Published in final edited form as:
Child Neuropsychol. 2017 May ; 23(4): 381–407. doi:10.1080/09297049.2015.1120860.

Cognitive and behavioral rating measures of executive function as predictors of academic outcomes in children

Elyssa H. Gerst, **Paul T. Cirino**, **Jack M. Fletcher**, and **Hanako Yoshida**
Department of Psychology, University of Houston, Texas, USA

Abstract

Interrelations of two measurement methods (cognitive versus behavioral ratings) for executive function (EF) were examined and related to reading comprehension and math calculations in fourth and fifth grade students ($n = 93$) in the context of a diverse urban student population. Relations among measures within four EF processes (working memory, planning, inhibition and shifting) were modest; relations to academics were stronger. EF measures contributed to both academic outcomes even in the context of relevant covariates (age, language and educational program). Working memory was particularly important for reading comprehension across measurement type. Cognitive measures from all EF processes, particularly inhibition and planning, and behavioral ratings of working memory were important for math.

Keywords

Executive function; academics; behavioral ratings; performance-based; assessment

The construct of executive function (EF) has been defined in multiple ways. A general description consists of the processes that enable engagement in goal-directed behavior, such as working memory, planning, inhibition and shifting (Lezak, Howieson, Bigler, & Tranel, 2012; Mahone et al., 2002). For children, executive dysfunction may negatively impact classroom and academic outcomes, as well as family and peer relationships (Anderson, 2002). However, we know less about which EF processes, measured in different ways, are predictive for achievement, at both absolute and relative levels, particularly within a model-driven framework. Executive dysfunction has been described in a number of neurological and neurodevelopmental disorders, but there is less research that focuses specifically on performance in populations facing more non-specific adversity. Understanding how the different processes of EF are associated with one another and with academic outcomes is informed by studies of both specific neurodevelopmental disorders and broader populations, and these two perspectives complement one another.

Therefore, the purpose of this study is to evaluate key EF processes, assessed via both cognitive and behavioral methods, in terms of their relation to one another, and towards the

CONTACT Elyssa H. Gerst egerst@central.uh.edu Department of Psychology, University of Houston, 4811 Calhoun Rd, Office 476, Houston, TX 77204, USA.

Disclosure statement
No potential conflict of interest was reported by the authors.

prediction of both reading and mathematical skills. The context for this study is an urban public school setting where students are diverse and frequently encounter linguistic and economic challenges. This context is important because children with limited socioeconomic resources and from minority populations may be considered as being at risk of experiencing executive dysfunction and/or academic failure (Finn & Rock, 1997; Noble, Farah, & McCandliss, 2006; Roy & Raver, 2014). The term “at risk” can also be used within school districts to mean being at risk of experiencing academic failure, with categories including students who have been retained or perform unsatisfactorily on state-mandated tests, or who are homeless. The limited number of studies examining populations where many students are at risk (e.g., low socioeconomic status [SES], free lunch, urban public schools) have found their EF and academic skills development to be worse than normative populations (Kishiyama, Boyce, Jimenez, Perry, & Knight, 2009; Noble et al., 2006; Waber, Gerber, Turcios, Wagner, & Forbes, 2006). Also, children of a minority population may have higher (more problematic) scores on teacher-rated measures of behavior than their Caucasian peers (Downey & Pribesh, 2004). Additionally, external influences, including SES, have been shown to impact the development of EF (Noble et al., 2006; Ursache, Blair, & Raver, 2012). This impairment is seen through the abnormal development of cortical areas key to EF, including areas within the prefrontal cortex (Farah et al., 2006; Hackman & Farah, 2009). Childhood poverty has been associated with differential outcomes for children from non-Caucasian (or minority) populations (Sirin, 2005). Children from minority populations have been found to have higher ratings on observable behavioral scales and worse performance on EF and academic measures than Caucasian children (Dwivedi & Banhatti, 2005; Epstein et al., 2005; Waber et al., 2006).

Below, we review salient prior work that motivates our hypotheses concerning the measurement of EF (the processes, the models from which they are derived, the means of assessing them, and how they relate to one another), as well as how EF relates to and is important for reading and mathematics achievement.

Measurement of EF

The two most common methods of assessing EF are through cognitive measures and behavioral rating scales. Cognitive measures of EF in children are similar to those used with adults. There are, however, several challenges in using cognitive measures to capture EF. First, the highly structured one-to-one setting, which has been hypothesized to “act as the examinee’s frontal lobe”, makes it more challenging to pick up on problems that are observed in less structured environments (Salthouse, Atkinson, & Berish, 2003). Second is the “task-impurity problem” (Burgess, Alderman, Evans, Emslie, & Wilson, 1998), where the target task includes variance that may be necessary but not sufficient for the EF process under study. Finally, novelty is a key characteristic of many EF tasks. In practice, the degree of novelty may not be the same across individuals, which can lead to the utilization of different strategies for the same task (Hughes & Graham, 2002; Walsh, 1978). Thus, it is not surprising that EF measures are weakly correlated among themselves over time, or with one another (Hughes & Graham, 2002; Miyake et al., 2000).

Behavioral rating scales are also used to assess EFs, and ameliorate some of the difficulties discussed above. Such measures are completed by observers (parents and/or teachers) and/or the child (depending on his or her age). Behavioral rating scales have been found to capture expected patterns of EF in different clinical populations, to correlate with biological markers associated with EF, and to produce evidence of relations between EF and real-world functioning (Isquith, Roth, & Gioia, 2013). However, behavioral rating scales may be susceptible to their own challenges, including rater differences, and the development of multiple reporter versions. Therefore, clinicians and researchers must rely on balancing multiple sources of information and reporter bias in making a diagnosis (Isquith et al., 2013). Some behavioral rating scales of EF provide only an overall total score which may not be sensitive enough to clarify certain types of observed behavior; for example, the Behavior Assessment System for Children – Second Edition (BASC-2; Reynolds & Kamphaus, 2004) offers a single Executive Functioning scale in the context of many content-area scales. An exception to this pattern is the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000), which has reliable subscales for a number of EF processes.

Previous studies examining the relation of cognitive measures with behavioral rating scales of EF have shown mixed results. Toplak, West, and Stanovich (2013) summarized much of this work in their review of 20 prior studies and found a median correlation of $r = .19$ across methods used in these studies. They concluded that the different types of measures capture different information: cognitive measures represent efficiency of performance in an optimal setting, whereas behavioral rating scales represent the frequency of goal achievement in typical settings. In some of these studies, global ratings of EF or cognitive EF composites were used to assess relations.

However, not all studies show weak relations. Shimoni, Engel-Yeger, and Tirosh (2012) found significant correlations between Behavioural Assessment of the Dysexecutive Syndrome in Children (BADS-C; Emslie, Wilson, Burden, Nimm-Smith, & Wilson, 2003) performance with BRIEF total score and subscores (Metacognition Index, and Emotional Control, Working Memory, Planning, Monitoring, and Inhibition subtests, range $|r| = .27$ to .47) in Israeli boys. Toplak, Bucciarelli, Jain, and Tannock (2008) examined four EF processes (inhibition, working memory, set shifting and planning) in an adolescent population with ADHD and controls. The BRIEF Teacher subscales of Inhibition and Working Memory were significantly correlated to their cognitive counterparts (Stop Signal Task, $r = .32$; Wechsler Adult Intelligence Scale-III (WAIS-III) Digit Span/Spatial Span, $r = .33$, respectively; Wechsler, 1997); shifting and planning measures did not significantly correlate across behavioral and cognitive domains. The stronger relations in these studies suggest that additional power may be gained by comparing specific rather than global measures of EF to behavioral rating scales. Alloway, Gathercole, Kirkwood, and Elliott (2009) found even stronger correlations between the listening recall and backwards digit recall measures of the author-developed battery Automated Working Memory Assessment (AWMA; Alloway, 2007) and the author-developed teacher rating scale Working Memory Rating Scale (WMRS; rs = −.57 and −.59, respectively; Alloway, 2008), though results of this magnitude are more an exception and would be bolstered by replication studies across populations or measurements.

In sum, cognitive measures and behavioral rating scales of EF are commonly used together in clinical settings to understand a child's current functional level, though most previous studies generally show weak interrelations. However, few studies evaluate the relations among specific processes of EF, where relations may be stronger.

Relation of executive function to academic skills

It is important to address the external correlates of EF because the ecological validity of EF measures is a recurrent issue (Burgess et al., 1998; Chaytor & Schmitter-Edgecombe, 2003). Children spend much of their time learning in school, so functional outcomes related to academic performance provide a reasonable target outcome by which to gauge the impact of EF on children. The EF processes studied here (working memory, planning, inhibition and shifting) have conceptual relevance for both reading comprehension and mathematics, and there is strong empirical support for the relation of EF to academic performance in children (Best, Miller, & Naglieri, 2011; Biederman et al., 2004; Waber et al., 2006).

Reading comprehension

By the late elementary school stage a child is expected to have mastered the basic skills of reading, like phonemic decoding and reading fluency, and begun to develop strategies for reading to learn (or reading comprehension; Sesma, Mahone, Levine, Eason, & Cutting, 2009). This developmental change is important as the requirements of the classroom and expectations of the child change, and students who were previously thought to be average readers may be identified as having challenges with reading comprehension (Biancarosa & Snow, 2004).

The EF processes described in this study all have logical relevance to reading comprehension. *Working memory* has been strongly linked to reading comprehension performance (Cain, Oakhill, & Bryant, 2004; Sesma et al., 2009; St Clair-Thompson & Gathercole, 2006) and is thought to allow students to hold previously read information while simultaneously accessing new information. As children begin to read to learn, strategies are used to gain information from passages, highlighting the conceptualized influence of *planning*, which has variable support for its contribution to reading comprehension (Locascio, Mahone, Eason, & Cutting, 2010; Sesma et al., 2009; Sikora, Haley, Edwards, & Butler, 2002). Students must also learn to focus on words and sentences that are relevant to the main topic and ignore additional information that is presented, thus demonstrating the role of *inhibition* in reading comprehension (Arrington, Kulesz, Francis, Fletcher, & Barnes, 2014). However, like planning, the contribution of inhibition to reading comprehension has variable support (Borella, Carretti, & Pelegrina, 2010; Christopher et al., 2012; St Clair-Thompson & Gathercole, 2006). Although there are few studies that demonstrate a strong link between shifting and reading comprehension, children may require *shifting* to move from one topic to another within a passage, or to integrate their background knowledge into text reading.

