Blankenship et al., 2019). The current study aims to extend this prior work and further understand the impacts of prior measures of EF, beginning in infancy, on academic achievement at age 9.

When considering the relationship between EF and academic achievement it is also important to consider the impact of verbal ability. Fuhs and Day's (2011) work with preschoolers showed verbal ability to be a prerequisite of EF development and supported a correlation between the development of the two areas. By considering EF and verbal ability independently when predicting academic achievement, we are able to better understand the impact of each variable.

Statement of Purpose

It is clear that EF is a complex conglomeration of cognitive abilities which develops beginning in infancy and through early childhood. There are both internal and external factors which are related with the development of these abilities. Different measurement approaches can also contribute differentially to an overall understanding of the EF abilities, and how they relate to academic achievement. The aim of the current study is to identify factors in infancy that influence latent factors of later EF and academic achievement, to ultimately identify areas for intervention with the potential to improve academic achievement.

Hypotheses (see Figures 1-6)

I. Some measures of EF at 10 months, 48 months, and 9 years will load onto a latent factor.

II. The main effect of infant temperament and maternal positive affect, (MPA) as well as prior measures of EF, will predict EF at each time point. The impact of infant temperament and MPA will decrease over time.

a. Temperament and MPA at 10 months will predict EF at 10 months.

b. EF, temperament and MPA at 10 months, will predict EF abilities at 48 months.

c. Prior measures of EF, and temperament and MPA at 10 months, will predict EF abilities at 9 years.

III. The interaction between temperament and maternal positive affect will impact EF at 10 months.

a. Infants with easy or difficult temperaments, who experience high levels of maternal positive affect will have higher EF scores.

b. Infants with easy or difficult temperaments, who experience low levels of maternal positive affect will have lower EF scores.

IV. EF and verbal ability at 48 months and 9 years will predict the relationship between EF at 10 months and academic achievement at 9 years.

Method

Participants

Participants were recruited as part of projects conducted by the Cognitive, Affect, and Psychophysiology Lab at Virginia Tech as well as a satellite recruitment location at the University of North Carolina at Greensboro. The sample included 304 infants (153 girls), who were recruited at 5 months. This sample includes only 2 of the 3 cohorts, as the first did not complete a 9 year follow up. All infants included in the study were healthy at the time of recruitment. On a demographics survey completed at the initial evaluation, 18% of children were reported to be Black/African American, 73% White, .03% Asian, 8% Other/multiracial, and 1% did not respond. Mother's education levels varied, 3% did not finish high school, 29% had a high school diploma, 45% had a bachelors or technical degree, and 21% had a graduate degree. At 9 years parents reported that 8 of the children had a developmental delay. These children were included in the sample to reflect the diversity of the population. Each child was assessed at 5, 10, 24 and 48 months as well as 6 and 9 years. Only the 10 month, 48 month, and 9 year data will be included in this study. A total of 233 children returned at 24 months, 202 at 48 months, and 167 at 9 years, which is a typical attrition rate for longitudinal studies of this type.

Procedures

Procedures were consistent between the two collection sites. Mothers and children came into the lab and after an explanation of the procedures consent was obtained from the parent. The parent also completed a general demographic survey when they arrived at 5 months.

Ten month visit

Maternal positive affect. Behaviors were coded during two interactions between the mother and child. First, a two minute free play task where mothers were asked to play typically with their infant sitting in a high chair, using the two toys provided was completed. Next, with the child still in the high chair, they were asked to play a structured game of peek-a-boo using their hands or a cloth. Interactions were coded in 30 second intervals for the mother's facial expressivity, specifically smiling, and vocal affect. The coding system was based on the work of Calkins and colleagues (2004). A score of 1 was given if there was no evidence of positive emotion, while a score of 4 indicated high levels of positive emotion exhibited, including laughing. Ratings were completed from video recordings by trained research assistants. A secondary observer coded 20% of interactions with toys and 30% of interactions while playing peek-a-boo to confirm the coding reliability. The Interclass Correlation Coefficient (ICC) was calculated to be .82 and .90 respectively.

Infant temperament. Infant temperament was measured by the Infant Behavioral Questionnaire- Revised (IBQ-R; Gartstein & Rothbart, 2003). Mothers completed this survey which measured their perception of their infant's temperament during the past two weeks. Responses were indicated on a 7-point Likert scale, from never (1) to always (7). The scores on the questionnaires are split into 14 subscales; activity level, distress to limitations, fear, orienting, smiling and laughing, high pleasure, low pleasure, soothability, falling reactivity, cuddliness, perceptual sensitivity, sadness, approachability, and vocal reactivity. These scales are split into three main factors; surgency/extraversion, negative affectivity, and orienting/regulation. Internal reliability of all subscales was acceptable (0.70 – 0.90) in a sample of infants 3 to 9 months (Gartstein & Rothbart, 2003). There was also moderate evidence of discriminant validity between subscales and convergent validity with observed measures of temperament (Parade & Leerkes, 2008).

A-not B-task. Infant inhibition and working memory was measured using the *A-not-B* behavioral task (Clearfield et al., 2006; Diamond, 2001). The child sat in the mother's lap across from the researcher who showed them a toy and then placed it under a cup. The researcher then distracted the child by snapping, and then would say the child's name and ask, "Where's the toy?". The eye movement of the child was coded as correct or incorrect according to the cup they looked towards and immediate verbal feedback was given. After the child completed two correct trials, the toy was switched and hidden under the opposite cup and the distraction procedure was lengthened to two seconds. A secondary observer coded 24% of interactions to confirm the coding reliability and the ICC was calculated to be .97.

Puppet task. The ability to regulate attention was measured using the *Puppet* task. The task is based on the premise that children with less developed executive abilities will take longer to disengage their fixation from an object or inhibit and shift their attention to other objects (Frick et al., 1999). Infants were seated across from the examiner and shown either a red or grey glove. The glove was worn by the researcher and had a painted face and objects dangling from it. The duration, number of looks, and longest look at each of the puppets was recorded. Coding of each variable was completed from video recordings by trained research assistants. A secondary observer coded 20% of interactions to confirm the coding reliability and the ICC was calculated to be .94.

48 month visit

BRIEF- P. When children returned at age four, parents were asked to complete the Behavior Rating Inventory of Executive Function Preschool Version (BRIEF- P). This measure assesses the executive function skills the parents observe their children exhibiting during day to day interactions. There are five major scales derived from the 63 Likert style questions; Inhibit, Shift, Emotional Control, Working Memory, and Plan/Organize. The scales are further combined to create three primary indexes, Inhibitory Self-Control Index (ISCI), Flexibility Index (FI), and Emergent Metacognition Index (EMI). The ISCI is made up of the inhibit scale, which measures the child's ability to control impulses and behavior, and the emotional control scale, which measures appropriate emotional responses to situations. The FI is made up of the shift scale, which measures the child's ability to solve problems flexibly, and the previously described emotional control scale. The EMI is made up of the working memory scale, which measures the child's ability to hold information in their consciousness to complete a task, and the plan/organize scales, which measure the ability to anticipate future events and implement steps ahead to complete a task. A Global Executive Composite (GEC) score is also produced. Scores on the BRIEF-P are interpreted so that a lower score indicates greater EF ability, while a higher score indicates potential dysfunction. As of 2019, the BRIEF-P has been used in 250 studies published in peer-reviewed journals, which attests to its validity in measuring EF in preschool aged children (Greene et al., 2019). The measure demonstrates high internal consistency for parent reports (.80 - .95) and moderate test-retest reliability (.78 - .90). It also exhibits convergent and discriminant validity with other clinical symptom measures (Gioia et al., 2000a).

Day/night task. This Stroop like task was used as a measure of inhibitory control (Gerstadt et al., 1994). During the *Day/night* task children were seated facing the tester and told they would be playing a silly game. The child was shown cards with either a moon and stars on a dark background or a sun on a light background. The experimenter instructed the child that when they saw the moon they should say “day”, and when they saw the sun they should say “night.” After the correct completion of two practice trials, the child was presented with 16 cards and the proportion of correct trials were recorded. The few studies which have considered the reliability of this task have found that it exhibits internal (Chasiotis et al., 2006; Rhoades et al., 2009) and test-retest reliability (Thorell & Wåhlstedt, 2006). ICC was calculated for 20% of the sample and was .96.

DCCS. The *Dimensional Change Card Sort* (DCCS) task was used as a measure of cognitive flexibility. It requires the child to sort cards by one specific dimension and then another, which is incompatible with the first. Children began in the pre- and post- switch conditions, if they correctly completed 5 of the 6 post-switch trials they were also asked to complete the borders condition. In the first trials the experimenter presented the child with 7 cards with a red or blue car or flower. The child was instructed to sort the cards by a shape or color, after they mastered the first rule they were told the rules were changed and they needed to sort by the opposite aspect. During the borders trial, children were given 12 cards and told to sort those with a border according to one dimension, but those without a border according to a different dimension. The proportion of correct trials were recorded. The task was used in over 150 studies from 1995-2015 and has been used widely as a measure of EF from ages 3-5 (Doebel & Zelazo, 2015). ICC was calculated for 19% of trials and was 1.00.

PPVT. The Peabody Picture Vocabulary Test (PPVT) is a short measure of a child's receptive vocabulary. The experimenter sat across the table from the child and spoke a word while revealing a page on the stimulus book. The child had to correctly identify which picture represents the word. Standard scores are derived based on the number of correct answers the child provided and their age. The age based internal consistency of the PPVT-4 for Form A is .89 - .97, alternate form reliability is .87 - .93, and test-retest reliability was .87 - .93 (Pearson, 2007).

