Journal of Educational and Developmental Psychology; Vol. 13, No. 2; 2023
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Executive Functioning and Adolescents' Academic Performance on Standardized Exams

William Ellery Samuels[1], Nelly Tournaki[2], Stanley Sacks[3], Sheldon Blackman[4], Theresa Peterford[4], Jo Ann Sacks[3] & Kenneth Byalin[4]

[1] School of Nursing, Hunter College, Brookdale Campus, New York, New York 10010, USA

[2] Department of Educational Studies, College of Staten Island, The City University of New York, New York 10010, USA

[3] National Development and Research Institutes, Inc., New York, New York 10010, USA

[4] Integration Charter Schools, New York, New York 10010, USA

Correspondence: William Ellery Samuels, School of Nursing, Hunter College, Brookdale Campus, Room 424 West, 425 East 25th Street, New York, New York 10010, USA. E-mail: ws1114@hunter.cuny.edu

Received: March 25, 2023 Accepted: May 15, 2023 Online Published: June 20, 2023

doi:10.5539/jedp.v13n2p10 URL: http://doi.org/10.5539/jedp.v13n2p10

Abstract

Executive functions (EFs) help regulate and direct thoughts, behaviors, and emotions. They also play vital roles in many areas of life. However, few studies address the role EFs play in adolescents' lives, including their academic performance. We investigated the effects of EFs on standardized exams in mathematics, reading, and English language arts. The main findings were that: 1) adolescents' EFs—especially when measured by their current teachers—predict performance on standardized academic assessments throughout the middle and high school grades; 2) this effect existed among a rather diverse sample of students both with and without diagnosed disabilities; 3) the predictiveness of EFs tended to increase across these grades when measured by the teachers, but not those gauged by the students themselves; and 4) EFs were somewhat more strongly associated with performance on standardized reading and English language arts exams than on math exams. In addition, students who identified as female tended to show stronger EFs; race/ethnicity showed some significance, but not in easily interpreted ways; and age was reliably associated with performance on these standardized exams such that older students tended to do better even though the exam scores were standardized by grade level. The results illustrate the contributions of EFs to standardized assessments for students with and without diagnosed disabilities.

Keywords: Executive functioning, standardized examination, academic achievement, individualized education programs (IEPs), middle school, high school, longitudinal research

1. Executive Functioning and Adolescents' Academic Performance on Standardized Exams

Executive functions (EFs) can be generally defined as a set of cognitive and behavioral control processes that individuals use to regulate and direct attention, memory, thoughts, emotional reactions, and behaviors so that they may attain both short- and long-term goals (Blair & Raver, 2012; Diamond, 2013; Gioia et al., 2015). These abilities to direct one's attention and behavior towards meeting a goal are necessary to complete most academic tasks. It is not surprising then that EFs present an area of growing interest among education researchers. The varied demands of reading (decoding and synthesizing lexical and phonological aspects) require the development and coordination of various executive functions (Ober et al., 2020). It is not surprising then that both Sesma et al. (2009) and Zelazo and Carlson (2012) contend that EFs are fundamental for language development and thus for literacy—itself a foundation for learning—as well as for the processing and organization of received information.

The main purpose of the present study is to examine the relationship between EFs and academic performance on standardized testing among adolescents. The analyses also take into consideration variables such as students' age, gender, disability status, race/ethnicity and socioeconomic status (SES). Prior research has established the relationship between EFs and assessment of academic performance through teacher grades (Samuels et al., 2016, 2019).

A recent meta-analysis on primarily elementary-aged students (Cortés Pascual et al., 2019) included 21 studies (total $N = 7{,}947$) worldwide. The results led the authors to conclude that EFs are a good predictor of academic performance (mean effect size = 0.36 ± 0.055), slightly more strongly for mathematics than language. They further reported that the EFs component that had the highest predictive weight for academic performance was working memory (compared to inhibition, cognitive flexibility, and planning). Cortés Pascual et al. also investigated potential moderating effects of gender and age, finding that gender could account for 49% of the variance in the relationship between EFs and academic performance; age, however, was not a significant moderator. These findings are especially important since age has traditionally been seen as a strong moderator of EFs (Best et al., 2009; Buckner et al., 2009; DePasquale et al., 2021; Lee et al., 2013; Karr et al., 2022).

Deer et al. (2020) and Gordon et al. (2018) indeed found that individual differences in EFs predict academic success across various developmental periods. López (2013) also found that academic results in language and mathematics are related to EFs among third graders.

Others (Best et al., 2011; Dubuc et al., 2020; Dutra et al., 2022; Nutley & Soderqvist, 2017; Ostrosky-Solís et al., 2007) have looked at particular EFs, especially working memory, a main component of EFs. Working memory develops rapidly during childhood and then much more slowly during adolescence (Huizinga & van der Molen, 2007). For example, a longitudinal study conducted by Ahmed et al. (2018) found that working memory at 54 months significantly predicts working memory at 15 years. Tsubomi and Watanabe (2017) also found that visual working memory develops rapidly until 10 years of age, while Hall et al. (2015) reported that primary memory capacity improves among children 5–8 years old.