Mathematical calculations

The study of mathematics is less well-developed than that of reading, though the situation is rapidly changing (Barnes, Fuchs, & Ewing-Cobbs, 2009). Mathematic development is hierarchical, evolving from math fact mastery in the earliest grades to later computational and application proficiency (Fuchs et al., 2006). The literature is much deeper with regard to the cognitive concomitants of math fact mastery and computational skill, relative to other mathematical domains such as applications, in part due to diversity in the way the latter is measured. It is also true that by the fourth or fifth grade, many math facts have become automatized (Ashcraft & Christy, 1995) and, therefore,the focus of this study is on broad computational performance.

As with reading comprehension, EF also has logical relevance to math computational performance. *Working memory* is required for computational proficiency because of the students' need to maintain and manipulate numbers in their mind while following a procedural algorithm, and has been shown to be robustly related to mathematics performance (Bull & Scerif, 2001; LeFevre et al., 2013; St Clair-Thompson & Gathercole, 2006). Choosing the appropriate algorithm to implement, and developing strategies to identify problem schemas, requires *planning*. While there is some support for the contribution of planning to mathematics performance (Sikora et al., 2002), there are no other known studies targeting this relation. Additionally, there is variable support for the contribution of *inhibition*, which is conceptualized as being necessary to focus on the multi-step procedure (Bull & Scerif, 2001; Espy et al., 2004; St Clair-Thompson & Gathercole, 2006; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). Finally, *shifting* is hypothesized to allow the student to change from one type of procedure or arithmetic operation across problems, but also has variable support for its contribution (Bull & Scerif, 2001; Espy et al., 2004; St Clair-Thompson & Gathercole, 2006; Van der Ven et al., 2012; Yeniad, Malda, Mesman, Van Ijzendoorn, & Pieper, 2013).

Other factors that impact academics

In addition to EFs, other cognitive and child-specific factors have been identified as having an impact on academic skill development. There is strong support in the literature for the utility of language processes (such as phonological awareness, oral comprehension and word decoding) as predictors of performance for both reading comprehension and math calculations, as the information presented to children involves more complex language as they progress through school (Hecht, Torgesen, Wagner, & Rashotte, 2001; Nation & Snowling, 2004). Of course age is highly relevant to EF given its protracted developmental course; while the students in this study are within a narrow developmental period (e.g., between the ages of 10 and 11), age is nonetheless associated with strong individual differences across EF processes (Anderson, 2008; Anderson, Damasio, Tranel, & Damasio, 2001; Levin et al., 1991). Given these data, language factors and relevant demographic characteristics are considered as potential covariates.

The present study

The goal of the present study was to clarify the nature of the relations of different types of EF with each other and with academic skill. The context is a population whose background is consistent with elevated academic risk, and in an age range particularly relevant for growth in terms of both EF and academic skill. We also evaluate these effects with regard to known predictors of achievement. We had three aims.

Aim 1

First, among the four individual EF processes (working memory, planning, inhibition, and shifting) identified from an important model of EF (Miyake et al., 2000) and the developmental neuropsychological literature (Anderson, 2002, 2008), we evaluated interrelations among cognitive versus behavioral rating scales of EF (e.g., comparing a cognitive task of planning to a behavioral rating scale of planning) and their relations to key areas of academic performance (reading comprehension and math calculations). Given the findings from previous studies, we hypothesized that *across* the two *types* of measures *within* each EF process, low-to-moderate zero-order correlations would be observed (Toplak et al., 2013); we also expected moderate correlations with both reading comprehension and math performance (McAuley, Chen, Goos, Schachar, & Crosbie, 2010; St Clair-Thompson & Gathercole, 2006).

Aim 2

The focus of the second aim was to examine the *relative* predictive power of each method of EF measurement, for both reading comprehension and mathematical outcomes, *within* each of the four EF processes. Since cognitive measures and behavioral rating scales are theorized to capture different information (Sadeh, Burns, & Sullivan, 2012; Toplak et al., 2013), we hypothesized that within each EF process the behavioral rating scales and cognitive measures of EF would both uniquely account for variance in academic performance. We expected that the cognitive measures would have qualitatively larger effect sizes than the behavioral rating measures, if only given the greater volume of empirical support for cognitive EF measures in the prediction of academic outcomes.

Aim 3

Finally, we evaluated the *relative* predictive value of the four theoretically identifiable EF processes for predicting reading comprehension and math calculation performance *across* both EF process *and* type of measurement. Because we expected (as per Aim 2) that at least one type of measure from each EF process would be a relevant predictor, we anticipated being able to evaluate prediction of achievement across all four EF processes. For reading comprehension we hypothesized that working memory, inhibition and planning would be significant predictors of performance, with shifting having a minimal unique contribution. For mathematical calculations we hypothesized that working memory, inhibition and shifting would be significant predictors of performance, with planning having a minimal unique contribution. Finally, we hypothesized that the effect size of working memory would be qualitatively larger than the effect sizes for the other processes in the prediction of performance on both reading comprehension and math calculations measures.

Method

Participants

A total of 93 fourth and fifth grade students enrolled in three public elementary schools in a large metropolitan area took part in this study (for participant characteristics, see Table 1). Although individual data was not available for being at risk or diagnostic status beyond that of Table 1, among the schools sampled, 61% of the student population was designated by the school district as being at risk of experiencing academic difficulties. Further, at the school level, 86% of the children enrolled were identified as being at an economic disadvantage, and 32% were identified as having limited English language proficiency. This study was a portion of a larger parent project focused on reading difficulties, approved by the University of Houston's Institutional Review Board.

Procedures

Data were collected in the schools in a quiet setting over the course of a 3-week data collection period in the spring semester by trained examiners. Cognitive EF data were collected over the course of two individual assessment sessions, typically on consecutive days. Academic outcome data were collected in one group assessment session. EF behavioral rating scale data were collected from 23 teachers, with each teacher completing an average of 4 reports each (range, 3–6). No teacher demographic information was available.

Measures

Behavioral rating measures

Behavior rating inventory of executive function – Teacher report—(BRIEF-T; Gioia et al., 2000). The BRIEF-T is an 86-item behavioral rating scale completed by teachers with eight subscales: Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials and Monitor. The subscales are combined into two commonly-used broad indices, namely Behavioral Regulation and Metacognition. Items are scored on a 1, 2, 3 scale with 1 indicating that the behavior has never been a problem, 2 indicating that the behavior is sometimes problem and 3 indicating that the behavior is often a problem. Raw scores were used as the dependent measure for consistency in this study, as some of the cognitive EF tasks in the study do not have age-based normative data (e.g., standard scores), and therefore age could be accounted for differentially for some measures relative to others in the model. Internal consistency coefficients of the Working Memory, Plan/Organize, Inhibit and Shift subscales ranges from .91 to .96 in the normative sample and .94 to .96 in this sample. The teacher form of this measure was chosen for the purpose of capturing classroom behavior as it relates to academics.

Cognitive measures

Working memory test battery for children (WMTB-C; Pickering & Gathercole, 2001) listening recall subtest—This subtest is often referred to as a complex span task (after Daneman & Carpenter, 1980; Kane, Conway, Miura, & Colflesh, 2007). The task requires the participant to listen to a sentence, decide whether it is true or false, and then remember the last word of the sentence; recall is assessed after an increasing series of sentences. There are six trials within each span length ranging from one to six words per trial. Test retest reliability is .61 for this subtest (Gathercole, Pickering, Ambridge, & Wearing, 2004). The total correct raw score was the dependent measure.

Inquisit tower task (Inquisit 3, 2003; Shallice, 1982)—This task is modeled after the widely used Tower of London task, which has been related to academic performance (Sikora et al., 2002). The Inquisit Tower Task is a computer-based planning task wherein different colored balls are loaded onto one of three sticks of different heights (presented in increasing order). An initial and target configuration are provided, and participants make their model match the target in as few moves as possible while obeying several rules. Two of the problems require a minimum of two moves, three require a minimum of three moves, four require a minimum of four moves and four require a minimum of five moves for a total of thirteen problems. The coefficient alpha for all thirteen problems within this sample was α = .51. The total raw score achieved was used in the analyses.

Delis–Kaplan Executive Function System (DKEFS; Delis, Kaplan, & Kramer, 2001)—The DKEFS is an age-normed battery of tests designed to capture EF. Two subtests were used. To assess inhibition, the Color Word Interference Test (CWIT), which has four conditions (Color Naming, Word Reading, Inhibition and Inhibition/Switching), was administered. The primary measure in this study is the Inhibition condition which requires participants to inhibit a prepotent response by reading the color of the ink of the word, rather than the word itself. Reliability for this measure ranges from .62 to .86. The CWIT subtest is widely used and modeled after the well-established Stroop (1935) task. The dependent measure was the time taken to complete this condition.

The DKEFS Trail Making Test (TMT) assesses shifting. The TMT has five conditions (Visual Scanning, Number Sequencing, Letter Sequencing, Number-Letter Switching and Motor Speed), each of which is timed. Reliability for this measure ranges from .57 to .81. The TMT is a well-used test, with several other versions common (Reitan, 1992).The variable used was the time taken to complete Number-Letter Switching, which requires participants to draw a line switching between connecting numbers and letters in consecutive order.

Academic measures

Gates MacGinitie Reading Tests – Fourth Edition (GMRT; MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000) passage comprehension—The GMRT Passage Comprehension task is a norm-referenced, untimed test of reading comprehension abilities. A total of 11 passages feature 48 questions targeting inference making, summarization, main idea, literal questions, and vocabulary. Reliability exceeds .90 for children in this age range. The coefficient alpha within this sample was α = .91. The dependent measure was the total correct score.