9-Year visit

BRIEF 2. When the child returned at 9, parents were asked to complete the Behavior Rating Inventory of Executive Function, Second Edition. Similar to the BRIEF-P, the BRIEF assesses the executive function skills the parents observe their children exhibiting during day to day interactions. Scores on this version are separated into eight scales; Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. These specific scales form the Behavioral Regulation Index (BRI) and Metacognition Index (MI) and from a sum of these two indices the Global Executive Composite (GEC) is derived. The BRI is made up of the Inhibit scale, which measures inhibitory control and impulsivity, the Shift scale, which measures the ability to move from one problem to another as needed, and the Emotional control scale, which measures the ability control emotional responses. The MI is made up of the Initiate scale, which measures the ability to begin a task and independently generate responses, the Working Memory scale, which measures the ability to hold information and apply it to complete a task, the Plan/Organize scale, which measures the ability to anticipate future events and implement steps ahead to complete a task, the Organization of Materials scale, which measures orderliness of play, work, and storage space, and the Monitor scale, which measures task monitoring and self-monitoring. The same scoring procedures followed in interpreting the BRIEF-P were followed in the interpretation of the BRIEF. The BRIEF has established convergent and divergent validity, as well as high internal consistency (0.8 – 0.98) and test-retest reliability on the parent form (0.82; Gioia, 2000).

Number Stroop. The child was seated in front of the computer and followed directions given by the experimenter. The task began by asking the child to count and enter the number of letters on the screen using the number key on a keyboard. After two practice trials, the child was instructed to "go as fast as you can, but get as many right as you can." In the next condition the child was told there would now be numbers on the screen, but all other rules remained the same. Two practice trials were again given. In the mixed condition the child was presented with both numbers and letters on the screen, but instructed to only count the number of digits. Reaction time during the mixed condition was recorded. Relatively few studies using computerized Stroop tests report reliability and validity data, but a recent meta-analysis supported the reliability and validity of this approach (Din & Tat Meng, 2019).

WCST. The *Wisconsin Card Sort Task* (WCST) was also administered on a computer. The child was presented with 64 cards and instructed to sort according to either color, shape, or number. The experimenter informed the child that the computer would alert them to whether cards were sorted correctly or incorrectly. The child was then instructed to correctly sort the cards as quickly as possible, being unaware that the sorting rule changed many times throughout the task.

WJ Achievement. Subtests of the Woodcock Johnson Test of Achievement Third Edition (WJ- Ach) were administered to measure academic performance. Reading performance was measured with the Reading Fluency and Passage Comprehension subtests. Mathematics performance was measured with the Calculation and Math Fluency subtests. Scores on the Math and Reading subtests were entered as raw scores into a latent variable which considers academic ability and fluency. A literature review on the test found the reliability and validity to be acceptable and supported its use (Abu-Hamour et al., 2012).

PPVT. Same as above.

Results

All analyses were performed in RStudio Version 1.2.5033 using the following packages from May 2020 to May 2021.

Table 1
RStudio packages used during analysis

Package	Description
haven (Wickham & Miller, 2020)	Import SAS data file
lavaan (Rosseel, 2012)	All structural equation modeling
psych (Revelle, 2019)	Descriptive statistics and moderated regression analysis
readexl (Wickham & Bryan, 2019)	Import Excel data file
tidyverse (Wickham et al., 2019)	Attrition analysis.
dplyr (Wickham et al., 2020)	Attrition analysis and correlations.
Hmisc (Harrel et al., 2020)	Create boxplots to visualize outliers and correlation tables.

Attrition Analysis

Two-hundred and fifty-eight infants took part in 10-month visit. There was 35% attrition from the 10-month visit to the 48-month visit and another 9% did not return to the study at 9-years. There was an overall attrition rate of 43% from the 10-month visit to the 9-year visit. The most widely reported reason for not returning to the study was moving away. To investigate the systemic differences between those children who remained in the study and those who did not, the sex, race, and maternal education levels were compared (Table 1). I was unable to run Little's Missing Completely at Random test because the package is no longer available through R, but percentage representation of each group stayed fairly consistent at all ages.

Table 2
Demographics of sample at each time point

	10 months	48 months	9 years
Boys	131 (51%)	93 (51%)	89 (53%)
African American/ Black	38 (15%)	31 (17%)	32 (19%)
White	200 (78%)	139 (76%)	120 (72%)
Other	11 (4%)	6 (3%)	9 (5%)
Multiracial	7 (3%)	4 (2%)	4 (2%)
Did not report race	2 (1%)	2 (1%)	2 (1%)
Mother did not finish high school	6 (2%)	4 (2%)	5 (3%)
Mother finished high school	65 (25%)	46 (25%)	43 (26%)
Mother finished college or technical school	124 (48%)	85 (47%)	78 (47%)
Mother finished graduate degree	57 (22%)	41 (23%)	36 (22%)
Mother did not report education	6 (2%)	6 (3%)	5 (3%)
N	258	182	167

Data Screening

Scores which were more than three standard deviations from the mean were identified by running boxplots for each variable, and determined to be true representation of children's responses. Skewness and kurtosis was calculated to determine the normality of all variables (Table 2). Responses on the Surgency scale of the IBQ completed by parents at 10 months was slightly more skewed (skew = 2.18) than the conventional accepted value. A log transformation of the variable was attempted, but significantly increased the leptokurtosis; therefore the original variable was retained.

Although the skew and kurtosis was within acceptable ranges, some problems emerged during analyses due to the vast differences in the standard deviations of variables. To address this challenge, the Stroop, WCST, BRIEF-P ISCI, BRIEF-P FI, BRIEF-P EMI, BRIEF Behavior Regulation, and BRIEF Metacognition variables were transformed. All above variables were set on a similar scale using the POMS method, which is recommended as an alternative to standardization appropriate to be used in longitudinal work (Moeller, 2015). Descriptive statistics of normalized variables are reported below (Table 3). Transformed variables were used for all analyses.

Table 3

Descriptive statistics of all measures

		N	Min	Max	Mean	(SD)	Kurtosis	Skewness
10 month measures	A-not-B	269	0	9	3.77	1.55	0.15	0.46
	Puppet	265	1.53	22.86	7.63	3.75	1.55	1.22
	MPA Boo	266	1.67	4	2.88	0.6	-0.76	0.28
	MPA Toys	266	1	4	2.46	0.53	0.2	0.55
	Negative Affectivity	267	1.9	5.09	3.45	0.6	0.05	-0.16
	Surgency	267	2.36	6.56	5.29	0.56	2.18	-0.66
48 month measures	Day/Night	173	0	1	0.75	0.19	0.90	-0.76
	DCCS post-switch	202	0	1	0.56	0.46	-1.84	-0.23
	DCCS borders	105	0.33	1	0.56	0.15	1.60	1.43
	BRIEF-P FI	216	35	86	50.32	9.84	0.54	0.78
	BRIE-P EMI	216	36	88	53.76	11.33	-0.89	0.31
	BRIEF-P ISCI	216	34	90	52.49	10.85	0.01	0.58
	BRIEF-P GEC	216	34	92	52.86	11.26	-0.34	0.45
9 year measures	Stroop	163	1234.18	4298.0	2290.64	590.56	0.03	0.63
	WCST	166	20	80	50.54	14.75	-0.51	0.56
	BRIEF BRI	205	35	84	51.28	10.40	0.15	0.71

BRIEF MI	205	31	80	52.34	10.36	-0.36	0.38
BRIEF GEC	205	32	83	52.05	10.22	0.04	0.52

Table 4
Descriptive statistics of normalized measures

		N	Min	Max	Mean	(SD)	Kurtosis	Skewness
48 month measures	BRIEF-P FI	216	0	1	0.3	0.2	0.45	0.77
	BRIE-P EMI	216	0	1	0.36	0.22	-0.97	0.26
	BRIEF-P ISCI	216	0	1	0.34	0.2	-0.15	0.51
9 year measures	Stroop	163	0	1	0.34	0.19	0.03	0.63
	WCST	166	0	1	0.51	0.25	-0.51	0.56
	BRIEF BRI	205	0	1	0.33	0.22	0.18	0.73
	BRIEF MI	205	0	1	0.39	0.23	-0.5	0.36

Control Variables

In order to consider the true effects of the variables of interest of EF, the correlations between the EF scores, and maternal education, and sex were examined (Table 5). Sex was significantly negatively correlated with the BRIEF-P scales of ISCI and EMI, as well as the BRIEF MI and BR indexes. This indicates that mothers reported a greater degree of executive dysfunction in boys. As this correlation did not appear in the task measures of EF, this may be due to gendered expectations of children reflected in the self-report. Maternal education was only related to DCCS, so will not be used in analyses. PPVT at 48 months was correlated with 48 month measures of DCCS and the ISCI and FI scales of the BRIEF-P. It was also related to the Stroop test, and the BR index of the BRIEF at age 9. PPVT at 9 years was correlated with DCCS and the Stroop test. PPVT variables will be used in analyses predicting academic achievement to isolate the impacts of EF. Sex will be used as a control in all further analyses.

Table 5

Correlations between tasks across time points.