There is good evidence that working memory well predicts academic performance in elementary school students (e.g, Albert et al., 2020; Artuso et al., 2019). However, this conclusion cannot be drawn with certainty for adolescents because the predictive power of this variable diminishes at around the age of 12 years (e.g., Aronen et al., 2005; Best et al., 2011; Lee et al., 2009; St Clair-Thompson & Gathercole, 2006).

So, while there is considerable evidence that EFs predict academic performance in the early years, there is a need to further examine these variables in adolescence. Theodoraki et al. (2020, p. 74) stated that "[t]he development of executive functions (EFs) has primarily been studied among younger children, despite research suggesting that particular aspects of EFs continue to develop throughout adolescence and into adulthood." In general, EFs are found to be associated with adolescent students' academic success (Best et al., 2011; Bierman et al., 2009; Kotsopoulos & Lee, 2012; Samuels et al., 2016, 2019; Vuontela et al., 2013; Waber et al., 2006). This is true for mathematics (Andersson 2008; Bull & Lee, 2014; Lee et al., 2009; van der Ven et al., 2012) as well as reading, writing, and science (Monette et al., 2011; St Clair-Thompson & Gathercole, 2006). Most studies, however, have included small samples; the few large studies on diverse, adolescent populations do not yet present a coherent picture (Ahmed et al. 2018; Best et al., 2011; Samuels et al., 2016, 2019). We therefore investigated predictive qualities of EFs in academic performance, an important form of success during adolescence.

The term "academic performance" can refer to a range of educational outcomes expressed through assessments such as grade point average (GPA), high school graduation rates, standardized test scores, ad hoc performance assessments, and college entrance exam scores. Types of assessments have long been a controversial issue among educators and researchers. Although GPA is the most-used measure of academic success (York et al., 2015), there is disagreement about what exactly are its function and importance; validity is important because the purpose of GPA is to accurately communicate to others the level of academic achievement that a student has obtained (Snowman & Biehler, 2003). Grades—and thus GPAs—are imperfect measures of a student's academic achievement that cannot convey all—or even sometimes sufficient—information about the level of the student's academic performance (Chiekem, 2015).

An alternative or complement to GPAs as indicators of academic performance are standardized tests that offer more generalizable assessments of student performance. In addition to allowing readier comparisons to students in other classes, schools, cities, and regions (Allensworth & Clark, 2020), they can guide curricular and content area choices, thus helping teachers align classroom materials with larger goals, including state and national content standards. They sometimes are also used as measures of accountability for what is learned within and between classes. Like any single assessment, standardized exams cannot measure every aspect of learning or thinking, but they do provide a more universal—and perhaps even more objective (Willingham et al., 2002)—standard for appraising that which they intend to measure. Zwick and Greif Green (2007), for example, found SAT scores much less affected by socioeconomic factors than high school grades.

Much work has been devoted to studying the predictiveness of standardized exams administered in highs school and GPA on subsequent success in college (e.g., Burton & Ramist, 2001; Kobrin et al., 2008). The results of these studies are mixed—and certainly neither standardized exams scores nor GPAs account for all that contributes to college success (Hannon, 2014); nor do either do so equally for all students (Allensworth & Clark, 2020; Hannon

& McNaughton-Cassill, 2011; Kobrin & Patterson, 2010). Nonetheless, standardized scores and high school GPAs are frequently found to be among the best predictors of academic performance in the first year of college (Burton & Ramist, 2001; Kobrin et al., 2008; Richardson et al., 2012; Sackett et al., 2012).

Although standardized exam scores and high school GPAs are correlated (Zwick et al., 2007), they appear to measure somewhat different aspects of academics. Frey and Detterman (2004) found that Scholastic Aptitude Test (SAT) scores correlated well (*r*s ≈ .8) with measures of general cognitive ability (*g*), while McNeish et al. (2015) reported that non-cognitive factors (such as parental involvement and study habits) were more strongly correlated with GPA than with American College Testing (ACT) scores. Noting that high school GPAs predict college graduation rates better than standardized exam scores (viz., the ACT & SAT), Galla et al. (2019) found evidence that this was at least in part due to GPAs being more strongly associated with self-regulation while standardized exam scores were more strongly associated with cognitive ability.

It may well be then that EFs predict performance on standardized exams. EFs have been found to be associated with both cognitive ability (at least in children, Foley et al., 2009) and middle and high school GPAs (Samuels et al., 2019). To the best of our knowledge, however, this association between EFs and standardized scores among *adolescents* has not been formally investigated; were EFs to predict performance on standardized assessments of academic performance, then their importance in educational interventions would be further supported. Also to the best of our knowledge, only three studies have reported correlations between EFs and standardized tests: Waber et al.'s (2006) and Wagner et al.'s (2019) samples included only children while Best et al. (2011) investigated the relationships between EFs and academic achievement among a sample of over 2,000 children and adolescents using the Planning scale of the Cognitive Assessment System (Naglieri & Das, 1997) and a standardized academic test; Best et al. found that EFs were moderately correlated with success in both math and reading achievement. There is therefore a need for research that explores the relationship between EFs and standardized tests as indicators of academic performance.