Woodcock–Johnson III (WJ-III; Woodcock, McGrew, & Mather, 2001) tests of academic achievement—The WJ-III Calculations task is a norm-referenced, untimed test of math computations, with 45 items of increasing difficulty. Items are scored 1 if correct and 0 if incorrect. The completion of simple-to-complex math calculations is a commonly-used measure of mathematical performance, as impaired performance on calculation measures is at present one of the main ways through which children are identified as having a math-related learning disability. This subtest has a median reliability of .85 in the age range of 5–19 years. The raw coefficient alpha was .82 for this sample. The specific dependent measure in this study was the total number correct.

Language measures as covariates

Woodcock–Johnson III (WJ-III; Woodcock et al., 2001) tests of academic achievement

Two subtests of the WJ-III were used. Oral Comprehension is a norm-referenced, untimed task of oral cloze procedure where participants are asked to provide one word to complete a passage. This subtest has a median reliability of .80 in the age range of 5–19 years. The coefficient alpha was .63 for this sample. The total raw score was the dependent variable. Letter-Word Identification is a norm-referenced, untimed test of oral word reading where the participant is asked to read up to 76 increasingly difficult English-language letters and words. This subtest has a median reliability of .91 in the age range of 5–19 years. The raw coefficient alpha was .91 for this sample. The total raw score was the dependent variable.

Comprehensive test of phonological processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) Elision

CTOPP Elision is a norm-referenced, untimed task of increasingly difficult phonological segmentation. The participant is presented with a word and is asked to repeat the word without saying a particular sound (an example item would be to say the word “cat” and then say the word “cat” without the /c/ sound, i.e., the participant is expected to respond with “at”; Wagner et al., 1999). This subtest has an average reliability range of .86 to .91 for this age range. The raw coefficient alpha was .91 for this sample. The total raw score was used for analyses.

Analyses

For all analyses, parametric assumptions were evaluated though graphic and statistical means, including the examination of variable distributions. Skewness and kurtosis were all generally within an acceptable range, and other visual evaluations of normality did not reveal strong deviations. The critical value was set to $p < .05$ despite the high number of correlations (though we note the difference between values uncorrected and corrected for multiple comparisons), as we strove to balance Type I and Type II errors; we also provide effect sizes where relevant to help contextualize the results.

To determine the relation among behavioral rating scales of EF, cognitive measures of EF, and academic outcomes, zero-order Pearson correlations were examined. The other primary analytic technique was multiple linear regression, to address the relative merit of both types of EF measures in predicting academic outcomes. Assumptions of regression were evaluated through diagnostic procedures of extremity and influence. Two participants had values greater than 3 standard deviations (*SD*s) from the mean for the cognitive measures of inhibition and/or shifting and were recoded to be within 1 *SD* of the next highest value (to maintain rank ordering), and the analyses proceeded from that point. Distributions of the residuals of the regression equation were examined for both pattern and influence; typically, residual distributions were without pattern, and where influential observations to regression were identified (Cook's distance > 4/n, studentized residuals, and leverage; Kleinbaum & Kleinbaum, 2008), their influence was further evaluated and is noted where relevant.

Statistical significance of the overall regression models and the total variance in the outcome variable (R^2) were evaluated. For models with covariates, an R^2 change ($R^2\Delta$) value (and accompanying *F* value) was calculated to determine the added predictive value of the EF tasks over the covariates. The effect size for each predictor variable is reported via the squared semi-partial correlation coefficient (sr^2) within each regression model.

Covariates

Age, race, education status (special education and gifted/talented), and relevant language variables (oral comprehension, letter-word identification and phonological awareness) are all significantly related to both reading comprehension and math computations (with the exception of age for math) as defined through correlational and general linear model analyses ($p < .05$). All relevant covariates together were initially examined, and the covariate set was trimmed to include only those which remained contributory in the context of the set, as detailed below.

Results

Preliminary results

Descriptive statistics for the covariates, EF cognitive tasks, and behavioral rating scales, and academic outcomes, appear in Table 2.

Aim 1: Individual relations among EF measures and with academic outcomes

The correlations among EF measures are presented in Table 3. Correlations between each EF cognitive and behavioral rating scale measures and partial-age correlations between each EF cognitive and behavioral rating scale measures are presented in Table 4. Within each EF process, the cognitive and behavioral rating scale measures had low-to-moderate correlations with one another; only the inhibition ($r = .25$, $p < .05$) and shifting ($r = .25$, $p < .05$) processes had significant correlations. Each cognitive measure of EF correlated significantly with reading comprehension (range $|r| = .32$ to .50) and math (range $|r| = .40$ to .50). Each behavioral rating scale measure of EF also correlated significantly with reading comprehension (range $|r| = .38$ to .55) and math (range $|r| = .29$ to .42). Use of a Bonferroni correction procedure (.05/21) for the correlations in Table 3 did not change substantive results, with only 3 of the 19 values no longer significant; notably, 2 of these were cross-method relations. In Table 4, of the 28 correlations (with or without age partialled) 22 were significant; with Bonferroni correction (.05/28) the remaining significant correlations were mostly among the BRIEF-T measures.

Aim 2: Predictive utility of both behavioral and cognitive EF measures in academic performance within EF process

Reading Comprehension—In models with both types of measures (cognitive and behavioral rating scales), for each of the four EF processes separately, and without covariates, all EF measures were unique predictors of performance for GMRT Reading Comprehension (all *p*s < .05). The pair of EF measures of working memory, planning, inhibition and shifting, accounted for 50%, 35%, 21% and 28% of the variance, respectively, in reading scores.

As shown in Table 5, in the covariate-only model, age, oral comprehension, phonological awareness, special education status and gifted and talented status accounted for 60% of the variance in reading comprehension. Letter-word identification was not a significant predictor in the context of the other covariates and was not included in future analyses.

Models that included the covariates, within each EF process (working memory, planning, inhibition and shifting), were all significant, all *p*s < .001. The working memory model accounted for 67% of the variance ($R^2\Delta$ for adding the EF measures to the model with covariates alone = .07, $p < .001$), with unique prediction for both the cognitive measure (WMTB-C Listening Recall, $p = .034$) and the behavioral rating scale (BRIEF-T Working Memory subscale, $p < .001$). The planning model (EF $R^2\Delta = .05$, $p = .006$) showed unique prediction only for the behavioral rating scale (BRIEF-T Plan/Organize subscale, $p = .006$). Similarly, the inhibition (EF $R^2\Delta = .04$, $p = .01$) and shifting models (EF $R^2\Delta = .04$, $p = .01$) showed only the behavioral rating scales to be significant predictors (BRIEF-T Inhibit subscale, $p = .041$; BRIEF-T Shift subscale, $p = .003$).

Math calculations—In models with the pair of measures within each of the four EF processes separately, and without covariates, both EF measures were uniquely predictive of performance on the WJ-III Calculations measure for working memory, planning and shifting ($p < .05$), accounting for 35%, 27% and 31% of the variance, respectively. The two inhibition measures accounted for 34% of the variance, though only the (cognitive) DKEFS Color Word Interference Inhibition trial ($p < .001$) was a unique predictor, whereas the (behavioral rating) BRIEF-T inhibition scale was $p = .080$.

The covariate-only model (with phonological awareness, special education status and gifted and talented status) accounted for 45% of the variance in math calculations performance. Age, oral comprehension and letter-word identification were not significant predictors in the context of the other covariates and were not included in future analyses.

Models that included the covariates, and each pair of measures within EF process separately, were all significant at $p < .001$. The working memory model accounted for 50% of the variance ($R^2\Delta$ for adding the EF measures to the model with covariates alone was .05, $p = .013$), with unique prediction for both the cognitive measure (WMTB-C Listening Recall, $p = .028$) and the behavioral rating scale (BRIEF-T Working Memory subscale, $p = .043$). The planning model (EF $R^2\Delta = .05$, $p = .021$) showed only the cognitive measure of planning (Tower of London, $p = .031$) to be a unique predictor. Regression diagnostics revealed eight observations with influential residuals, and their removal produced a model with neither the behavioral rating scale (BRIEF-T Planning, $p = .055$) nor the cognitive measure ($p = .219$) as a significant predictor $F(5, 79) = 18.43$, $p < .001$, $R^2 = .54$. The full inhibition model (EF $R^2\Delta = .12$, $p = .001$) and the full shifting model ($R^2\Delta = .08$, $p < .001$) both only showed the cognitive measures (DKEFS, Color Word Interference–Inhibition, $p < .001$; DKEFS TMT–Number-Letter Switching, $p = .001$, respectively) to be significant predictors (Table 6).

Aim 3: Predictive utility of EF in academic performance

Final regression models for this aim are summarized in Table 7. These analyses were built by including significant predictors from the more discrete Aim 2 models, which had the effect of keeping the number of predictors in these final models at a reasonable number for the sample size. However, these final models consider EF not only across *type of measure* but also across EF *process*.

Reading comprehension

The inclusion of multiple BRIEF-T variables resulted in multicollinearity issues (see Table 4 for intercorrelations). As such, only a single behavioral rating measure (that with the greatest zero-order correlation) was included (BRIEF-T Working Memory variable, $r = .55$). This resulted in a reading comprehension model being identical to the “working memory only” model from Aim 2. Therefore, in order to more directly address the hypotheses across EF processes, a model was constructed that included the other cognitive EF measures (Tower of London, Color Word Interference–Inhibition, and TMT–Number-Letter Switching) despite their non-significance in the covariate models of Aim 2. The overall model was significant $F(10, 82) = 17.43$, $p < .001$, $R^2 = .68$. However, the only uniquely contributing EF variable in the model was the BRIEF-T Working Memory ($p = .001$), along with age ($p < .001$), phonological awareness ($p = .002$), and gifted and talented status ($p = .039$). Regression diagnostics revealed six observations with influential residuals. Their removal produced an overall significant model $F(10, 76) = 16.43$, $p < .001$, $R^2 = .68$, with the cognitive measure of Working Memory (WMTB-C, $p = .011$) and the BRIEF-T Working Memory ($p = .002$) as significant predictors in addition to age ($p < .001$), oral comprehension ($p = .038$), phonological awareness ($p = .005$), and gifted and talented status ($p = .020$).