		Basic Info		10 month tasks		48 month											
		Sex	Mom's Education	A-not-B	Puppet	Day/ Night	DCCS pre-switch (border)	DCCS post-switch	PPVT	ISCI	FI	EMI	Stroop	WCST	PPVT	BRI	MCI
Basic Info	Sex	1															
	Mom's Education	-0.11	1														
10 month tasks	A-not-B	0.07	0.07	1													
	Puppet	0.01	0.05	0.16**	1												
48 moth tasks	Day/Night	-0.02	0.04	-0.11	0.08	1											
	DCCS pre-switch (border)	-0.03	0.08	-0.10	-0.04	0.14	1										
	DCCS post-switch	0.06	0.34**	0.09	0.06	0.05	-0.04	1									
	PPVT	0.02	0.41**	0.03	0.02	0.08	0.33**	0.47**	1								
	ISCI	-0.16*	-0.06	-0.10	0.00	-0.10	0.10	-0.15*	-0.15*	1							
	FI	0.00	-0.01	-0.11	-0.04	-0.05	0.06	-0.18*	-0.17*	0.78**	1						
	EMI	-0.14*	-0.13	-0.06	-0.06	-0.04	0.02	-0.13	-0.13	0.79**	0.66**	1					
9 year tasks	Stroop	0.03	-0.12	-0.15	0.05	0.07	0.10	-0.33**	-0.26**	-0.06	-0.01	-0.01	1				
	WCST	0.10	0.03	0.13	0.10	0.07	0.03	-0.02	0.00	-0.06	-0.07	-0.06	-0.04	1			
	PPVT	-0.19*	0.35**	0.02	-0.06	0.12	0.25*	0.34**	0.55**	-0.04	-0.07	-0.07	-0.16*	0.10	1		
	BRI	-0.22**	-0.04	-0.09	-0.08	-0.01	0.02	-0.12	-0.19*	0.55**	0.50**	0.42**	-0.04	-0.05	-0.05	1	
	MCI	-0.28**	-0.11	-0.02	-0.01	0.01	-0.01	-0.18	-0.12	0.50**	0.37**	0.52**	0.03	0.05	0.01	0.71**	1

Note. * p < .05, ** p < .01

Structural Equation Modeling (SEM)

SEM combines the statistical methods of factor analysis and multiple regression analyses to analyze specific structural relationships. It is the preferred method to be used with longitudinal data because latent constructs can be considered over time. SEM is also useful in conceptualizing factors like EF which cannot be accurately measured using one simple task (MacCallum & Austin, 2000). This approach is superior to creating aggregate measures of EF based on all tasks at an age point, as not all tasks will equally contribute to the overall measure and variance may be shared among some measures (MacCallum & Austin, 2000).

Hypothesis I: Some measures of EF at 10 months, 48 months, and 9 years will load onto a latent factor.

Model 1. A model including all EF measures at each age did not converge after 3861 iterations.

10 Month Measures. *A-not-B* task and *Puppet* task

48 Month Measures. *Day/Night* task, *DCCS*, ISCI, FI, and EMI

9 Year Measures. *Stroop* task, *WCST*, BRI, and MI

Model 2. The *Day/Night* and *DCSS border prop* task at 48 months and the *WCST* and *Stroop* tasks at 9 years were not correlated with any other EF variables, so were removed from the model to improve the fit. This model fit better (x^2= 40.89, df = 17, p = .001, RMSEA = 0.07, CFI = 0.96), but not all factor loadings were significant.

Model 3. Modification indices suggested allowing the BRIEF EMI variable at 48 months to covary with the BRIEF MI variable at 9 years. This aligns theoretically both are measures of metacognition at different ages. This improved model fit (x^2 = 19.90, df = 16, p = .225, RMSEA = 0.03, CFI = .99), but the 10 month factor loading was still not significant.

Model 4. The *Puppet* task was removed to improve the factor loading of the 10 month variable. It is surprising that the tasks are not significant together, because they were correlated. This change slightly decreased overall model fit (x^2 = 15.24, df = 11, p = .000, RMSEA = 0.04, CFI = .99).

Model 5. The *A-not-B* task was removed to eliminate issues associated with the 10 month measures. The overall model fit slightly improved (x^2 = 12.29, df = 7, p = .092, RMSEA = 0.05, CFI = .99; Table 5, Figure 1). The measures of EF at 48 months included in the final model are the BRIEF-P FI, EMI, and ISCI, as well as the post switch condition of the *DCCS* task. The measures of EF at 9 years included in the final model are the BRIEF MI and BRI.

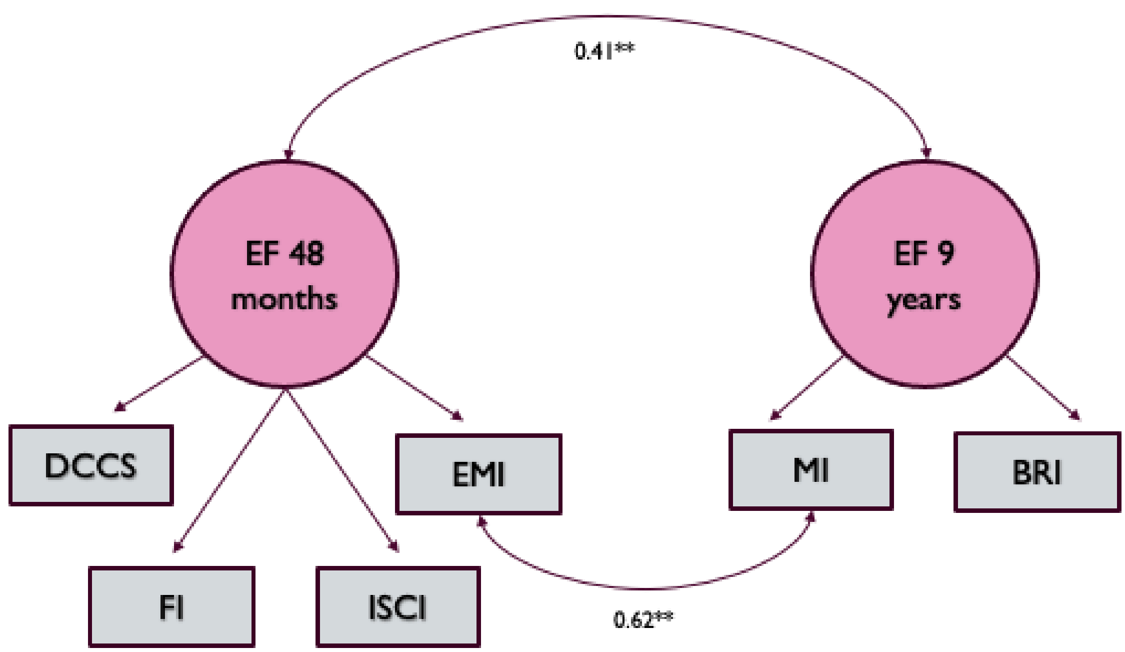

Figure 1. Model 5 final CFA model with standardized coefficients.

Table 6
CFA models.

Model 1 CFA
Did not converge

Model 2 CFA

		B	SE	*b*	*p*
10 month tasks					
	A-not-B	1.00			
	Puppet	1.17	2.25	0.52	.601
48 month tasks					
	ISCI	1.00			
	FI	0.85	0.05	16.36	.000
	EMI	0.95	0.06	16.67	.000
	DCCS post-switch	-0.40	0.18	-2.28	.022
9 year tasks					
	MCI	1.00			
	BRI	1.03	0.12	8.91	.000
Fit	x^2= 40.89, df = 17, p = .001, RMSEA = 0.07, CFI = 0.96				

Model 3 CFA

		B	SE	*b*	*p*
10 month tasks					
	A-not-B	1.00			
	Puppet	1.18	2.14	0.55	.582
48 month tasks					
	ISCI	1.00			
	FI	0.86	0.05	16.17	.000
	EMI	0.95	0.06	16.70	.000
	DCCS post-switch	-0.40	0.18	-2.26	.024
9 year tasks					
	MCI	1.00			
	BRI	1.10	0.13	8.46	.000
Fit	x^2 = 19.90, df = 16, p = .225, RMSEA = 0.03, CFI = .99				

Model 4 CFA

		B	SE	*b*	*p*
10 month tasks					
	A-not-B	1.00			
48 month tasks					
	ISCI	1.00			
	FI	0.86	0.05	16.16	.000
	EMI	0.95	0.06	16.70	.000
	DCCS post-switch	-0.40	0.18	-2.25	.024
9 year tasks					
	MCI	1.00			
	BRI	1.10	0.13	8.53	.000
Fit	x^2 = 15.24, df = 11, p = .000, RMSEA = 0.04, CFI = .99				

Model 5 CFA

		B	SE	*b*	*p*
48 month tasks					
	ISCI	1.00			
	FI	0.86	0.05	16.14	.000
	EMI	0.95	0.06	16.70	.000
	DCCS post-switch	-0.40	0.18	-2.25	.025
9 year tasks					
	MCI	1.00			
	BRI	1.09	0.13	8.70	.000
Fit	x^2 = 12.29, df = 7, p = .092, RMSEA = 0.05, CFI = .99				

Hypothesis II: The main effect of infant temperament and maternal positive affect, (MPA) as well as prior measures of EF, will predict EF at each time point. The impact of infant temperament and MPA will decrease over time.

a. Temperament and MPA at 10 months will predict EF at 10 months.

There are no significant correlations between measures of EF at 10 months and these factors, therefore a 10 month factor of EF was not included in further analyses. This hypothesis was not supported.

b. EF, temperament, and MPA at 10 months, will predict EF abilities at 48 months. Because there were no significant correlations between EF at 10 months and other factors, temperament and MPA were used to predict EF at 48 months. Maternal positive affect and infant temperament were allowed to covary, based on my hypothesis of their reciprocal relationship. The model fit well, x^2 = 24.8, df = 18, p = .130, RMSEA = 0.04, CFI = .99. Infant negative affect was a significant predictor of EF at 48 months and there was a significant relationship between infant negative affect and maternal positive affect, although maternal positive affect was not related to EF. This hypothesis was partially supported.