In this study, adolescent students' EFs were measured by asking them to complete the student-reported Behavior Rating Inventory of Executive Function–Self-Report (BRIEF-SR; Gioia et al., 2000) and also by asking their teachers to complete the teacher-reported Behavior Rating Inventory of Executive Function (BRIEF) about those same students. Versions of the BRIEF have been to study EFs in academic settings, but most of these studies used versions of the BRIEF to investigate relationships among children (e.g., Clark et al., 2010; Locascio et al., 2010). Only, Langberg et al. (2013) and Samuels et al. (2019) used versions of the BRIEF to study these relationships among adolescents. Langberg et al. investigated academic outcomes among ~100 adolescents diagnosed with ADHD and found that teacher-rated scores on the Plan/Organize subscale of the BRIEF significantly contributed to the prediction of these students' overall GPAs beyond that made by the number of parent-reported ADHD symptoms. Samuels et al. found that both BRIEF and BRIEF-SR scores predicted overall GPAs above and beyond what was already predicted by students' demographic variables (gender, IEP status, and eligibility for free / reduced school lunch), supporting similar results found by Albert et al. (2020) that EFs (measured through a battery of different instruments) can partially explain differences in academic achievement between adolescents from lower and higher socioeconomic statuses.

The BRIEF-SR has been used less frequently than the BRIEF in research. Boschloo et al. (2014) did not find a significant relationship between some sub-scores on a Dutch version of the BRIEF-SR and grades in Dutch, English, and mathematics; they also did not find that grades were predicted by behavioral measures on EFs from the Delis-Kaplan Executive Functions System. It may be that studies like that reported by Boschloo et al. represent similar null findings that others find and do not publish, or that adolescents' insights into their own EFs remain an understudied area. Adolescents have been found to be able to rate their own behaviors accurately (Wichstrøm, 1995); nonetheless, individuals of many ages who are still developing an ability are often not so good at rating themselves on that ability (Dunning et al. 2003), and being able to monitor aspects of one's own performance is itself an EF, so adolescents who are still developing the ability to monitor their own behaviors may not be so able to accurately rate themselves. One of the goals of the present study is to investigate the relationship of BRIEF and BRIEF-SR in predicting academic performance, comparing them against each other and conducting an initial foray into the role of self-monitoring on the predictive aspects of the BRIEF-SR's validity here.

In the present longitudinal study, we investigated the association between standardized tests of academic performance (in math, reading, and English language arts) and EFs among adolescents throughout those students' careers at one middle and one high school. The sample included students with and without disabilities sharing inclusive classrooms—a special feature of the school. The disability status of each student was controlled in the analysis along with students' age, gender, race/ethnicity and SES. We considered it important to focus on EFs since most of the studies over the last decade have found this variable to be more significant for academic performance than the intelligence quotient, the variable traditionally viewed to be the best predictor of academic success (Ren et al., 2015).

1.1 Goals and Hypotheses

The primary goals of the current study were to:

(1) investigate the predictive validity of the **BRIEF** by analyzing the contribution of BRIEF General Executive Composite (GEC) scores to predictions of academic performance on standardized tests among 6th to 12th graders,

(2) examine the predictive validity of the **BRIEF-SR** by analyzing the contribution of BRIEF-SR GEC scores to the predictions of these same outcomes,

(3) compare the contributions of the BRIEF with those of the BRIEF-SR for their uses as experimental tools and in light of relevant demographic variables (age, gender, race/ethnicity, SES, and disability), and

(4) evaluate the results for both components of EF: metacognition and behavioral regulation.

We hypothesized that BRIEF and BRIEF-SR GEC scores would predict performance on a standardized test among adolescent students. However, we anticipated that the valid predictive use of BRIEF-SR GEC scores here may not be as well supported (i.e., will not predict academic performance as well as BRIEF GEC scores) given the equivocal findings on the use of the BRIEF-SR in academic settings outlined earlier.

2. Methods

2.1 Participants/Demographics

With informed consent and IRB approval, data were collected in a charter school of New York City. The school was chosen because it provides an inclusive environment for all students. It welcomes all students, including those with disabilities and integrates them in all classrooms and activities. For further context, it is worth noting that the school's mission statement indicates that full integration of students empowers them to break down barriers through the power of their daily academic and social experience, enabling them to develop the academic skill, emotional fluency, and confidence required to be successful students today and thoughtful, open-minded leaders tomorrow.

Therefore, all students attended the same middle school and then the same high school. The students at these two schools are diverse and represent a general—if under-served—population, but are all experiencing the same overall academic environment. We would have liked to collect data from other schools as well, but the intensity of the yearly data collection and breadth of the data collected have so far prevented us from being able to doing so at more schools.