Math calculations

The final model for math calculations included covariates of phonological awareness, gifted and talented status, and special education status, along with one behavioral rating scale predictor of EF (BRIEF-T Working Memory) and all four cognitive EF predictors (WMTB-C Listening Comprehension, Tower of London, DKEFS Color-Word Interference Inhibition, DKEFS TMT–Number-Letter Switching). The overall model was significant (EF $R^2\Delta = .16$, $p < .001$). However, the only uniquely contributing EF variable in this model was the DKEFS Color-Word Interference Inhibition ($p = .002$), along with the covariates of phonological awareness ($p = .045$) and gifted and talented status ($p = .001$). Regression diagnostics revealed five observations with influential residuals. Their removal produced an overall significant model $F(8, 79) = 21.08$, $p < .001$, $R^2 = .68$, with the cognitive measure of planning (Tower of London, $p = .02$) as a significant predictor in addition to the DKEFS Color-Word Interference Inhibition ($p < .001$), phonological awareness ($p = .011$) and gifted and talented status ($p < .001$).

Discussion

The purpose of the present study was to better understand the interrelations between cognitive and behavioral ratings of EF and their joint relation to academic measures, and to do so in the context of an understudied population. The advantages of this particular study are the inclusion of both cognitive and behavioral measures of each of four model-based processes of EF (working memory, planning, inhibition and shifting; Anderson, 2002, 2008; Miyake et al., 2000). In addition to consideration of their interrelations, we also considered whether they differed in their prediction of important outcomes, namely academic functions, while including relevant covariates.

Modest relations were found between cognitive and behavioral rating measures of EF within each of the four processes. However, there were stronger yet moderate signifi-cant relations found between all of the EF measures and reading comprehension and math performance across domain and measure type (see Table 3). All but one of the cognitive and behavioral rating measures of EF, when considered within each of the four EF processes separately, was a significant predictor of performance of *both* reading comprehension and math measures (the exception being the inhibition rating scale for math). Inclusion of relevant covariates modified the utility of each measure, with behavioral rating scales being more relevant for reading comprehension and cognitive measures being more relevant for math calculations. When strong covariates and all EF processes were simultaneously considered, working memory had the strongest unique contribution to reading comprehension, whereas inhibition and planning were strongest for mathematics computations.

Findings for Aim 1: Cognitive and behavioral rating scale measures capture complementary perspectives of EF

Although both the cognitive and behavioral rating scale measures of EF were moderately related to each academic outcome, the two types of EF measures had only modest relations with one another (range $|r|$ = .19 to .25). Even when "matched" on specific EF processes, only inhibition and shifting had significant correlations across measure types. The only known previous study that examined relations within the same four EF processes found slightly different results. Toplak et al. (2008) found significant relations in the processes of working memory (with parent and teacher reports) and inhibition (with teacher reports only), but no significant relations in the processes of planning or shifting. Toplak et al. included an older population (44 adolescents with a diagnosis of ADHD and 44 comparison controls) of only native English-speaking backgrounds, used different EF cognitive measures, and the correlational analyses did not separate ADHD participants and controls. Additionally, although academic measures were collected they were not examined as outcome measures. Despite the differences in methodology between this study and previous studies, our findings are consistent and indicate that behavioral rating scales and cognitive measures of EF are weakly related to each other, even at the level of individual EF processes, which allows for the use of both types of measures without concern of collecting redundant information.

We had hypothesized that the comparison of process-specific measures on a one-to-one basis would lead to stronger interrelations. However, this was not the case, at least in this population. The four scales of the BRIEF-T showed high overlap and very strong relations to one other across EF process (range |*r*| = .72 to .90), whereas the cognitive measures were separable and did not relate highly to one another across EF process (range |*r*| = .22 to .47).

It appears that the robust interrelations of the subscales of the BRIEF-T indicate that a composite of this measure (like those developed by the test publishers) may provide similar information to the subscales as assessments of the daily influence of EF in children. It is possible that such strong relations may be in part due to rater effects. Evaluation of behavioral rating scales by a single person (e.g., teacher) whose interactions emphasize a single (though quite important) environment may lead to less specific knowledge of individual EF skills across process. Previous studies have found that behavioral ratings from informants from different situational settings (i.e., parents, teachers, mental health workers) have limited relations with one another (Achenbach, McConaughy, & Howell, 1987); however, some have identified predictive utility for both parent and teacher reports (Verhulst, Koot, & der Ende, 1994). This study focused on teacher ratings, given the specific focus on academic outcomes, for which teachers may be better suited as raters. Additionally, Toplak et al. (2008) included both parent and teacher ratings in their study and came to the same conclusion regarding the relation between behavioral rating scales and cognitive measures of EF, despite finding two significant relations with parent ratings and only one with teacher ratings. This rater effect is important to consider when examining the role of behavioral rating scales in cognition, which may be addressed clinically by including parents and teachers, and/or multiple teacher ratings.

That the cognitive measures showed weak interrelations may lead to their relating differentially to outcomes, and also implies that clinicians and researchers should be wary of combining such scores into composites to represent "overall cognitive EF", except perhaps when considered in a latent variable framework. It is however true that the cognitive measures used here were chosen specifically to represent *different* EF processes (Anderson, 2002, 2008; Miyake et al., 2000) and so future research is needed to determine the extent to which cognitive EF measures might be fruitfully combined *within* a specific process.

What can be concluded about the relations between behavioral rating and cognitive measures of EF? The measures are not capturing the same information, given the present results, and those of Toplak and colleagues (2008, 2013). It is also unlikely that behavioral rating and cognitive measures of EF are mutually exclusive; modest relations were identified here, and some other studies have identified stronger relations (Alloway et al., 2009; Shimoni et al., 2012). However, in the Shimoni et al. (2012) study, this was only true for the BRIEF (parent) Working Memory and Planning subscales, both in relation to a total score on the BADS-C. In the Alloway et al. (2009) study the authors only focused on working memory. The most likely scenario is that both types of EF measures are capturing complementary perspectives of the same, or at least similar, construct. Cognitive measures may represent a child's ability demonstrate EF skill at its maximum in a highly structured setting, whereas behavioral rating scale score may represent the observable behaviors associated with that EF in everyday settings. A similar pattern has been noted in children with ADHD (Barkley, 1991).

Findings for Aim 2. Both cognitive and behavioral rating scale measures within EF processes are predictive of academic performance

Consistent with previous literature, there was a significant relation between the cognitive measures of each EF process and reading comprehension performance, with working memory having the strongest absolute predictive utility (Carretti, Borella, Cornoldi, & De Beni, 2009; Chiappe, Hasher, & Siegel, 2000; Sesma et al., 2009; St Clair-Thompson & Gathercole, 2006). McAuley et al. (2010) previously demonstrated a relation between the BRIEF Metacognitive Index (parent report) and math and reading performance, and there is some evidence that the BRIEF Working Memory subscale (parent report) has been related to reading disorders and poor reading performance (Gioia, Isquith, Kenworthy, & Barton, 2002; Locascio et al., 2010). The present study adds to this extant knowledge by demonstrating that both cognitive and behavioral ratings in three additional EF processes—shifting, inhibition, and planning—are also predictive of reading when considered separately (and without covariates). Thus, while the present results would indicate that both behavioral rating scales and cognitive measures of EF do not substitute for one another, they also suggest that each contributes uniquely to the prediction of reading comprehension.

The consideration of a child's language capabilities significantly altered the contribution of both types of EF measures across processes in predicting reading comprehension in this study. Despite the strong influence of language skills in a child's reading comprehension performance, the *additional* variance accounted for by the inclusion of the pair of EF measures was pertinent (range $R^2\Delta$ = .04 to .07), within each EF process. However, of the EF processes, only the working memory model held a unique contribution for both cognitive and behavioral measures when covariates were included. This pattern may be more indicative of shared variance between the measures used to capture language and the cognitive measures of EF (i.e., the "task impurity problem"; Burgess et al., 1998) than a lack of utility for cognitive measures in predicting reading comprehension. The inclusion of behavioral rating scale measures of EF may provide information about a child's behavioral interaction with their environment that may influence their use of cognitive skills in the service of a goal-directed task (e.g., if a child is observed to have challenges with working memory, they may also experience challenges when undertaking academic tasks that require them to hold complicated instructions in their minds).

For mathematical calculations, each pair of behavioral rating scale and cognitive measures of EF were predictive. This finding contributes to the understanding of the connection between EF and mathematics in the current literature (Bull & Scerif, 2001; Friso-van den Bos, van der Ven, Kroesbergen, & Van Luit, 2013; St Clair-Thompson & Gathercole, 2006), which is less developed than the literature examining the role of EF in reading skills and has focused on cognitive measures of EF, particularly working memory. Although rating measures of behavioral inattention have been frequently associated with math performance (Finn, Pannozzo, & Voelkl, 1995; Raghubar et al., 2009), few studies have examined the role of EF behavioral rating measures. An exception is Mahone et al. (2002), who found that parent report working memory and inhibition subscales and other BRIEF indices were related to math performance, as measured by the math composite score from the Wechsler Individual Achievement Test (WIAT; Wechsler, 1992). Therefore, the present results add to the limited empirical base, and suggest that, like reading comprehension, both behavioral rating and cognitive measures of EF are relevant in empirical and clinical studies of mathematic skills.