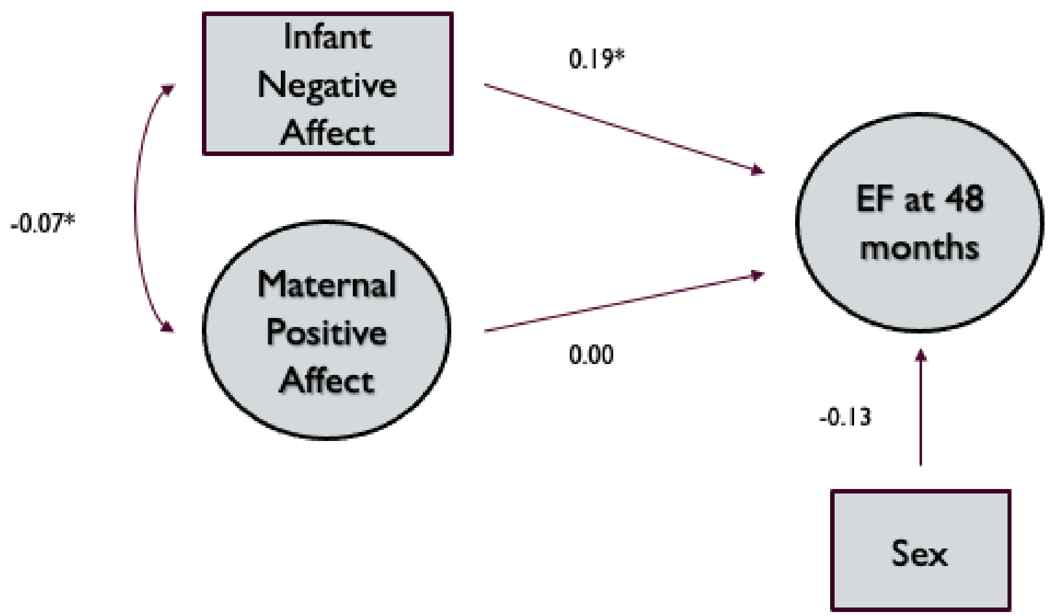

Figure 2. Model of hypothesis II-b with standardized coefficients.

c. Prior measures of EF, and temperament and MPA at 10 months, will predict EF abilities at 9 years.

As discussed above, measures of EF from 10 months were removed from the models. Maternal positive affect and infant temperament were allowed to covary, based on my hypothesis of their reciprocal relationship. The Emergent Metacognition Index on the BRIEF-P and the Metacognition Index were allowed to covary, as they were in the CFA. The model fit well, x^2 = 47.21, df = 31, p = .031, RMSEA = 0.04, CFI = .98. Infant negative affect was a significant predictor of EF at 48 months, which then predicted EF at 9 years. There were significant relationships between infant negative affect and maternal positive affect, although maternal positive affect was not a significant predictor of EF. This hypothesis was partially supported.

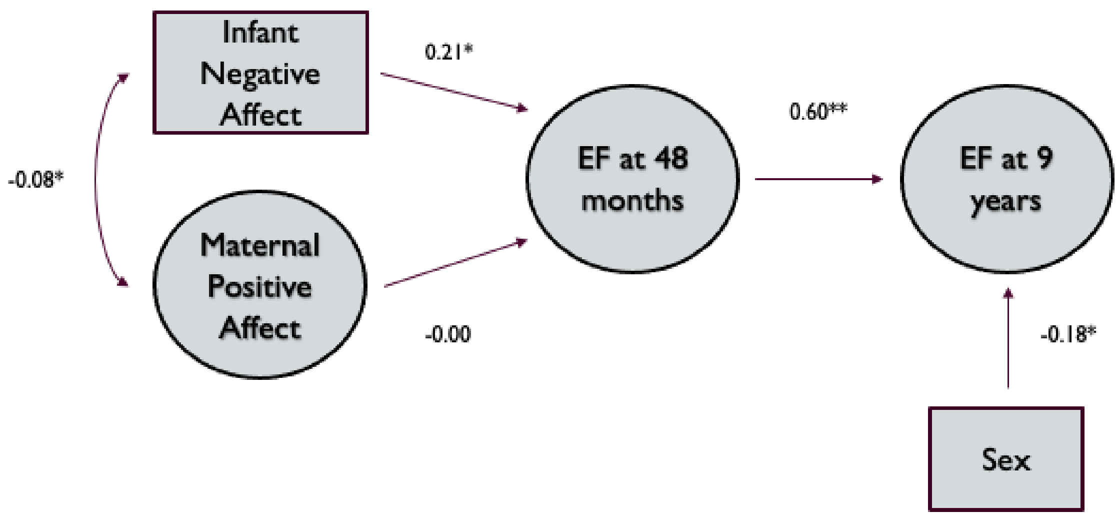

Figure 3. Model of hypothesis II-c with standardized coefficients.

Hypothesis III: The interaction between temperament and maternal positive affect will impact EF at 10 months.

a. Infants with easy or difficult temperaments, who experience high levels of maternal positive affect will have higher EF scores.
b. Infants with easy or difficult temperaments, who experience low levels of maternal positive affect will have lower EF scores.

This hypothesis was tested using a moderation analyses to test the potential impact of maternal positive affect on the relationship between infant negative affect and EF at 48 months. 48-month EF was used because it was related to infant negative affectivity in previous models. Maternal positive affect was examined as a moderator of the relation between infant negative affect and EF skills. Infant negative affect and maternal positive affect were entered in the first step of the regression analysis. In the second step of the regression analysis, the interaction term between infant negative affect and maternal positive affect was entered, but it did not explain a significant difference in later EF abilities. This hypothesis was not supported.

Hypothesis IV: EF and verbal ability at 48 months and 9 years will predict the relationship between EF at 10 months and academic achievement at 9 years.

Measures of EF at 10 months were not included in this model, as mentioned previously. Modification indices suggested allowing the measure of reading fluency and the measure of math fluency to covary, and this was accepted as both may be indicators of underlying processing speed abilities (Rohde & Thomas, 2007). The model fit moderately well, x^2 = 147.68, df = 57, p = .000, RMSEA = 0.09, CFI = .90. Verbal ability measures were positively and significantly associated with math and reading abilities, but executive function was not. Therefore this hypothesis is not supported.

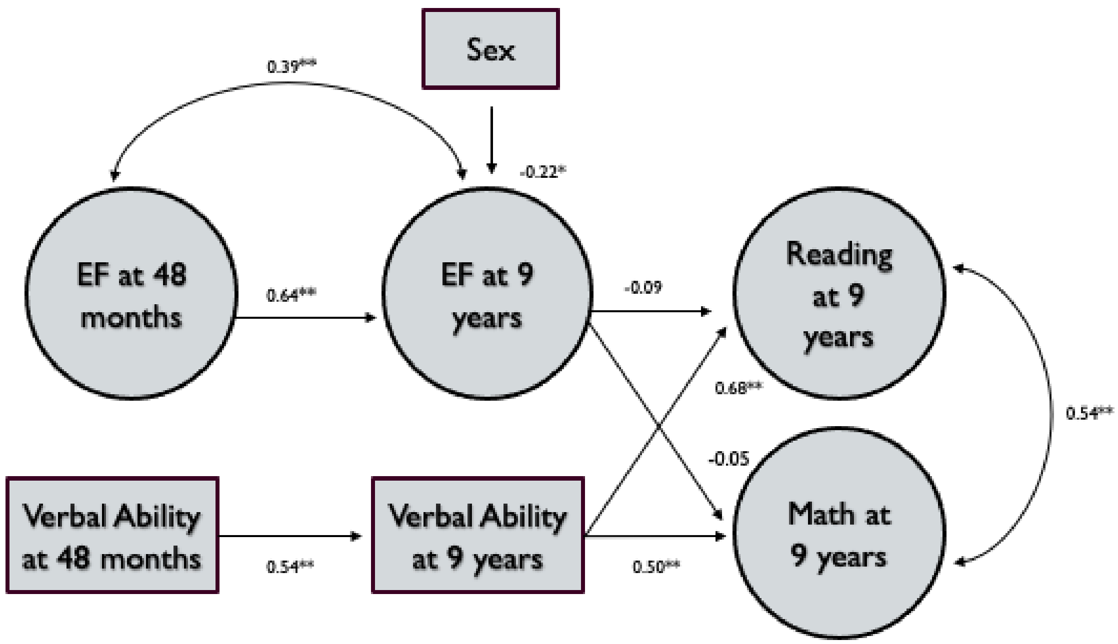

Figure 4. Model of hypothesis IV with standardized coefficients.

Post Hoc Analyses

I was intrigued by the lack of support for Hypothesis IV and wanted to further investigate the relationship of academic achievement and EF by separating task and survey measures of EF. When testing Hypothesis IV using the measures of EF which fit in the CFA, I noticed that modification indices indicated allowing DCCS, the only task measure of EF, to predict reading and math ability. Based on this and the literature indicating both shared and unique predictive value of survey and task measures of EF (Dekker et al., 2017), I decided to create two new path models similar to Hypothesis IV, but considering only task measures or survey measures of EF.