Scantron© Scaled Scores in mathematics were available for 688 students, in reading for 719 students, and in Language Arts for 717. For this group, which spanned grades 6–12. The mean age for the students with at least one Scaled Score and BRIEF / BRIEF-SR score was 13.57 (*SD* = 1.91) years. Forty-seven percent identified as female; 40% identified as African-, 4% as Asian-, 12% as European-, and 33% as Latin-American. (Race and ethnicity were recorded as exclusive categories, with only 1% identifying as members of multiple races.) Thirty-four percent were classified as having a disability and had individualized education programs (IEPs). The classifications included in the IEPs were: social emotional impairment, autism spectrum disorder, learning disability, other health impairment, speech and language impairment, intellectual disability.

2.2 Instruments

2.2.1 The Behavior Rating Inventory of Executive Function (BRIEF)

The Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000) is an 86-item instrument developed to assess—via parent and/or teacher reports—EF manifestations in the everyday lives of children and adolescents aged 5–18 years. The BRIEF has been widely used in clinical applications as well as in a variety of research studies involving children and adolescents who are typically and atypically developing (for reviews see Isquith et al., 2013; Roth et al., 2014). Researchers who examined the discriminant validity of the BRIEF reported that it successfully differentiates between children and adolescents with and without ADHD (Reddy et al., 2011; Toplak et al., 2008). It has been widely used to assess outcomes following a variety of interventions (Isquith et al., 2013) and is associated with academic performance (Clark et al., 2010; Langberg et al., 2013; Locascio et al., 2010; Roth et al., 2014; Samuels et al., 2016, 2019).

The BRIEF has demonstrated good inter-item and test-retest reliability (Gioia et al., 2000). It has also been found to be a practical tool showing valid uses in school and clinical settings as well as in research; there are over 400 peer-reviewed publications supporting the reliability, clinical utility, and valid uses of the BRIEF, mostly among children. Overall, reviews of the BRIEF have been positive (Baron, 2000; Goldstein, 2001; Strauss et al., 2006). To our knowledge, however, no studies have yet investigated its use to predict performance on standardized exams in schools and among diverse, community-dwelling adolescents.

2.2.2 The Behavior Rating Inventory of Executive Function–Self-Report Version (BRIEF-SR)

The Behavior Rating Inventory of Executive Function–Self-Report Version (BRIEF-SR) offers another method to measure EFs among older children and adolescents. The BRIEF-SR is designed for those aged 11–18 years to self-report the frequency of various EF-related behaviors through 80 items that measure nearly the same domains as the BRIEF (Guy et al., 2004). The use of the BRIEF-SR may therefore allow for investigations of EFs among adolescents while relying on a different source for information that may reduce the burden on any one participant while also providing a complimentary—or perhaps even an alternate—vehicle for measurement.

Guy et al. (2004) provided evidence for the BRIEF-SR's ability to validly measure EFs, including through its relationship with the Behavior Assessment System for Children Parent Rating Scales (BASC-PRS) and with Teacher Rating Scales (BASC-TRS)—but, importantly, not directly against the BRIEF. Indeed, few studies have compared versions of the BRIEF side-by-side with the BRIEF-SR, but the present study undertakes this task. The uses of the BRIEF and BRIEF-SR here have been purchased from PAR, Inc., their publisher.

2.2.3 Structure of the BRIEF and BRIEF-SR

The BRIEF and BRIEF-SR were constructed to measure two general areas of EF: Metacognition and Behavioral Regulation (Gioia et al., 2000; Guy et al., 2004), themselves each comprised of further subscales that interrelate but nonetheless represent different executive functions (Karr et al., 2022; Keller et al., 2023). Exploratory factor analyses of the eight subscale divisions of the parent and teacher forms of BRIEF showed the same two-factor solution in both normal controls and specific clinical subjects (Gioia et al., 2000). The metacognition and behavioral regulation areas can be combined to create an overall Global Executive Composite (GEC) score.

As operationalized by the BRIEF, metacognition includes the "ability to initiate, plan, organize, and sustain future-oriented problem solving in working memory" (Gioia et al., 2000, p.20). Behavioral regulation involves the "ability to shift cognitive set and modulate emotions and behavior via appropriate inhibitory control" while allowing "metacognitive processes to successfully guide active, systematic problem solving (and supports) appropriate self-regulation" (p.20).

2.2.4 General Administration

Both the BRIEF and BRIEF-SR were administered once per academic year (AY), within a few weeks of the end of the AY. All students completed the BRIEF-SR on the same day during the same class (Wellness, a course that roughly corresponds to a combined Health and Civics course). Teachers of this course completed the BRIEF for each of their students within one week of when the students completed the BRIEF-SR.

2.3 Assessment of Academic Performance Through Scantron© Scaled Scores

Students were assessed in reading, English language arts, and mathematics using the Scantron© Performance Series, a series of standardized assessments of various content areas, including of course math, reading, and English language arts. The Performance Series are computer adaptive diagnostic assessments that each take between 45 and 60 minutes to complete and consists of 52–68 items, depending on how quickly a student's score can be determined.