As with reading comprehension, the role of EF measures in the prediction of mathematical calculations performance were altered when relevant covariates were taken into consideration, though in a different way. Although both phonological awareness and educational program status accounted for a significant amount of variance in the prediction model, the inclusion of the EF measures added additional predictive utility, as indicated by the significant R^2 change for each model. Unlike the case for reading, here each of the cognitive measures of EF maintained significant predictive utility when considered with its behavioral rating scale counterpart, even when covariates were included. In contrast, only one of the four models (working memory) showed a unique contribution for behavioral rating scale EF. This type of result is generally consistent with prior work that emphasizes the role of cognitive measures of EF as important predictors of mathematical performance (Bull & Scerif, 2001; St Clair-Thompson & Gathercole, 2006; Van der Ven et al., 2012).

Findings of Aim 3: EF predictors of academic skills

Challenges created by the multicollinearity of the BRIEF-T variables and the influence of key covariates resulted in a limited ability to also examine the comparative utility of all four EF processes as represented by both cognitive and behavior rating measures. However, when the cognitive measures of planning, inhibition and shifting were placed together with the working memory cognitive and behavior rating measures, their relative contribution was minimal. The findings from this study add additional support for working memory as a key predictor of reading comprehension, previously identified with both cognitive measures (Carretti et al., 2009) and behavioral rating scales (Gioia et al., 2002).

For mathematical calculations, the current study adds additional support for the role of specific EF processes, in the context of others, in the development of mathematical calculation skill. The results from previous studies have shown mixed results (Bull & Scerif, 2001; Espy et al., 2004; Sikora et al., 2002; St Clair-Thompson & Gathercole, 2006; Van der Ven et al., 2012; Yeniad et al., 2013). While the cognitive measure of inhibition was clearly the strongest predictor of mathematical calculations, as evidenced by the greatest portion of the variance accounted for, the measure of planning also showed unique significance when influential observations were excluded. Additionally, the combination of the other EF processes (working memory and shifting) accounted for a portion of shared variance and, together with relevant covariates, accounted for a significant proportion of the outcome. Furthermore, given that the cognitive measures in this study had modest interrelations, the shared variance likely does not come from the measures capturing the same cognitive ability, but from overlapping variance in their prediction of mathematical calculations. The academic process of mathematical calculations appears to involve all four EF processes (working memory, planning, inhibition and shifting), with inhibition and planning being particularly strong predictors of performance.

Limitations

The participants in this study had relatively high BRIEF-T scores (see Table 2). The relatively high scores compared with the normative sample may reflect the phenomena of minority students being rated higher on behavioral rating measures (Dwivedi & Banhatti, 2005; Epstein et al., 2005; Waber et al., 2006). Differences in the make-up of the published normative sample and the population used in this study are also relevant, as the normative sample for the Teacher form is made up of 72% Caucasian, 14% African American, 4% Hispanic, 6% Asian/Pacific Islander and <1% Native American children, which although consistent with census-based distributions differs from the student body sampled from here (see Table 1). To some extent, the elevated scores may reflect that the development of EF is impacted by external factors, including low SES (Ursache et al., 2012). The impact of the elevated scores in this study are mitigated to the extent that the relations identified between EF measures were commensurate with findings from previous studies. The context of the study would also have been enhanced if further individual data were available for these students, or if diagnostic evaluations had been conducted to rule in/out specific disorders, although the descriptive data available at the school level does suggest some elevated risk in this sample.

Other limitations reflect measurement issues. The strong correlations between BRIEF-T subscales making differential interpretations among them difficult. If separable behavioral rating scales (either by reporter or scale composition) had been utilized, perhaps the results would have been different, and future studies may separate out these methodological characteristics. Although both cognitive and behavioral approaches have limitations—e.g., overt structure for cognitive (Salthouse et al., 2003), response bias for ratings (Isquith et al., 2013)—using both together demonstrated their complementary contribution. Finally, although the cognitive measures in this study were chosen from theoretically motivated models, given that researchers and clinicians use a variety of EF measures to tap into these processes, the results may have been different had different cognitive measures been utilized. The reliabilities for the cognitive measures of EF were lower than the reliabilities for the behavioral rating scales and cognitive measures of academics both within this sample and based on the normative development of the measures. Given that the aim of an assessment of EF is to capture goal-directed behavior in a novel situation, it has been suggested that low test-retest reliability for cognitive measures of EF is expected, and that it is a product of the loss of novelty and differential utilization of strategies at the second assessment (Denckla, 1996; Rabbit, 1997). Additionally, it has been argued that a comparison of tasks with low reliabilities will produce a low correlation that likely reflects the nature of the internal inconsistences of the measures (Miyake et al., 2000). Thus, the weaker relations among the cognitive measures of EF in this sample may be indicative of the lower reliabilities of these measures. Latent variables or other means may have produced more generalizable results for the cognitive measures of EF. Clearly, however, more work is needed regarding the structure of EF in order to better identify what may represent “gold standards” for EF measurement.

Conclusions

This study examined the interrelations between behavioral rating scales and cognitive measures of EF within four processes (working memory, planning, inhibition and shifting). Both the behavioral rating scales and the cognitive measures of EF were important in predicting both reading comprehension and math outcomes, even when strong covariates were considered. These relations manifest to different extents for both reading comprehension and math. The processes of working memory and the behavioral rating scales were more strongly related to reading comprehension and the processes of inhibition and planning, and cognitive measures were more strongly related to mathematical calculations. In a clinical setting the findings of this study would support the use of both types of measures within and across EF processes in an academic or neuropsychological assessment. Even in a population where environmental factors likely have a strong influence on cognitive development, both types of EF measures can be used in a complementary and advantageous fashion to inform academic achievement.

Acknowledgements

The attitudes and opinions expressed in this presentations are those of the authors and do not necessarily reflect those of the funding agency. The authors thank the coworkers, parents, teachers, and school and district officials who made this research possible.

Funding

This research was supported by Award Number P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health & Human Development to the University of Houston. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health.

References

Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychological Bulletin. 1987; 101(2):213–232. doi:10.1037/0033-2909.101.2.213. [PubMed: 3562706]

Alloway, TP. Automated Working Memory Assessment (AWMA): Manual. Pearson Education; Oxford: 2007.

Alloway, TP. Working Memory Rating Scale (WMRS): Manual. Pearson Education; Oxford: 2008.

Alloway TP, Gathercole SE, Kirkwood HJ, Elliott J. The working memory rating scale: A classroom-based behavioral assessment of working memory. Learning and Individual Differences. 2009; 19:242–245. doi:10.1016/j.lindif.2008.10.003.

Anderson P. Assessment and development of executive function (EF) during childhood. Child Neuropsychology. 2002; 8(2):71–82. doi:10.1076/chin.8.2.71.8724. [PubMed: 12638061]

Anderson, PJ. Towards a developmental model of executive function. In: Anderson, V.Jacobs, R., Anderson, PJ., editors. Executive functions and the frontal lobes: A lifespan perspective. Taylor & Francis; Philadelphia, PA: 2008. p. 3-21.

Anderson SW, Damasio H, Tranel D, Damasio AR. Long-term sequelae of prefrontal cortex damage acquired in early childhood. Developmental Neuropsychology. 2001; 18(3):281–296. doi:10.1207/S1532694202. [PubMed: 11385828]

Arrington CN, Kulesz PA, Francis DJ, Fletcher JM, Barnes MA. The contribution of attentional control and working memory to reading comprehension and decoding. Scientific Studies of Reading. 2014; 18(5):325–346. doi:10.1080/10888438.2014.902461.

Ashcraft MH, Christy KS. The frequency of arithmetic facts in elementary texts: Addition and multiplication in grades 1-6. Journal for Research in Mathematics Education. 1995; 26(5):396–421. doi:10.2307/749430.

Barkley RA. The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. Journal of Abnormal Child Psychology. 1991; 19(2):149–178. doi:10.1007/BF00909976. [PubMed: 2056161]

Barnes, M., Fuchs, L., Ewing-Cobbs, L. Pediatric neuropsychology, second edition: research, theory, and practice. 2nd ed. Guilford Press; New York, NY: 2009. Math disabilities.

Best JR, Miller PH, Naglieri JA. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. Learning and Individual Differences. 2011; 21(4):327–336. doi:10.1016/j.lindif.2011.01.007. [PubMed: 21845021]

Biancarosa, G., Snow, CE. Reading next: A vision for action and research in middle and high school literacy: A report from Carnegie corporation of New York. Alliance for Excellent Education; Washington, DC: 2004.

Biederman J, Monuteaux MC, Doyle AE, Seidman LJ, Wilens TE, Ferrero F, Faraone SV. Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. Journal of Consulting and Clinical Psychology. 2004; 72(5):757–766. doi:10.1037/0022-006X.72.5.757. [PubMed: 15482034]

Borella E, Carretti B, Pelegrina S. The specific role of inhibition in reading comprehension in good and poor comprehenders. Journal of Learning Disabilities. 2010; 43(6):541–552. doi:10.1177/0022219410371676. [PubMed: 20606207]

Bull R, Scerif G. Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. Developmental Neuropsychology. 2001; 19(3):273–293. doi:10.1207/S15326942DN1903_3. [PubMed: 11758669]

Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. Journal of the International Neuropsychological Society. 1998; 4(6):547–558. doi:10.1017/S1355617798466037. [PubMed: 10050359]

Cain K, Oakhill J, Bryant P. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. Journal of Educational Psychology. 2004; 96(1):31–42. doi:10.1037/0022-0663.96.1.31.

Carretti B, Borella E, Cornoldi C, De Beni R. Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. Learning and Individual Differences. 2009; 19(2):246–251. doi:10.1016/j.lindif.2008.10.002.

Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. Neuropsychology Review. 2003; 13(4):181–197. doi:10.1023/B:NERV.0000009483.91468.fb. [PubMed: 15000225]

Chiappe P, Hasher L, Siegel LS. Working memory, inhibitory control, and reading disability. Memory & Cognition. 2000; 28(1):8–17. doi:10.3758/BF03211570. [PubMed: 10714133]

Christopher ME, Miyake A, Keenan JM, Pennington B, DeFries JC, Wadsworth SJ, Olson RK. Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. Journal of Experimental Psychology: General. 2012; 141(3):470–488. doi:10.1037/a0027375. [PubMed: 22352396]

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19(4):450–466. doi:10.1016/S0022-5371(80)90312-6.