Model 1: This model included only task measures of EF. I attempted to fit a CFA model using the task measures only, but it did not fit well. I chose the DCCS task as the measure of EF at 48 months and the Stroop Number task as the measure of EF at 9 years, as these were the only task measures of EF which were correlated across ages. The model fit moderately well, x^2 = 58.30, df = 20, p = .000, RMSEA = 0.10, CFI = .88, and supported EF as a significant predictor of reading and math ability.

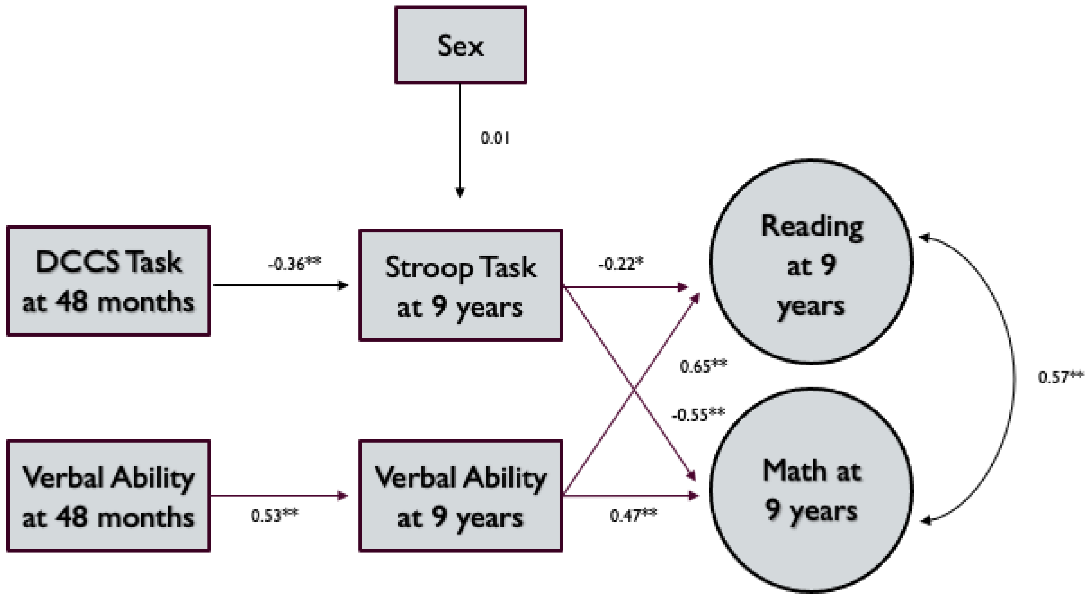

Figure 5. Post hoc model with only task measures of EF with standardized coefficients.

Model 2: This model only included parental report measures of EF from the BRIEF-P and BRIEF. The model fit well, x^2 = 85.41, df = 46, p = .000, RMSEA = 0.07, CFI = .95, but EF was not a significant predictor of reading or math abilities.

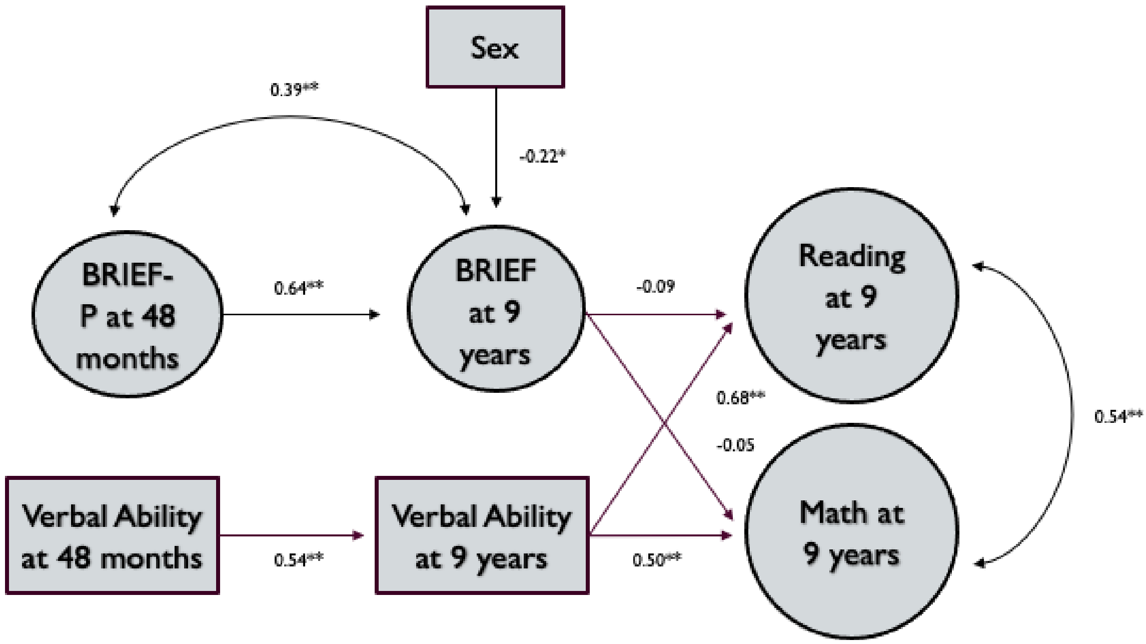

Figure 6. Post hoc model with only survey measures of EF with standardized coefficients.

Discussion

I tested several hypotheses related to predicting EF development from other factors, and then using different measures of EF to predict academic achievement. By considering infant negative affect and creating latent factors of 48 month and 9 year EF, I showed that infant negative affect predicted EF at 48 months, and EF at 48 months predicted EF at 9 years. Post hoc analysis showed significant paths from EF to math and reading achievement when using task measures of EF, but not when using parent report. Although few hypotheses were supported, I demonstrated that infant temperament impacts EF development and EF impacts academic achievement. I also highlighted the importance of considering which EF measurement tools are used in understanding the relationship between EF and academic achievement.

Hypothesis I: Some measures of EF at 10 months, 48 months, and 9 years will load onto a latent factor. This hypothesis was supported.

Although the first models failed to converge, a model of EF was created with statistically and theoretically related measures. Initial analyses revealed that behavioral and rating scale measures at each age point were not significantly correlated, except for the BRIEF-P FI and ISCI and the *DCCS* post switch measure. Rating scale measures from the BRIEF-P and BRIEF appeared to fit the CFA model best, so I chose to retain them and the *DCCS* post switch variable in creating the final model. Although I hypothesized that the *Day/Night* tasks and scores on *DCCS* would be related based on work by Broomell (2019) and Kraybill and Bell (2013), there are differences in the methods applied in these studies compared to the current approach. Broomell (2019) considered both tasks in the analysis, but also included many other task measures of inhibitory control in creating a latent factor of inhibitory control at 48 months. Differences in results may be due to only considering two task measures, thus limiting the model power and giving a less comprehensive measure of abilities. The decision in this study to include only two task measures of EF at each time point was made in order to focus the scope of the project and further investigate the use of both self-report and task measures of EF. It also could be impacted by using two task measures of EF that primarily measure inhibitory control, while using BRIEF-P measures scales of inhibitory self-control, flexibility, and emergent metacognition. The tasks may have fit the model better if only the inhibitory self-control scale was included. Kraybill and Bell used the *DCCS*, *Yes/No,* and *Pig/Bull* 48-month tasks as a measure of EF in their work (2013). The *Yes/No* task is thought to be conceptually very similar to the *Day/Night* task included in this analysis. Kraybill and Bell found all the tasks to be intercorrelated and combined the three tasks to create a 4-year-EF composite (2013). Although the *Day/Night* and *Yes/No* tasks are considered to be similar, it is possible that the differential task expectations measure slightly different abilities. The choice to include the *WCST* and *Stroop*

Number tasks were again based on Broomell (2019). Similar to my analysis, the *WCST* task did not fit this model well and was not retained in the CFA, but *Stroop* was included in the final model. In Broomell's work these measures were used as the only predictors of EF at age 9, so differences may be due to the inclusion of BRIEF scales.

The second CFA model with the *Day/Night* and *DCSS border prop* task at 48 months and the *WCST* and *Stroop* tasks at 9 years removed converged but did not fit well. Modification indices were run to determine how to improve the overall model fit. Results suggested including a correlation between the BRIEF-P *EMI* and the BRIEF *MI*. This is theoretically supported as both are measures of metacognition at different age points. This change improved the fit on Model 3, but the 10-month factor loading remained insignificant. The *Puppet* task did not appear to fit well in the model and was removed in Model 4. Theoretically, the two 10 month tasks were chosen based on previous work which used each task as measures of EF at this age (Cuevas & Bell, 2014; Blankenship et al., 2019). It is somewhat surprising that the two measures did not have a significant factor loading when creating a latent factor, considering they were significantly correlated in initial analyses. The *Puppet* task may not have fit the model due to differences in the approach Cuevas and Bell took by dichotomizing the variable in their work, separating children into short and long looker, versus using the times as a continuous variable (2014). Preserving continuous variables is preferable in maximum likelihood estimation (Flora & Curran, 2004). Another factor which may have impacted the relationship between the *Puppet* task and *A-not-B* task is the difference in the underlying ability being measured. Although the *Puppet* task was supported by Cuevas and Bell (2014) as a measure of EF, infants may be relying primarily on the orienting network to sustain visual attention (Colombo & Cheatham, 2006).

Therefore the *Puppet* task may be a better measure of the orienting network, and not the executive network or inhibitory control, as measured by the *A-not-B* task (Broomell, 2019).