Since the Performance Series use adaptive testing (the items and their difficulty levels changes based on a student's responses), one cannot use classical test theory's definition (often operationalized as Cronbach's α) to determine reliability. Instead, the exam is usually stopped once the standard error of measurement on an exam is less than a pre-determined threshold. Test-retest reliability for these exams tends to be reasonable (*r*s > .65; Scantron, 2015).

Evidence of the content-related validity of the Performance Series comes from their alignment with New York State Learning Standards (New York State Education Department, 2015–2016) and with standards proposed by national content pedagogy organizations including the National Council of Teachers of Mathematics (NCTM) and the National Council of Teachers of English (NCTE). Scores on the Performance Series mathematics exams also tend to correlate well with elementary and middle school students' performance on New York State mathematics tests (*r*s > .69; Scantron, 2015); correlations with other content areas were not available.

Among the available metrics generated by the Scantron Performance Series exams, we elected to focus on the Scaled Scores. Scaled Scores are standardized transformations of students' raw scores on any given administration of an exam. These scores are standardized using Rasch modeling of the item difficulties (Scantron, 2015). Scaled

Scores are grade-level independent, making them useful for comparing changes over time.

2.3.1 Administration

At the schools studied here, Performance Series exams were administered in each subject area usually twice per year; usually two different content area Performance Series exams are given, one during the fall and an other during the spring. In other words, students complete Scatron Performance Series exams on two occasions in an AY, but usually only complete a given content area exam at most once per year. There were instances when students took the same Scantron exam more than once in a given year; this would occur when there was a mis-administration, technical issues, or if the student or teacher felt that the performance on the exam was not indicative of their true ability and another chance to assess was given. Scantron indeed allows teachers to “spoil” an exam, halting it to allow students to retake it at a later time when either feels the exam will better measure the student’s ability. At this school, it is the administration that decides if an exam should be spoiled. This occurred on 7.9% of the times the tests were administered; on 6.6% ($n = 45$) of the occasions, a student took it one or two times more than prescribed, and on 1.3% ($n = 9$) occasions they took a given exam up to as many seven times more than the minimum. Since some students retook a given Scantron exam, averaged a student’s Scale Scores by academic year (AY) to insure the comparability of the data for each student.

There was an average of 3.43 AYs of data for Math Scaled Scores, 4.38 AYs for Reading Scaled Scores, and 4.16 AYs for Language Arts Scaled Scores. Note, however, that students did not complete a given content area Scantron exam every year; occasionally (~9% of the time), either the BRIEF or BRIEF-SR were not completed in a given year. Given this, students did not often have *both* a Scantron© Scaled Score *and* BRIEF or BRIEF-SR scores for the same AY. In nearly all cases (≥89%), data were available for both a Scaled Score and BRIEF / BRIEF-SR score for only one AY; otherwise, data were available for both Scantron and BRIEF / BRIEF-SR scores for two AYs, sufficient to perform the planned analyses exploring longitudinal effects.

2.4 Analytic Plan

To summarize the data collection procedure, the BRIEF and BRIEF-SR were administered to teachers and students respectively within a few weeks of the end of each academic year. The teachers completed the BRIEF for a random subset of students whom they taught that academic year; the students completed the BRIEF-SR about themselves. The BRIEF and BRIEF-SR were collected every year for each student, with different teachers completing it for a given student each academic year. Students completed the Scantron Performance Series exams as they were normally administered during the academic year, usually a given Scantron exam was administered once per AY in either semester.

All instruments collected during the same academic year were coded as being collected during the same AY. Although we knew the exact dates on which the instruments were measured, including this level of exactness did not improve the power of the analyses. The years of data were nested within student so that we could analyze within-student changes in both EFs and performance on the standardized exams.

Our investigation into the relationships between EFs and standardized scores included three general steps to strengthen the analysis. First, we assessed the simple (zero-order) correlations between BRIEF and BRIEF-SR scores and sub-scores with Scantron© Scaled Scores in math, reading, and English language arts; we also examined the relationships of important demographics within the models. Second, we used “predictive modeling” to examine the interactive effects of the variables. Third, we employed multilevel models (also commonly called as “hierarchical linear models”) to test the joint effects of the variables. All variables were either standardized (for continuous variables) or included as dummy variables (for nominal variables).

In these analyses, we first created a “base” model that included the demographic variables that were both found to be of interest through the zero-order correlations and that were also found to be of interest through previous research (e.g., Best et al., 2009; Best et al., 2011; Danielsson et al., 2010; Torske et al., 2017; Vogan et al., 2018). The demographic variables included in the base models were age, gender, race/ethnicity, SES, and IEP status.

These base models served as comparisons for the models that then also included EF-related terms: a main effect of the given EF (metacognition, etc.) and an EF-by-time interaction term. The EF-by-time interaction term tested whether any effect of the EF on that standardized score changed over time, e.g., if the EF become more impactful as the student matured.

Creating base models like this are commonly done so that one can not only test if a given term (e.g., EF-by-time interaction) is significant, but also if adding that and related terms (e.g., also the EF main effect) significantly improve the overall fit of the model. Improvements over the base model can let us look at a group of variables collectively and thus answer broader, theoretical questions. In other words, these base models lets us look first at non-EF factors and then focus specifically on EFs per se.