Delis, D., Kaplan, E., Kramer, J. Delis-kaplan executive function scale. The Psychological Corporation; San Antonio, TX: 2001.

Denckla, MB. A theory and model of executive function: A neuropsychological perspective. In: Lyon, GR., Krasnegor, NA., editors. Attention, memory, and executive function. Brookes; Baltimore, MD: 1996. p. 263-278.

Downey DB, Pribesh S. When race matters: Teachers' evaluations of students' classroom behavior. Sociology of Education. 2004; 77(4):267–282. doi:10.1177/003804070407700401.

Dwivedi KN, Banhatti RG. Attention deficit/hyperactivity disorder and ethnicity. Archives of Disease in Childhood. 2005; 90(suppl 1):i10–i12. doi:10.1136/adc.2004.058180. [PubMed: 15665149]

Emslie, H., Wilson, FC., Burden, V., Nimmo-Smith, I., Wilson, BA. Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C). Harcourt Assessment/The Psychological Corporation; London: 2003.

Epstein JN, Willoughby M, Valencia EY, Tonev ST, Abikoff HB, Eugene L, Hinshaw SP. The role of children's ethnicity in the relationship between teacher ratings of attention-deficit/hyperactivity disorder and observed classroom behavior. Journal of Consulting and Clinical Psychology. 2005; 73(3):424–434. doi:10.1037/0022-006X.73.3.424. [PubMed: 15982140]

Espy KA, McDiarmid MM, Cwik MF, Stalets MM, Hamby A, Senn TE. The contribution of executive functions to emergent mathematic skills in preschool children. Developmental Neuropsychology. 2004; 26(1):465–486. doi:10.1207/s15326942dn2601_6. [PubMed: 15276905]

Farah MJ, Shera DM, Savage JH, Betancourt L, Giannetta JM, Brodsky NL, Hurt H. Childhood poverty: Specific associations with neurocognitive development. Brain Research. 2006; 1110(1): 166–174. doi:10.1016/j.brainres.2006.06.072. [PubMed: 16879809]

Finn JD, Pannozzo GM, Voelkl KE. Disruptive and inattentive-withdrawn behavior and achievement among fourth graders. The Elementary School Journal. 1995; 95(5):421–434. doi:10.1086/461853.

Finn JD, Rock DA. Academic success among students at risk for school failure. Journal of Applied Psychology. 1997; 82(2):221–234. doi:10.1037/0021-9010.82.2.221. [PubMed: 9109280]

Friso-Van Den Bos I, van der Ven SHG, Kroesbergen EH, Van Luit JEH. Working memory and mathematics in primary school children: A meta-analysis. Educational Research Review. 2013; 10:29–44. doi:10.1016/j.edurev.2013.05.003.

Fuchs LS, Fuchs D, Compton DL, Powell SR, Seethaler PM, Capizzi AM, Fletcher JM. The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. Journal of Educational Psychology. 2006; 98(1):29–43. doi:10.1037/0022-0663.98.1.29.

Gathercole SE, Pickering SJ, Ambridge B, Wearing H. The structure of working memory from 4 to 15 years of age. Developmental Psychology. 2004; 40(2):177–190. doi:10.1037/0012-1649.40.2.177. [PubMed: 14979759]

Gioia, GA., Isquith, PA., Guy, SC., Kenworthy, L. Behavior rating inventory of executive function. Psychological Assessment Resources; Odessa, FL: 2000.

Gioia GA, Isquith PK, Kenworthy L, Barton RM. Profiles of everyday executive function in acquired and developmental disorders. Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence. 2002; 8(2):121–137. doi:10.1076/chin.8.2.121.8727. [PubMed: 12638065]

Hackman DA, Farah MJ. Socioeconomic status and the developing brain. Trends in Cognitive Sciences. 2009; 13(2):65–73. doi:10.1016/j.tics.2008.11.003. [PubMed: 19135405]

Hecht SA, Torgesen JK, Wagner RK, Rashotte CA. The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: a longitudinal study from second to fifth grades. Journal of Experimental Child Psychology. 2001; 79(2):192–227. doi:10.1006/jecp.2000.2586. [PubMed: 11343408]

Hughes C, Graham A. Measuring executive functions in childhood: Problems and solutions? Child and Adolescent Mental Health. 2002; 7(3):131–142. doi:10.1111/1475-3588.00024.

Inquisit 3. Millisecond Software; Seattle, WA: 2003.

Isquith PK, Roth RM, Gioia G. Contribution of rating scales to the assessment of executive functions. Applied Neuropsychology: Child. 2013; 2(2):125–132. doi:10.1080/21622965.2013.748389. [PubMed: 23442015]

Kane MJ, Conway ARA, Miura TK, Colflesh GJH. Working memory, attention control, and the n-back task: A question of construct validity. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2007; 33(3):615–622. doi:10.1037/0278-7393.33.3.615.

Kishiyama MM, Boyce WT, Jimenez AM, Perry LM, Knight RT. Socioeconomic disparities affect prefrontal function in children. Journal of Cognitive Neuroscience. 2009; 21(6):1106–1115. doi: 10.1162/jocn.2009.21101. [PubMed: 18752394]

Kleinbaum, DG., Kleinbaum, DG. Applied regression analysis and other multivariable methods. Brooks/Cole; Australia; Belmont, CA: 2008.

LeFevre J-A, Berrigan L, Vendetti C, Kamawar D, Bisanz J, Skwarchuk S-L, Smith-Chant BL. The role of executive attention in the acquisition of mathematical skills for children in Grades 2 through 4. Journal of Experimental Child Psychology. 2013; 114(2):243–261. doi:10.1016/j.jecp.2012.10.005. [PubMed: 23168083]

Levin HS, Culhane KA, Hartmann J, Evankovich K, Mattson AJ, Harward H, Fletcher JM. Developmental changes in performance on tests of purported frontal lobe functioning. Developmental Neuropsychology. 1991; 7(3):377–395. doi:10.1080/87565649109540499.

Lezak, MD., Howieson, DB., Bigler, ED., Tranel, D. Neuropsychological assessment. 5th ed. Oxford University Press; New York, NY: 2012.

Locascio G, Mahone EM, Eason SH, Cutting LE. Executive dysfunction among children with reading comprehension deficits. Journal of Learning Disabilities. 2010; 43(5):441–454. doi: 10.1177/0022219409355476. [PubMed: 20375294]

MacGinitie, W., MacGinitie, R., Maria, K., Dreyer, LG., Hughes, KE. Gates-macginitie reading tests (GMRT). 4th ed. Riverside Publishing; Itasca, IL: 2000.

Mahone EM, Cirino PT, Cutting LE, Cerrone PM, Hagelthorn KM, Hiemenz JR, Denckla MB. Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. Archives of Clinical Neuropsychology. 2002; 17(7):643–662. doi:10.1016/S0887-6177(01)00168-8. [PubMed: 14591848]

McAuley T, Chen S, Goos L, Schachar R, Crosbie J. Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? Journal of the International Neuropsychological Society. 2010; 16(3):495–505. doi:10.1017/S1355617710000093. [PubMed: 20188014]

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive Psychology. 2000; 41(1):49–100. doi:10.1006/cogp.1999.0734. [PubMed: 10945922]

Nation K, Snowling MJ. Beyond phonological skills: Broader language skills contribute to the development of reading. Journal of Research in Reading. 2004; 27(4):342–356. doi:10.1111/j.1467-9817.2004.00238.x.

Noble KG, Farah MJ, McCandliss BD. Socioeconomic background modulates cognition–achievement relationships in reading. Cognitive Development. 2006; 21(3):349–368. doi:10.1016/j.cogdev.2006.01.007. [PubMed: 19789717]

Pickering, SJ., Gathercole, SE. Working memory test battery for children. Psychological Corporation Europe; London: 2001.

Rabbit, P. Introduction: Methodologies and models in the study of executive function. In: Rabbit, P., editor. Methodology of frontal and executive function. Psychology Press; Hove, UK: 1997. p. 1-38.

Raghubar K, Cirino P, Barnes M, Ewing-Cobbs L, Fletcher J, Fuchs L. Errors in multi-digit arithmetic and behavioral inattention in children with math difficulties. Journal of Learning Disabilities. 2009; 42(4):356–371. doi:10.1177/0022219409335211. [PubMed: 19380494]

Reitan, RM. Trail making test: Manual for administration and scoring. Reitan Neuropsychology Laboratory; 1992.

Reynolds, CR., Kamphaus, RW. Behavior assessment system for children-second edition (BASC-2). AGS; Circle Pines, MN: 2004.

Roy AL, Raver CC. Are all risks equal? Early experiences of poverty-related risk and children's functioning. Journal of Family Psychology. 2014; 28(3):391–400. doi:10.1037/a0036683. [PubMed: 24749652]

Sadeh S, Burns M, Sullivan A. Examining an executive function rating scale as a predictor of achievement in children at risk for behavior problems. School Psychology Quarterly. 2012; 27(4): 236–246. doi:10.1037/spq0000012. [PubMed: 23294237]

Salthouse TA, Atkinson TM, Berish DE. Executive functioning as a potential mediator of age-related cognitive decline in normal adults. Journal of Experimental Psychology: General. 2003; 132(4): 566–594. doi:10.1037/0096-3445.132.4.566. [PubMed: 14640849]

Sesma HW, Mahone EM, Levine T, Eason SH, Cutting LE. The contribution of executive skills to reading comprehension. Child Neuropsychology. 2009; 15(3):232–246. doi: 10.1080/09297040802220029. [PubMed: 18629674]

Shallice T. Specific impairments of planning. Philosophical Transactions of the Royal Society of London. B, Biological Sciences. 1982; 298(1089):199–209. doi:10.1098/rstb.1982.0082. [PubMed: 6125971]

Shimoni M, Engel-Yeger B, Tirosh E. Executive dysfunctions among boys with attention deficit hyperactivity disorder (ADHD): Performance-based test and parents report. Research in Developmental Disabilities. 2012; 33(3):858–865. doi:10.1016/j.ridd.2011.12.014. [PubMed: 22230238]

Sikora DM, Haley P, Edwards J, Butler RW. Tower of London test performance in children with poor arithmetic skills. Developmental Neuropsychology. 2002; 21(3):243–254. doi:10.1207/S15326942DN2103_2. [PubMed: 12233937]

Sirin SR. Socioeconomic status and academic achievement: A meta-analytic review of research. Review of Educational Research. 2005; 75(3):417–453. doi:10.3102/00346543075003417.