Removing the *Puppet* task resulted in the 10-month factor becoming significant, but did not improve model fit. In the fifth iteration of the CFA model, the *A-not-B* task was also removed, because it did not fit further models and the overall model fit slightly improved. Infant executive function is a difficult construct to measure accurately and many studies that consider it have used multiple task measures of EF (Devine et al., 2019) or a battery of EF tasks and brain electrical activity (Broomell, 2019; Wolfe & Bell, 2007). The *A-not-B* task was successfully used as the only 10-month EF predictor of later academic achievement in another investigation, although 5 month attention was also included in the path model and only task measures of EF were included (Blakenship et al., 2019). The mixed measurement approach and reliance on parent measures of EF in the CFA may have misconstrued any relationship between 10 month EF and 48 month EF, thus no significant path relationships from 10 month EF were observed and it was removed from all models to improve parsimony. Including alternative measures of executive function in infancy may have increased the likelihood that early EF measures would significantly predict later outcomes. Alternative measures of EF in this time period including response inhibition tasks in which the child is instructed not to reach for a visually attractive toy, working memory tasks in which children search for hidden toys in different locations, and flexibility tasks which require sorting according to changing rules (Miller & Marcovitch, 2015). It should be noted that previous work (Wolfe & Bell, 2007) similarly did not support the relationship between measures of working memory and inhibitory control at 8 months and 4.5 years. The researchers contributed this to potential mismatch between cognitive measures at different time points, which may also be impacting the outcome of this study.

Hypothesis II: The main effect of infant temperament and maternal positive affect, (MPA) as well as prior measures of EF, will predict EF at each time point. The impact of infant temperament and MPA will decrease over time.

A. **Temperament and MPA at 10 months will predict EF at 10 months. This hypothesis was not supported.**

As discussed above, the 10-month measures of EF did not fit the model well and were removed from the CFA. Despite this, regression analyses were performed in an attempt to fit this model, but no significant relationships were found between EF measures at 10 months and maternal positive affect or infant temperament.

B. **EF, temperament and MPA at 10 months, will predict EF abilities at 48 months. This hypothesis was partially supported.**

After removing the 10-month measures of EF, there was a significant path from infant negative affect to EF at 48 months. The direction of this relationship indicates that infants with higher parental ratings of negative affect at 10 months, had higher reports of EF difficulties at 48 months. This finding contributes to an expanding literature basis, supporting the negative relationship between negative affectivity in infancy and effortful control and executive function in preschool. Parental positive affect was not significantly associated with EF abilities at any age point. Although previous work in this area supported the relationship between parental positive affect at 10 months and EF abilities at 4 years, the difference may be explained by the consideration of primarily rating scale measures of EF rather than behavioral measures (Kraybill & Bell, 2013). Although the previous study found a predictive relationship between maternal positive affect and the BRIEF-P measure of Post-Kindergarten EF, the sample size was considerably smaller (n = 56). Maternal positive affect was also measured only during a structured play task with a toy, whereas in this study a toy task and a peek-a-boo task were considered (Kraybill & Bell, 2013). Despite both behavioral and rating scales having internal consistency, there is some evidence that scores between them are not always correlated (Toplak et al., 2013). Although maternal positive affect did not independently impact EF, there was a significant negative relationship between maternal positive affect and infant negative affect. This indicates that these two variables are related, such that infants in dyads with higher ratings of maternal positive affect had lower ratings of negative affect.

C. Prior measures of EF, and temperament and MPA at 10 months, will predict EF abilities at 9 years. This hypothesis was supported.

When measures of 9-year EF abilities were added into the model, the relationship between infant negative affectivity and EF at 48 months remained significant. Furthermore, EF abilities at 48 months significantly predicted EF abilities at 9 years, exhibiting the strong predictive power of earlier measures of EF. These findings support the impact of differences in infancy, especially in temperamental style, shaping later cognitive function. The negative relationship between EF at 9 years and sex indicates that parent reported higher levels of EF issues in boys than girls.

Hypothesis III: The interaction between temperament and maternal positive affect will impact EF at 10 months.

a. Infants with easy or difficult temperaments, who experience high levels of maternal positive affect will have higher EF scores.

b. Infants with easy or difficult temperaments, who experience low levels of maternal positive affect will have lower EF scores.

This hypothesis was not supported.

This hypothesis was an attempt to further understand the nature of the relationships between maternal positive affect and infant negative affectivity, and EF. Based on the Differential Susceptibility model (Belsky & Pleuss, 2009), I hypothesized that not only children with easy temperaments would benefit from high levels of maternal positive affect, but also those with difficult temperaments. For this analysis I defined difficult temperament as high ratings of infant negative affectivity. Because 10-month measures of EF were not reliable in previous models, and a relationship was found between measures of 48-month EF and infant negative affectivity, this model was run using 48-month measures of EF. A moderated regression analysis did not indicate significant relationships between maternal positive affect, infant negative affect, or the interaction between the two, with EF. Because no relationship was found between EF and measures and maternal positive affect in previous analyses, it is not surprising that it did not function as a moderator of the relationship between infant negative affectivity and EF at 48 months.

Hypothesis IV: EF and verbal ability at 48 months and 9 years will predict the relationship between EF at 10 months and academic achievement at 9 years. This model was partially supported.

I chose a model that considered the impacts of both verbal ability and EF in predicting academic achievement to better understand the true impacts of EF. Because 10-month measures of EF were not related to later measures, it was not included in the model. The model ran but did not fit very well. The path from Verbal ability at 48 months to Verbal Ability at 9 years and to math and reading achievement was significant. In other words, higher Verbal Ability scores at 48 months predicted higher Verbal Ability scores at 9 years, which predicter greater academic achievement in both reading and math. The relationship between EF and academic achievement was not significant for math or reading, although EF ability at 48 months did predict EF abilities at 9 years. Sex was again negatively related to EF at 9 years, indicating that parents reported greater executive dysfunction in boys. Not discovering a significant path from EF abilities to academic achievement may be due to differences in EF measurement. My CFA included primarily BRIEF scales from parent reports, whereas many other studies used task measures of EF. I further investigated this through a post-hoc analysis.

Post Hoc Analysis

In the Post Hoc Analysis I maintained the two-path model of predicting academic achievement using EF and Verbal Ability, but chose to create separate models using only task or parent report measures of EF. The separate models highlighted important differences in the underlying concepts of EF being measured. The task based model showed significant paths of EF and Verbal Ability predicting academic achievement, while in the survey model only verbal ability was a significant predictor of math and reading ability. Sex was significantly related to EF abilities in the model using survey measures of EF, but not the model using task measures of EF. Parent perceptions of behavioral difficulties may have impacted their ratings of EF, which could be driving the differences observed across measures of EF. Another reason that task measures of EF may have been better predictors of academic achievement is due to behavior differences in children across environments. Ratings on the BRIEF reflect a parent’s perception of their child’s skills, so a teacher BRIEF report may have more accurately indicated a child’s academic ability.

Contributions and Limitations

This study was novel in considering both internal and external factors which impact executive functioning ability longitudinally, although only infant negative affectivity was significant. It also showed the path from infant negative affectivity to later EF abilities, and then the path from earlier EF abilities and verbal ability to reading and math achievement. It also highlighted the impact different measurement approaches of EF may have on findings. This contributes to the literature supporting the complexity of EF development and the many factors which may impact it during various stages, as well as supporting the connection between EF and academic abilities.

The current study was also faced with many limitations. Using a previously collected longitudinal data set allowed me to conduct many analyses with variables of interest, but also restricted what could be investigated. Data on maternal positive affect was only collected during one visit during two short tasks. The EF tasks considered may also measure primarily inhibitory control, and not working memory or cognitive flexibility. In retrospect, it may have been useful to include measures of early EF collected at a later time point, such as 24 months, for a more stable early measure that could have been included in all models. One recent work identified that individual trajectories of inhibitory control stabilized at 24 months (Broomell, 2019), thus measures at this time point may be better predictors of later life outcomes.

Future Directions

Future work should be done which considers both child and parenting variables and their interaction in predicting EF development. These studies will include measures of infant temperament and childhood temperament at multiple time points and more parenting variables. These may include intrusiveness, mind-mindedness, and guided learning. Researchers should consider how the interaction between parenting and temperament variables impact each other and therefore EF development. Additionally, more work should be done to understand the impacts of EF on academic achievement. Further investigation is needed to understand how differences in EF measurement may be impacting the relationships between EF and academic achievement, and which underlying constructs of EF may be differentially evaluated across measurement approaches. Exploration should also occur in determining how early indicators of executive dysfunction present opportunities for early interventions. EF deficits may be able to be detected earlier than deficits in early academic skills and targeting these areas through intervention at this point could have important lasting academic benefits.

Applications

This investigation contributes to the current discussion in school psychology regarding the impact of EF skills on academic success. Recent work has supported the potential for EF as a target area for intervention in closing the achievement gap (Zelazo & Carlson, 2020). Identifying early indicators of potential EF dysfunction provides the opportunity to begin early interventions to address future behavioral or academic challenges, and infant temperament could be useful as one indicator of risk. Developmentally appropriate play-based interventions beginning in preschool have been found to improve EF, as well as academic performance (Rosas et al., 2019), Access to these early interventions may have a lasting impact on students' academic careers. In order to fully understand the impacts of EF and EF based interventions on school success, it is essential to have a reliable measure of EF which can be feasibly administered within the school context. For the ease of administration on large scale, many schools may turn to rating scale measures of EF. The current work highlighted potential concerns in using only parent report measures of EF in considering the impact of EF on academic achievement and should serve as a warning in the exclusive use of survey measures of EF.