All times-varying terms (viz., age, EF scores, and standardized exam scores) were nested within students. Students, however, were not nested within their classrooms since this changed every year and since students could be grouped differently within a year for different courses.

All analyses were conducted with R, version 3.0.2 (R Core Team, 2013). R packages used included captioner (Alathea, 2015), dplyr (Wickham et al., 2022), knitr (Xie, 2022), Hmisc (Harrell, 2022), lmerTest (Kuznetsova, Brockhoff, Christensen, 2017), and psych (Revelle, 2022).

3. Results

This section presents the Results of all features of the analytic plan (Tables 1 – 4) and a summary of the main findings (Table 5).

3.1 Correlations

Table 1 presents the correlations with Scantron© Scaled Scores and BRIEF and BRIEF-SR sub-scores and total scores, IEP variables, and demographics. Teacher-reported BRIEF Global Executive Composite (GEC) scores—the overall measure of executive functioning—were strongly and significantly correlated with all three Scaled Scores (*r*s = -.38 – -.45) (Note 1) Student-reported BRIEF-SR GEC scores correlated weakly with the Scaled Scores; the correlation with Language Arts was moderate (r = -.16) while the correlations with Math and Reading were both low (*r*s ≈ -.08). BRIEF and BRIEF-SR GEC scores themselves correlated moderately (r = .19).

Table 1. Correlations between Scantron© Scaled Scores and Executive Functioning Scores, IEP Variables, and Demographic Variables. Bold-faced correlations are significant at α = .05

Variable Group	Variable	Math Scaled Score	Reading Scaled Score	Language Arts Scaled Score
Scantron© Scaled Scores	Math Scaled Score		**.69**	**.72**
	Reading Scaled Score	**.69**		**.75**
	Language Arts Scaled Score	**.72**	**.75**	
BRIEF	**Global Executive Composite**	**-.38**	**-.39**	**-.45**
	Behavioral Regulation Index	**-.35**	**-.37**	**-.44**
	Metacognitive Index	**-.39**	**-.39**	**-.46**
	Emotional Control	**-.36**	**-.36**	**-.43**
	Inhibit	**-.32**	**-.35**	**-.40**
	Initiate	**-.34**	**-.32**	**-.43**
	Monitor	**-.39**	**-.39**	**-.44**
	Organization of Materials	**-.33**	**-.32**	**-.39**
	Plan/Organize	**-.36**	**-.39**	**-.44**
	Shift	**-.32**	**-.37**	**-.43**
	Working Memory	**-.39**	**-.39**	**-.45**
BRIEF-SR	**Global Executive Composite**	**-.09**	**-.08**	**-.16**
	Behavioral Regulation Index	**-.03**	**-.09**	**-.13**
	Metacognitive Index	**-.13**	**-.07**	**-.17**
	Emotional Control	<.01	-.06	-.03
	Inhibit	**-.03**	**-.07**	**-.20**
	Monitor	**-.05**	**-.10**	**-.18**
	Organization of Materials	**-.09**	**-.03**	**-.16**
	Plan/Organize	**-.10**	**-.05**	**-.16**
	Shift	-.04	-.10	-.04
	Task Completion	**-.20**	**-.16**	**-.23**
	Working Memory	-.06	<.01	-.06

IEP Variables	Special Education Status	**-.34**	**-.36**	**-.36**
	Social Emotional Impairment	**-.06**	**-.02**	**-.04**
	Autism	-.08	-.11	-.09
	Learning Disability	**-.12**	**-.12**	**-.16**
	Other Health Impairment	-.01	<.01	.01
	Speech or Language Impairment	**-.17**	**-.28**	**-.22**
	Intellectual Impairment	**-.24**	**-.21**	**-.24**
Demographic Variables	Age	.23	.22	.08
	Gender	.01	**.07**	**.13**
	Lunch Status	-.01	.04	.02
	Hours Doing Homework	.06	**.33**	.16
	African	-.04	-.08	-.07
	Asian	**.12**	**.09**	**.10**
	European	.01	.02	**.06**
	Latin	-.10	-.10	-.14
	Multiracial	.02	-.01	.02
	Other Race/Ethnicity	.08	.06	.06

The GEC scores are comprised of two sub-scores—the Behavioral Regulation Index (BRI) and the Metacognitive Index (MCI)—that are themselves comprised of scores on individual EFs (Note 2). Both BRI and MCI demonstrated similar levels of correlation with Scaled Scores as did the GEC for both the BRIEF and the BRIEF-SR. The correlations of the BRIEF BRI and MCI scores with the Scaled Scores ranged from -.35 to -.48 (compared with GEC correlations of -.38 – -.45). BRIEF-SR BRI and MCI scores showed roughly comparable levels of correlation with Scaled Scores (MCI *r*s = -.07 – -.17; BRI *r*s = -.03 – -.13).