St Clair-Thompson HL, Gathercole SE. Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. The Quarterly Journal of Experimental Psychology. 2006; 59(4):745–759. doi:10.1080/17470210500162854. [PubMed: 16707360]

Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology. 1935; 18(6):643–662. doi:10.1037/h0054651.

Toplak ME, Bucciarelli SM, Jain U, Tannock R. Executive functions: Performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). Child Neuropsychology. 2008; 15(1):53–72. doi:10.1080/09297040802070929. [PubMed: 18608232]

Toplak ME, West RF, Stanovich KE. Do performance-based measures and ratings of executive function assess the same construct? Journal of Child Psychology and Psychiatry. 2013; 54(2):131–143. doi:10.1111/jcpp.12001. [PubMed: 23057693]

Ursache A, Blair C, Raver CC. The promotion of self-regulation as a means of enhancing school readiness and early achievement in children at risk for school failure. Child Development Perspectives. 2012; 6(2):122–128. doi:10.1111/j.1750-8606.2011.00209.x.

Van Der Ven SHG, Kroesbergen EH, Boom J, Leseman PPM. The development of executive functions and early mathematics: A dynamic relationship. British Journal of Educational Psychology. 2012; 82(1):100–119. doi:10.1111/j.2044-8279.2011.02035.x. [PubMed: 22429060]

Verhulst FC, Koot HM, der Ende JV. Differential predictive value of parents' and teachers' reports of children's problem behaviors: A longitudinal study. Journal of Abnormal Child Psychology. 1994; 22(5):531–546. doi:10.1007/BF02168936. [PubMed: 7822627]

Waber DP, Gerber EB, Turcios VY, Wagner ER, Forbes PW. Executive functions and performance on high-stakes testing in children from urban schools. Developmental Neuropsychology. 2006; 29(3):459–477. doi:10.1207/s15326942dn2903_5. [PubMed: 16671862]

Wagner, RK., Torgesen, JK., Rashotte, CA. Comprehensive test of phonological processing: CTOPP. Pro-ed; Austin, TX: 1999.

Walsh, KW. Neuropsychology: A clinical approach. Churchill Livingstone; Oxford, UK: 1978.

Wechsler, D. Wechsler Individual Achievement Test. Psychological Corporation; San Antonio, TX: 1992.

Wechsler, D. WAIS-III: Administration and scoring manual: Wechsler Adult Intelligence Scale. Psychological Corporation; San Antonio, TX: 1997.

Woodcock, RW., McGrew, KS., Mather, N. The Woodcock-Johnson III tests of achievement. Riverside Publishing; Illinois: 2001.

Yeniad N, Malda M, Mesman J, van Ijzendoorn MH, Pieper S. Shifting ability predicts math and reading performance in children: A meta-analytical study. Learning and Individual Differences. 2013; 23:1–9. doi:10.1016/j.lindif.2012.10.004.

Table 1

Demographic characteristics.

Variable	(*n* = 93)
Age in years: Mean (*SD*)	10.91 (0.72)
Grade	58.06% fourth grade
Gender	55.91% female
Ethnicity	38.71% African American
	51.61% Hispanic
	9.68% Other
Education Program	5.38% Special Education
	8.60% Gifted and Talented

Table 2

Descriptive statistics for predictor and outcome variables ($n = 93$).

Variable	Category/Scale	Mean	*SD*	Skewness	Kurtosis
Covariates					
WJ-III Oral Comprehension					
Raw Score	0–34	18.33	4.32	−0.19	−0.51
T-Score		94.80	14.19	−0.30	−0.11
CTOPP Elision					
Raw Score	0–20	13.08	4.81	−0.24	−1.38
Scaled Score		8.43	3.11	0.04	−1.12
Behavioral Rating Scales					
BRIEF-T Working Memory					
Raw Score	10–30	17.50	6.13	0.36	−0.94
T-Score		65.08	17.13	0.26	−1.03
BRIEF-T Plan/Organize					
Raw Score	10–30	18.22	6.13	0.18	−1.12
T-Score		63.95	15.52	0.16	−1.09
BRIEF-T Inhibit					
Raw Score	10–30	16.20	6.28	0.73	−0.66
T-Score		63.39	20.18	1.05	0.23
BRIEF-T Shift					
Raw Score	10–30	15.71	5.63	0.87	0.14
T-Score		63.44	18.61	0.93	0.54
Cognitive Measures					
WMBT-C Listening Recall					
Total Raw Score	0–36	13.46	3.99	0.36	0.94
Standard Score		101.71	19.80	−0.23	−0.25
Tower of London					
Total Raw Score	0–26	17.30	3.48	−0.78	1.54
DKEFS CWIT Inhibition					
Total Time	0–180	69.07	19.98	0.68	0.28

Variable	Category/Scale	Mean	*SD*	Skewness	Kurtosis
Scaled Score		10.72	2.70	−0.76	0.32
DKEFS TMT Number-Letter Switching					
Total Time	0–240	143.70	55.14	0.30	−0.98
Scaled Score		7.29	4.08	−0.22	−1.14
Outcome Measures					
Gates-MacGinitie Reading Comprehension					
Raw Score	0–48	27.34	10.01	0.04	−1.17
Standard Score		96.26	12.90	0.11	−0.12
WJ-III Math Calculations					
Raw Score	0–45	20.59	3.78	−0.07	0.44
Standard Score		106.01	12.79	−0.15	−0.29

Note. BRIEF-T = Behavior Rating Inventory of Executive Function – Teacher Form; CTOPP = Comprehensive Test of Phonological Processing; CWIT = Color Word Interference Test; DKEFS = Delis–Kaplan Executive Function System; SD = standard deviation; TMT = Trail Making Test; WJ-III = Woodcock–Johnson III; WMBT-C = Working Memory Test Battery for Children.

Table 3

Aim 1: Correlations between executive function measures and academic outcomes.

		Cognitive Measures					
		WM	**Plan**	**Inhibit**	**Shift**	**Math**	**Reading**
Behavioral Rating Scales	WM	−0.20				−0.42**	−0.55**
	Plan		−0.19			−0.40**	0.52**
	Inhibit			*0.25***		*− 0.29***	−0.38*
	Shift				*0.25***	−0.36*	−0.49**
Math		0.49**	0.40**	−0.57**	−0.50**		0.55**
Reading		0.55**	0.37*	−0.32*	−0.32*		

* *Note*. $p < .05$;

** $p < .001$.

WM = working memory. Values in italics no longer significant after correction ($p < .002$).

Table 4

Aim 1: Correlations between cognitive and behavioral rating scale executive function measures.

		Behavioral Rating Scales				Cognitive Measures			
		WM	**Plan**	**Inhibit**	**Shift**	**WM**	**Plan**	**Inhibit**	**Shift**
Behavioral Rating Scales	WM	1	0.91 **	0.76 **	0.72 **	−0.20	_− 0.24_ *	0.40 **	_0.31_ *
	Plan	0.90 **	1	0.74 **	0.76 **	−0.12	−0.18	0.32 *	_0.24_ *
	Inhibit	0.76 **	0.73 **	1	0.84 **	−0.04	_− 0.24_ *	_0.25_ *	0.19
	Shift	0.72 **	0.76 **	0.84 **	1	−0.08	_− 0.23_ *	_0.30_ *	_0.25_ *
Cognitive Measures	WM	−0.20	−0.13	−0.04	−0.08	1	_0.24_ *	−0.37 **	−0.36 *
	Plan	_− 0.25_ *	−0.19	_− 0.24_ *	_− 0.23_ *	_0.24_ *	1	_− 0.25_ *	_− 0.22_ *
	Inhibit	0.38 **	_0.30_ *	_0.25_ *	_0.29_ *	−0.37 **	_− 0.25_ *	1	0.48 **
	Shift	_0.32_ *	_0.24_ *	0.19	_0.25_ *	−0.36 *	_− 0.22_ *	0.47 **	1

* _Note._ $p < .05$;

** $p < .001$.

WM = working memory. Bottom left = Pearson correlations, top right = partial age correlations. Values in italics no longer significant after correction.

Table 5

Aim 2: Regression statistics for each EF process and reading comprehension.