Summary and Conclusions

I was successful in demonstrating how rating scale measures and some behavioral measures could be used concurrently in creating latent factors of EF at different ages. I exhibited the impact of infant temperament on EF development and many measures of preschool EF to be predictive of later elementary EF. I found that an interaction between maternal affect and infant negative affectivity did not moderate the relationship between infant negative affectivity and EF. Additionally, I showed that only verbal ability predicted academic achievement, not EF, when using rating scale measures of EF. Lastly, I presented a model predicting academic achievement with independent contributions from both verbal ability and EF, when using task measures of EF.

REFERENCES

Abu-Hamour, B., Hmouz, H. A., Mattar, J., & Muhaidat, M. (2012). The Use of Woodcock-Johnson Tests for Identifying Students with Special Needs-a Comprehensive Literature Review. *Procedia - Social and Behavioral Sciences*, *47*, 665–673. https://doi.org/10.1016/j.sbspro.2012.06.714

Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Developmental Psychology, 50*, 2368 –2379. http://dx.doi.org/10.1037/ a0037493

Barkley R. A. (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology review*, *11*(1), 1–29. https://doi.org/10.1023/a:1009085417776

Bathelt, J., Holmes, J., Astle, D. E., Holmes, J., Gathercole, S., Astle, D., Manly, T., & Kievit, R. (2018). Data-driven subtyping of executive function–related behavioral problems in children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(4), 252-262.e4. https://doi.org/10.1016/j.jaac.2018.01.014

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, *135*(6), 885–908. https://doi.org/10.1037/a0017376

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

Blankenship, T. L., Slough, M. A., Calkins, S. D., Deater-Deckard, K., Kim-Spoon, J., & Bell, M. A. (2019). Attention and executive functioning in infancy: Links to childhood executive function and reading achievement. *Developmental Science*, *22*(6), e12824. https://doi.org/10.1111/desc.12824

Bornstein, M. H., Putnick, D. L., Gartstein, M. A., Hahn, C.-S., Auestad, N., & O'Connor, D. L. (2015). Infant temperament: Stability by age, gender, birth order, term status, and socioeconomic status. *Child Development*, *86*(3), 844–863. https://doi.org/10.1111/cdev.12367

Broomell, A. P. R. (2019). Development of Neuroconnectivity and Inhibitory Control: Relation to Social Cognition in Late Childhood. [Doctoral Dissertation, Virginia Polytechnic Institute and State University]. Electronic Thesis Depository

Broomell, A. P. R., & Bell, M. A. (2017). Inclusion of a Mixed Condition Makes the Day/Night Task More Analogous to the Adult Stroop. *Developmental Neuropsychology*, *42*(3), 241–252. https://doi.org/10.1080/87565641.2017.1309655

Calkins, S. D., Hungerford, A., & Dedmon, S. E. (2004). Mothers' interactions with temperamentally frustrated infants. *Infant Mental Health Journal, 25*, 219–239. https://psycnet.apa.org/doi/10.1002/imhj.20002

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool Children. *Developmental Neuropsychology*, *28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Center on the Developing Child at Harvard University (2011). Building the Brain's "Air Traffic Control" System: How Early Experiences Shape the Development of Executive Function: Working Paper No. 11. http://www.developing child.harvard.edu

Chasiotis, A., Kiessling, F., Hofer, J., & Campos, D. (2006). Theory of mind and inhibitory control in three cultures: Conflict inhibition predicts false belief understanding in Germany, Costa Rica and Cameroon. *International Journal of Behavioral Development*, *30*(3), 249–260. https://doi.org/10.1177/0165025406066759

Cheng, N., Lu, S., Archer, M., & Wang, Z. (2018). Quality of maternal parenting of 9-month-old infants predicts executive function performance at 2 and 3 years of age. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02293

Clearfield, M. W., Diedrich, F. J., Smith, L. B., & Thelen, E. (2006). Young infants reach correctly in A-not-B tasks: On the development of stability and perseveration. *Infant Behavior and Development*, *29*(3), 435–444. https://doi.org/10.1016/j.infbeh.2006.03.001

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior*, *34*, 283–322. https://doi.org/10.1016/S0065-2407(06)80010-8

Conway, A., & Stifter, C. A. (2012). Longitudinal Antecedents of Executive Function in Preschoolers. *Child Development*, *83*(3), 1022–1036. https://doi.org/10.1111/j.1467-8624.2012.01756.x

Cuevas, K., & Bell, M. A. (2014). Infant Attention and Early Childhood Executive Function. *Child Development*, *85*(2), 397–404. https://doi.org/10.1111/cdev.12126

Cuevas, K., Deater-Deckard, K., Kim-Spoon, J., Watson, A. J., Morasch, K. C., & Bell, M. A. (2014). What’s mom got to do with it? Contributions of maternal executive function and caregiving to the development of executive function across early childhood. *Developmental Science*, *17*(2), 224–238. https://doi.org/10.1111/desc.12073

Dekker, M., Ziermans, T., Spruijt, A., & Swaab, H. (2017). Cognitive, Parent and Teacher Rating Measures of Executive Functioning: Shared and Unique Influences on School Achievement. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00048

Devine, R. T., Ribner, A., & Hughes, C. (2019). Measuring and Predicting Individual Differences in Executive Functions at 14 Months: A Longitudinal Study. *Child Development*, *90*(5), e618–e636. https://doi.org/10.1111/cdev.13217

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases, of inhibitory control in reaching. Annals of the New York Academy of Sciences, 608, 637–676. https://doi.org/10.1111/j.1749-6632.1990.tb48913.x

Diamond, A. (2001). Looking closely at infants' performance and experimental procedures in the A-not-B task. *Behavioral and Brain Sciences*, *24*(1), 38–41. https://doi.org/10.1017/S0140525X01253916

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development, 62,* 1–205. https://psycnet.apa.org/doi/10.2307/1166208

Din, N.C. & Tat Meng, E.C. (2019). Computerized Stroop Tests: A Review. *Journal of Psychology and Psychotherapy, 9*(1), 5. doi:10.4172/2161-0487.1000353

Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children. *Developmental Review : DR*, *38*, 241–268. https://doi.org/10.1016/j.dr.2015.09.001

Duckworth, A. L., & Seligman, M. E. P. (2005).Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*, 939–944.http://dx.doi.org/10.1111/j.1467-9280.2005.01641.x

Engelhardt, L. E., Harden, K. P., Tucker-Drob, E. M., & Church, J. A. (2019). The neural architecture of executive functions is established by middle childhood. *NeuroImage; Amsterdam*, *185*, 479–489. http://dx.doi.org/10.1016/j.neuroimage.2018.10.024

Fay-Stammbach, T., Hawes, D. J., & Meredith, P. (2014). Parenting influences on executive function in early childhood: A review. *Child Development Perspectives*, *8*(4), 258–264. https://doi.org/10.1111/cdep.12095

Flora, D. B., & Curran, P. J. (2004). An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis With Ordinal Data. Psychological Methods, 9(4), 466–491. https://doi.org/10.1037/1082-989X.9.4.466

Frick, J. E., Colombo, J., & Saxon, T. F. (1999). Individual and Developmental Differences in Disengagement of Fixation in Early Infancy. *Child Development*, *70*(3), 537–548. https://doi.org/10.1111/1467-8624.00039

Frick, M. A., Forslund, T., Fransson, M., Johansson, M., Bohlin, G., & Brocki, K. C. (2018). The role of sustained attention, maternal sensitivity, and infant temperament in the development of early self-regulation. *British Journal of Psychology*, *109*(2), 277–298. https://doi.org/10.1111/bjop.12266

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning in preschoolers at head start. *Developmental Psychology*, *47(2)*, 404–416. http://dx.doi.org/10.1037/a0021065.

Gartstein, M.A. & Rothbart, M.K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behavior & Development.* ;26:64–86. https://doi.org/10.1016/S0163-6383(02)00169-8

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 312–7 years old on a stroop- like day-night test. *Cognition*, *53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Gioia, G. A., Epsy, K. A., Guy, S. C., & Isquith, P. K. (2000a). *Behavior Rating Inventory of Executive Function®- Preschool Version (BRIEF®-P).*

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000b). *Behavior Rating Inventory of Executive Function®, Second Edition (BRIEF®2)*. 2.

Graziano, P. A., Calkins, S. D., & Keane, S. P. (2011). Sustained attention development during the toddlerhood to preschool period: Associations with toddlers’ emotion regulation strategies and maternal behaviour. *Infant and Child Development*, *20*(6), 389–408. https://doi.org/10.1002/icd.731

Greene, J. A., Trujillo, S., Isquith, P. K., Gioia, G. A., Espy, K. A. (2019). *Enhanced Interpretation of the Behavior Rating Inventory of Executive Function–Preschool Version (BRIEF-P)* (white paper). Lutz, FL: PAR.