The various individual BRIEF EFs—Emotional Control, etc.—also showed relatively similar correlations with Scaled Scores. The correlations with Language Arts (*r*s = -.44 – -.48) tended to be stronger than those for either Reading (*r*s = -.33 – -.41) or Math (*r*s = -.32 – -.39), but no one EF stood out as especially-well correlated.

There was somewhat more variation among the individual BRIEF-SR EFs. Task Completion showed relatively stronger correlations than other executive functions with all three Scaled Scores (*r*s = -.16 – -.23). Inhibit, Monitor, Organization of Materials, and Plan/Organize all correlated more strongly with Language Arts Scaled Scores (*r*s = -.16 – -.20) than with either Math or Reading Scaled Score (*r*s = -.03 – -.10). Emotional Control and Working Memory were not significantly correlated with any of the Scaled Scores (largest *r* = -.06).

For demographic variables, students' genders correlated significantly with Language Arts and Reading (*r*s = .13 & .07, respectively) but not Math (*r* = .01) Scaled Scores, indicating that girls showed a small but significant tendency to out-perform boys on linguistic-related exams. Whether a student was eligible for free / reduced-priced school lunch was not significantly correlated. Identifying as a member of most races/ethnicities categorized here tended not to significantly correlate with Scaled Scores, except for identifying as Asian-American (*r*s = .09 – .12) or European-American's correlations with Language Arts scores (*r* = .06), although even these correlations were small.

In summary, EFs correlated well with performance on these standardized exams. Teacher-reported BRIEF scores correlated more strongly than student self-reported BRIEF-SR scores. Various EFs within the BRIEF and BRIEF-SR tended to show rather similar levels of correlations with standardized scores, although student-reported BRIEF-SR scores varied more. Gender also correlated with Reading and Language Arts scores, but not Math; other demographic factors were not strongly reliably correlated with the exam scores.

Having an IEP correlated significantly and similarly with all three Scaled Scores (*r*s = -.34 – -.36). The negative correlation indicates that those with IEPs tended to receive lower Scantron© Scaled Scores.

In addition, the various diagnoses within the IEPs showed variable levels of correlations with Scaled Scores. Intellectual impairment correlated about -.23 with all three Scaled Scores, and speech or language impairment correlated strongly with Reading (*r* = -.28) and Language Arts (*r* = -.22), and less so with Math (*r* = -.17). Being diagnosed with a learning disability (*r*s = -.12 – -.16) and even with social/emotional impairment (*r*s = -.02 – -.06) correlated significantly with Scaled Scores, but the magnitudes of these correlations were small enough to warrant little attention.

The student's age was *not* significantly correlated with Scaled Scores (r = .08 – .22). However, we found an association between age and Scaled Scores when age was nested within student in the multilevel models reported below.

The Scaled Scores themselves were highly intercorrelated (rs = .63 – .70). The entire matrix of correlations between the variables is provided in Appendix 1, available online.

3.2 Predictive Modeling: Interactive Effects of the Variables

The zero-order correlations with Scaled Scores suggest that EFs (especially as measured by the teachers) as well as students' IEP status, and perhaps demographics provide insights into these students' performance on the Scantron exams. This provides important insight into understanding their success here. However, these correlations only show how each of the variables relate to the Scaled Scores—not how they act together. The matrix of correlations for all of the variables (presented in Appendix 1, available online) indicates that several of the other variables correlate with each other. The individual EFs measured by the BRIEF correlated strongly (mean r = .91), as do those measured in the BRIEF-SR (mean r = .58), and measures between these instruments also correlate mildly (mean r = .14). Special education status correlates as well with BRIEF scores (mean r = .25, r with GEC = .26) and gender (r = -.17), but not well with BRIEF-SR GEC score (r = .09) or sub-scores (mean r = .07)—or with free / reduced lunch status (r = .05), or race/ethnicity (mean $|r|$ = .06).

3.2.1 Multilevel (Hierarchical) Linear Models: Joint Effects of the Variables

The picture, however, is not as simple as these correlations present. We must investigate their joint effects to see a more complete picture. We did this through a series of multilevel models, first creating a base (null) model without EF-related terms. The base models in Tables 2–4 therefore include all model terms *except* those related to executive functioning.

We then added to these base models either terms for BRIEF GEC (i.e., total) scores or for BRIEF-SR GEC scores. The BRIEF columns Tables 2–4 present the changes in model terms after adding the teacher GEC main effect term and the GEC × Age interaction to the base model; the BRIEF-SR columns after adding those terms for student self-reports. Creating a base model allows us to test whether considering EFs more holistically can improve our understanding of how well students perform on the respective Scaled Score. This improvement can be tested by comparing the Bayesian information criterion (BIC) for the base model against the BIC for the model containing the terms for executive functioning. In all cases, adding the executive functioning terms greatly improved the model fits (smallest χ^2_2 = 4916.37; p < .001 for Math Scaled Score predicted by BRIEF-SR terms).

Of course, we then also investigated the size and significance of each of the model terms in relation to the scores in math, reading, and language arts.