Predictor	*b*	*SE*(*b*)	*β*	*t*	*p*	*sr* 2
Covariate Model						
Intercept	41.61	12.12	0.00	3.43	<.001	
Age	−3.42	0.98	−0.25	−3.49	<.001	.06
WJ-III Oral Comp	0.72	0.18	0.31	4.01	<.001	.07
CTOPP Elision	0.73	0.15	0.35	4.69	<.001	.10
Special Education	−5.55	3.13	−0.13	−1.77	.08	.01
Gifted/Talented	8.44	2.65	0.24	3.19	<.001	.05
Total Model $R^2 = .60$, $F(5, 87) = 26.34$, $p < .001$						
Working Memory						
Intercept	52.28	11.48	0.00	4.55	<.001	
Age	−3.43	0.91	−0.25	−3.78	<.001	.06
WJ-III Oral Comp	0.38	0.20	0.16	1.89	.063	.01
CTOPP Elision	0.52	0.15	0.25	3.45	.001	.05
Special Education	−3.78	2.94	−0.09	−1.29	.202	.01
Gifted/Talented	5.96	2.56	0.17	2.33	.022	.02
BRIEF-T WM	−0.45	0.12	−0.28	−3.84	<.001	.06
WMTB-C	0.47	0.22	0.19	2.16	.034	.02
Total Model $R^2 = .67$, $F(7, 85) = 24.66$, $p < .001$						
$R^2\Delta = .07$, $F(2, 85) = 8.74$, $p < .001$						
Planning						
Intercept	41.41	12.07	0.00	3.43	.001	
Age	−2.99	0.95	−0.22	−3.17	.002	.04
WJ-III Oral Comp	0.64	0.17	0.28	3.71	<.001	.06
CTOPP Elision	0.56	0.16	0.27	3.59	.001	.05
Special Education	−3.59	3.05	−0.08	−1.18	.243	.01
Gifted/Talented	7.21	2.56	0.20	2.82	.006	.03
BRIEF-T Plan/Organize	−0.34	0.12	−0.21	−2.85	.006	.03
TOL	0.31	0.20	0.11	1.55	.125	.01

Predictor	*b*	*SE*(*b*)	*β*	*t*	*p*	sr^2
Total Model R^2 = .65, *F*(7, 85) = 22.24, *p* < .001						
$R^2\Delta$ = .05, *F*(2, 85) = 5.37, *p* = .006						
Inhibition						
Intercept	55.38	12.52	0.00	4.42	<.001	
Age	−3.64	0.95	−0.26	−3.85	<.001	.06
WJ-III Oral Comp	0.71	0.17	0.31	4.11	<.001	.07
CTOPP Elision	0.59	0.15	0.29	3.81	<.001	.06
Special Education	−4.37	3.03	−0.10	−1.44	.153	.01
Gifted/Talented	6.94	2.60	0.20	2.67	.009	.03
BRIEF-T Inhibit	−0.23	0.11	−0.15	−2.08	.041	.02
DKEFS CWIT	−0.07	0.04	−0.14	−1.97	.052	.02
Total Model R^2 = .64, *F*(7, 85) = 21.88, *p* < .001						
$R^2\Delta$ = .04, *F*(2, 85) = 4.87, *p* = .01						
Shifting						
Intercept	50.60	12.12	0.00	4.17	<.001	
Age	−3.33	0.94	−0.24	−3.54	.001	.05
WJ-III Oral Comp	0.71	0.17	0.31	4.15	<.001	.07
CTOPP Elision	0.51	0.17	0.25	3.08	.003	.04
Special Education	−3.75	3.06	−0.08	−1.22	.224	.01
Gifted/Talented	8.01	2.55	0.23	3.14	.002	.04
BRIEF-T Shift	−0.41	0.13	−0.23	−3.07	.003	.04
DKEFS TMT	−0.01	0.01	−0.03	−0.40	.693	<.01
Total Model R^2 = .64, *F*(7, 85) = 21.91, *p* < .001						
$R^2\Delta$ = .04, *F*(2, 85) = 4.90, *p* = .01						

Note. *b* = parameter estimate; BRIEF-T = Behavior Rating Inventory of Executive Function – Teacher Form; *β* = standardized estimate; CTOPP = Comprehensive Test of Phonological Processing; CWIT = Color Word Interference Test; DKEFS = Delis-Kaplan Executive Function System; *SE*(*b*) = standard error of parameter estimate; sr^2 = squared semi-partial correlation; TMT = Trail Making Test; TOL = Tower of London; WJ-III Oral Comp = Woodcock–Johnson III Oral Comprehension; WM = working memory; WMBT-C = Working Memory Test Battery for Children. $R^2\Delta$ accounts for the addition of EF variables as a whole to the model with covariates alone.

Table 6

Aim 2: Regression statistics for each EF process and math calculations.

Predictor	*b*	*SE*(*b*)	*β*	*t*	*p*	*sr*²
Covariates						
Intercept	16.91	0.90	0.00	18.82	<.001	
CTOPP Elision	0.26	0.07	0.34	3.97	<.001	.10
Special Education	−3.46	1.34	−0.21	−2.59	.011	.04
Gifted/Talented	5.44	1.12	0.41	4.87	<.001	.15
Total Model R^2 = .45, $F(3, 89)$ = 24.40, p < .001						
Working Memory						
Intercept	17.37	1.66	0.00	10.48	<.001	
CTOPP Elision	0.18	0.07	0.23	2.58	.012	.04
Special Education	−2.66	1.32	−0.16	−2.01	.048	.02
Gifted/Talented	4.21	1.16	0.31	3.63	.001	.08
BRIEF-T WM	−0.11	0.05	−0.17	−2.05	.043	.02
WMTB-C	0.19	0.09	0.20	2.24	.028	.03
Total Model R^2 = .50, $F(5, 87)$ = 17.64, p < .001						
$R^2\Delta$ = .05, $F(5, 87)$ = 4.57, p = .013						
Planning						
Intercept	17.63	1.90	0.00	9.30	<.001	
CTOPP Elision	0.23	0.06	0.34	3.79	<.001	.08
Special Education	−4.20	1.68	−0.19	−2.50	.015	.04
Gifted/Talented	4.22	0.98	0.35	4.34	<.001	.11
BRIEF-T Plan/Organize	−0.09	0.05	−0.16	−1.95	.055	.02
TOL	0.10	0.08	0.10	1.24	.219	<.01
Total Model R^2 = .54, $F(5, 79)$ = 18.43, p < .001						
$R^2\Delta$ = .09, $F(5, 79)$ = 5.54, p = .032						
Inhibition						
Intercept	23.41	1.69	0.00	13.86	<.001	
CTOPP Elision	0.21	0.06	0.27	3.35	.001	.06
Special Education	−2.83	1.21	−0.17	−2.34	.022	.03

Predictor	b	$SE(b)$	β	t	p	sr^2
Gifted/Talented	4.21	1.03	0.31	4.08	<.001	.08
BRIEF-T Inhibit	−0.02	0.05	−0.03	−0.35	.731	<.01
DKEFS CWIT	−0.07	0.02	−0.37	−4.77	<.001	.11
Total Model R^2 = .57, F(5, 87) = 23.04, p < .001						
$R^2\Delta$ = .12, F(5, 87) = 7.13, p = .001						
Shifting						
Intercept	21.75	1.78	0.00	12.22	<.001	
CTOPP Elision	0.17	0.07	0.22	2.41	.018	.03
Special Education	−2.75	1.28	−0.17	−2.15	.035	.03
Gifted/Talented	4.94	1.06	0.37	4.68	<.001	.12
BRIEF-T Shift	−0.05	0.06	−0.07	−0.86	.393	<.01
DKEFS TMT	−0.02	0.01	−0.29	−3.60	.001	.07
Total Model R^2 = .53, F(5, 87) = 19.50, p < .001						
$R^2\Delta$ = .08, F(5, 87) = 11.98, p < .001						

Note. b = parameter estimate; BRIEF-T = Behavior Rating Inventory of Executive Function – Teacher Form; β = standardized estimate; CTOPP = Comprehensive Test of Phonological Processing; CWIT = Color Word Interference Test; DKEFS = Delis-Kaplan Executive Function System; $SE(b)$ = standard error of parameter estimate; sr^2 = squared semi-partial correlation; TMT = Trail Making Test; TOL = Tower of London; WJ-III Oral Comp = Woodcock–Johnson III Oral Comprehension; WM = working memory; WMBT-C = Working Memory Test Battery for Children. $R^2\Delta$ accounts for the addition of EF variables as a whole to the model with covariates alone. Influential residuals removed from planning model.

Table 7

Aim 3: Regression statistics for model with combined EF measures and reading comprehension and math calculations.

Predictor	*b*	*SE(b)*	*β*	*t*	*p*	*sr* ^{2}
Reading Comprehension						
Intercept	55.39	12.44	0.00	4.45	<.001	
Age	−4.15	0.91	−0.31	−4.56	<.001	.09
WJ-III Oral Comp	0.43	0.20	0.18	2.12	.038	.02
CTOPP Elision	0.45	0.15	0.22	2.91	.005	.04
Special Education	−4.30	3.53	−0.08	−1.22	.227	.01
Gifted/Talented	6.04	2.53	0.17	2.38	.020	.02
BRIEF-T WM	−0.40	0.12	−0.25	−3.20	.002	.04
WMTB-C	0.61	0.24	0.24	2.59	.011	.03
TOL	0.17	0.20	0.06	0.86	.390	<.01
DKEFS CWIT	−0.03	0.04	−0.07	−0.83	.411	<.01
DKEFS TMT	0.01	0.01	0.07	0.89	.374	<.01
Total Model R^2 = .68, F(10, 76) = 16.43, p = .001						
$R^2\Delta$ = .08, F(10, 76) = 3.36, p < .001						
Math Calculations						
Intercept	20.43	2.11	0.00	9.68	<.001	
CTOPP Elision	0.14	0.05	0.20	2.60	.011	.03
Special Education	−1.71	1.08	−0.10	−1.58	.118	.01
Gifted/Talented	3.55	0.88	0.30	4.03	<.001	.07
BRIEF-T WM	−0.04	0.04	−0.07	−1.02	.313	<.01
WMTB-C	0.08	0.07	0.09	1.17	.245	.01
TOL	0.16	0.07	0.16	2.37	.020	.02
DKEFS CWIT	−0.05	0.01	−0.27	−3.50	<.001	.05
DKEFS TMT	−0.01	0.01	−0.13	−1.72	.089	.01
Total Model R^2 = .68, F(8, 79) = 21.08, p < .001						
$R^2\Delta$ = .23, F(8, 79) = 9.46, p < .001						

Note. *b* = parameter estimate; BRIEF-T = Behavior Rating Inventory of Executive Function – Teacher Form; *β* = standardized estimate; CTOPP = Comprehensive Test of Phonological Processing; CWIT = Color Word Interference Test; DKEFS = Delis-Kaplan Executive Function System; *SE*(*b*) = standard error of parameter estimate; sr^2 = squared semi-partial correlation; TMT = Trail Making Test; TOL =

Tower of London; WJ-III Oral Comp = Woodcock–Johnson III Oral Comprehension; WM = working memory; WMBT-C = Working Memory Test Battery for Children. $R^2\Delta$ accounts for the addition of EF variables as a whole to the model with covariates alone. Influential residuals removed from presented models.