Grossman, A. W., Churchill, J. D., McKinney, B. C., Kodish, I. M., Otte, S. L., & Greenough, W. T. (2003). Experience effects on brain development: Possible contributions to psychopathology. *Journal of Child Psychology and Psychiatry*, *44*(1), 33–63. https://doi.org/10.1111/1469-7610.t01-1-00102

Harrell, F. E. Jr., Dupont, C., et al. (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-0. https://CRAN.R-project.org/package=Hmisc

Hodel, A. S. (2018). Rapid infant prefrontal cortex development and sensitivity to early environmental experience. *Developmental Review*, *48*, 113–144. https://doi.org/10.1016/j.dr.2018.02.003

Hughes, C., & Devine, R. T. (2019). For better or for worse? Positive and negative parental influences on young children's executive function. *Child Development*, *90*(2), 593–609. https://doi.org/10.1111/cdev.12915

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive functions to improve academic achievement: A review. *Review of Educational Research, 85,* 512-552. https://doi.org/10.3102%2F0034654314561338

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kim, S., & Kochanska, G. (2012). Child temperament moderates effects of parent–child mutuality on self-regulation: A relationship-based path for emotionally negative infants. *Child Development*, *83*(4), 1275–1289. https://doi.org/10.1111/j.1467-8624.2012.01778.x

Kraybill, J. H. & Bell, M. A. (2013). Infancy predictors of preschool and post-kindergarten executive function. *Developmental Psychobiology, 55*, 530-538. doi:10.1002/dev.21057

Liu, R., Blankenship, T. L., Broomell, A. P. R., Garcia-Meza, T., Calkins, S. D., & Bell, M. A. (2018). Executive function mediates the association between toddler negative affectivity and early academic achievement. *Early Education & Development,* 29(5), 641–654. https://doi.org/10.1080/10409289.2018.1446880

MacCallum, R. C., & Austin, J. T. (2000). Applications of Structural Equation Modeling in

Psychological Research. Annual Review of Psychology, 51(1), 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

Mcauley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function*? Journal of the International Neuropsychological Society*, *16*(3), 495–505. https://doi.org/10.1017/S1355617710000093

Miller, S. E., & Marcovitch, S. (2015). Examining executive function in the second year of life: Coherence, stability, and relations to joint attention and language. *Developmental Psychology*, *51*(1), 101–114. https://doi.org/10.1037/a0038359

Miranda, A., Colomer, C., Mercader, J., Fernández, M. I., & Presentación, M. J. (2015). Performance-based tests versus behavioral ratings in the assessment of executive functioning in preschoolers: Associations with ADHD symptoms and reading achievement. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00545

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moeller, J. (2015). A word on standardization in longitudinal studies: Don’t. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01389

Monette, S., Bigras, M., & Guay, M.-C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, *109*(2), 158–173. https://doi.org/10.1016/j.jecp.2011.01.008

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*(3), 308–330. https://doi.org/10.1016/j.dr.2010.07.001

Morgan, P.L., Farkas, G., Hillemeier, M.M., Pun, W.H. and Maczuga, S. (2019), Kindergarten Children's Executive Functions Predict Their Second-Grade Academic Achievement and Behavior. *Child Development*, *90:* 1802-1816. https://doi.org/10.1111/cdev.13095

Nelson, C. A., Ewing-Cobbs, L., Luciana, M., Angulo-Barroso, R. M., Aslin, R. N., Bardhan, N. P., Bauman, M. D., Clayards, M. A., Csibra, G., & Dannemiller, J. L. (2008). *Handbook of Developmental Cognitive Neuroscience*. MIT Press. http://ebookcentral.proquest.com/lib/hunter-ebooks/detail.action?docID=3338909

Cortés Pascual, A., Moyano Muñoz, N., & Robres, A.Q. (2019). The Relationship Between Executive Functions and Academic Performance in Primary Education: Review and Meta-Analysis. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01582

Parade, S. H., & Leerkes, E. M. (2008). The reliability and validity of the Infant Behavior Questionnaire-Revised. *Infant behavior & development*, *31*(4), 637–646. https://doi.org/10.1016/j.infbeh.2008.07.009

Pearson (2007). *Technical Information.* Retrieved from *https://support.pearson.com/usclinical/servlet/fileField?entityId=ka00N000000sibPQAQ&field=Attachments__Body__s*

Piaget, J. (1954). *The Construction of Reality in the Child*. Routledge & Kegan Paul.

Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development*, *17*(4), 387–405. https://doi.org/10.1002/icd.582

Razza, R. A., & Blair, C. (2009). Associations among false-belief understanding, executive function, and social competence: A longitudinal analysis. *Journal of Applied Developmental Psychology*, *30*(3), 332–343. https://doi.org/10.1016/j.appdev.2008.12.020

Revelle, W. (2019) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.9.12.

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social–emotional competence. *Journal of Applied Developmental Psychology*, *30*(3), 310–320. https://doi.org/10.1016/j.appdev.2008.12.012

Rimm-Kaufman, Sara & Pianta, Robert & Cox, Martha. (2000). Teachers' judgments of problems in the transition to school. *Early Childhood Research Quarterly. 15.* 147-166. https://doi.org/10.1016/S0885-2006(00)00049-1

Rochette, É., & Bernier, A. (2016). Parenting and preschoolers' executive functioning: A case of differential susceptibility? *International Journal of Behavioral Development*, *40*(2), 151–161. http://dx.doi.org.proxy195.nclive.org/10.1177/0165025414557370

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*(1), 83–92. https://doi.org/10.1016/j.intell.2006.05.004

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. URL http://www.jstatsoft.org/v48/i02/.

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, *52*, 569–578. https://doi.org/10.2307/1129176

Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science*, *16*(4), 207–212. https://doi.org/10.1111/j.1467-8721.2007.00505.x

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, *78*(1), 122–135. https://doi.org/10.1037/0022-3514.78.1.122

Rothbart, M. K., Ellis, L. K., & Posner, M. I. (2013). Temperament and self-regulation. In K. D. Vohs & R. F. Baumeister (Eds.), Handbook of self-regulation, second edition: Research, theory, and applications. New York: Guilford Press.

Rosas, R., Espinoza, V., Porflitt, F., & Ceric, F. (2019). Executive Functions Can Be Improved in Preschoolers Through Systematic Playing in Educational Settings: Evidence From a Longitudinal Study. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02024

Schmitz, N., Rubia, K., Daly, E., Smith, A., Williams, S., & Murphy, D. G. M. (2006). Neural correlates of executive function in autistic spectrum disorders. *Biological Psychiatry*, *59*(1), 7–16. https://doi.org/10.1016/j.biopsych.2005.06.007

Shallice, T., Marzocchi, G. M., Coser, S., Savio, M. D., Meuter, R. F., & Rumiati, R. I. (2002). Executive Function Profile of Children With Attention Deficit Hyperactivity Disorder. *Developmental Neuropsychology*, *21*(1), 43–71. https://doi.org/10.1207/S15326942DN2101_3

Soto, E. F., Kofler, M. J., Singh, L. J., Wells, E. L., Irwin, L. N., Groves, N. B., & Miller, C. E. (2020). Executive functioning rating scales: Ecologically valid or construct invalid? *Neuropsychology*, *34*(6), 605–619. https://doi.org/10.1037/neu0000681

Spiegel, J. A., Lonigan, C. J., & Phillips, B. M. (2017). Factor Structure and Utility of the Behavior Rating Inventory of Executive Function-Preschool Version. *Psychological Assessment*, *29*(2), 172–185. https://doi.org/10.1037/pas0000324

Suor, J. H., Sturge-Apple, M. L., Davies, P. T., & Jones-Gordils, H. R. (2019). The interplay between parenting and temperament in associations with children's executive function. *Journal of Family Psychology*, *33*(7), 841–850. https://doi.org/10.1037/fam0000558

Swingler, M. M., Perry, N. B., Calkins, S. D., & Bell, M. A. (2017). Maternal behavior predicts infant neurophysiological and behavioral attention processes in the first year. *Developmental Psychology*, *53*(1), 13–27. https://doi.org/10.1037/dev0000187

Thomas, A., & Chess, S. (1977). *Temperament and development*. Brunner/Mazel.

Thorell, L. B., & Wåhlstedt, C. (2006). Executive functioning deficits in relation to symptoms of ADHD and/or ODD in preschool children. *Infant and Child Development*, *15*(5), 503–518. https://doi.org/10.1002/icd.475

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, 54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Valcan, D. S., Davis, H., & Pino-Pasternak, D. (2018). Parental behaviours predicting early childhood executive functions: A meta-analysis. *Educational Psychology Review*, *30*(3), 607–649. https://doi.org/10.1007/s10648-017-9411-9

Watkins, L. H., Sahakian, B. J., Robertson, M. M., Veale, D. M., Rogers, R. D., Pickard, K. M., Aitken, M. R. F., & Robbins, T. W. (2005). Executive function in Tourette's syndrome and obsessive–compulsive disorder. *Psychological Medicine*, *35*(4), 571–582. https://doi.org/10.1017/S0033291704003691

Wolfe, C. D., & Bell, M. A. (2007). The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. *Brain and Cognition, 65*(1), 3–13. https://doi.org/10.1016/j.bandc.2006.01.009

Wickham et al., (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686, https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl

Wickham, H., Francois, R., Henry, L., & Müeller, K. (2020). dplyr: A Grammar of Data Manipulation. R package version .8.5. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Miller, E. (2020). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.3.1. https://CRAN.R-project.org/package=haven

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

Willoughby, M. T., Magnus, B., Vernon-Feagans, L., Blair, C. B., & the Family Life Project Investigators. (2017). Developmental delays in executive function from 3 to 5 years of age predict kindergarten academic readiness. *Journal of Learning Disabilities, 50,* 359 –372. http:// dx.doi.org/10.1177/0022219415619754

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children : Abstract : Nature Protocols. *Nat. Protocols*, *1*(1), 297– 301. https://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., & Carlson, S. M. (2020). The neurodevelopment of executive function skills: Implications for academic achievement gaps. *Psychology & Neuroscience*, *13*(3), 273. https://doi.org/10.1037/pne0000208

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., … Carlson, S. M. (2003). The Development of Executive Function in Early Childhood. *Monographs of the Society for Research in Child Development*, *68*(3), i–151. Retrieved from JSTOR.