3.2.2 Math Scaled Scores

The base model in Table 2 shows that when students' age, gender, free / reduced lunch status, IEP status, and race/ethnicity are all significant when added together to predict Math Scaled Scores. The effect sizes (β-weights) for each term are also rather large. Using the effect size guidelines given by Kraft (2020) (Note 3) suggests that these are all rather "large" effects for education interventions; even Cohen (1988) would consider them to be between "small-" and "medium-" sized effects.

Table 2. Summary of General Linear Regression Model Predicting **Math** Scantron© Scaled Score. Bold-faced terms are significant at α = .05

		Base Model			**Model with BRIEF GEC Terms**			**Model with BRIEF-SR GEC Terms**			**Model with *both* BRIEF *and* BRIEF-SR GEC Terms**		
Model Term		**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***
Age		**0.23** (0.16 – 0.29)	**7.01**	**< 0.001**	**0.32** (0.09 – 0.55)	**2.76**	**0.006**	**0.49** (0.22 – 0.75)	**3.59**	**< 0.001**	**0.36** (0.07 – 0.66)	**2.41**	**0.017**
Gender		**0.17** (0.10 – 0.24)	**4.80**	**< 0.001**	0.08 (-0.14 – 0.29)	0.71	0.478	0.19 (-0.03 – 0.40)	1.70	0.091	0.08 (0.07 – 0.66)	0.65	0.518
IEP Status		**-0.49** (-0.57 – -0.42)	**-12.40**	**< 0.001**	**-0.29** (-0.53 – -0.05)	**-2.37**	**0.018**	**-0.57** (-0.82 – -0.32)	**-4.54**	**< 0.001**	**-0.42** (0.07 – 0.66)	**-3.12**	**0.002**
Free / Reduced Lunch Eligibility		**0.18** (0.09 – 0.27)	**4.03**	**< 0.001**	0.19 (-0.06 – 0.43)	1.50	0.134	**0.30** (0.03 – 0.56)	**2.19**	**0.029**	**0.31** (0.07 – 0.66)	**2.25**	**0.026**
Race/Ethnicity	Asian	**0.74** (0.56 – 0.93)	**7.91**	**< 0.001**	**0.87** (0.34 – 1.39)	**3.27**	**0.001**	**1.17** (0.66 – 1.69)	**4.48**	**< 0.001**	**0.83** (0.07 – 0.66)	**2.93**	**0.004**
	Black	**0.19** (0.10 – 0.29)	**4.02**	**< 0.001**	**0.58** (0.34 – 0.83)	**4.73**	**< 0.001**	**0.56** (0.29 – 0.83)	**4.05**	**< 0.001**	**0.55** (0.07 – 0.66)	**3.80**	**< 0.001**
	Latinx	**0.32** (0.22 – 0.42)	**6.41**	**< 0.001**	**0.38** (0.12 – 0.65)	**2.84**	**0.005**	**0.49** (0.21 – 0.78)	**3.41**	**< 0.001**	**0.41** (0.07 – 0.66)	**2.75**	**0.007**
	White	**0.26** (0.13 – 0.40)	**3.87**	**< 0.001**	**0.47** (0.11 – 0.82)	**2.59**	**0.010**	**0.49** (0.11 – 0.87)	**2.52**	**0.012**	**0.42** (0.07 – 0.66)	**2.12**	**0.035**
BRIEF	GEC				**-0.20** (-0.36 – -0.03)	**-2.35**	**0.019**				**-0.34** (0.07 – 0.66)	**-3.23**	**0.001**
	GEC × Age				0.18 (-0.08 – 0.44)	1.35	0.179				-0.15 (0.07 – 0.66)	-1.45	0.148
BRIEF-SR	GEC							**-0.21** (-0.39 – -0.02)	**-2.18**	**0.031**	-0.04 (0.07 – 0.66)	-0.27	0.786
	GEC × Age							0.04 (-0.23 – 0.32)	0.31	0.754	0.00 (0.07 – 0.66)	0.01	0.988

Table 3. Summary of General Linear Regression Model Predicting **Reading** Scantron© Scaled Score. Bold-faced terms are significant at α = .05.

		Base Model			**Model with BRIEF GEC Terms**			**Model with BRIEF-SR GEC Terms**			**Model with *both* BRIEF *and* BRIEF-SR GEC Terms**		
Model Term		**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***
Age		**0.24** (0.20 – 0.28)	**11.89**	**< 0.001**	**0.36** (0.26 – 0.45)	**7.45**	**< 0.001**	**0.41** (0.32 – 0.51)	**8.58**	**< 0.001**	**0.44** (0.33 – 0.54)	**7.88**	**< 0.001**
Gender		**0.31** (0.25 – 0.37)	**10.12**	**< 0.001**	**0.20** (0.05 – 0.36)	**2.57**	**0.011**	**0.39** (0.24 – 0.54)	**5.05**	**< 0.001**	**0.29** (0.33 – 0.54)	**3.45**	**< 0.001**
IEP Status		**-0.52** (-0.59 – -0.45)	**-15.06**	**< 0.001**	**-0.29** (-0.47 – -0.11)	**-3.18**	**0.002**	**-0.55** (-0.73 – -0.38)	**-6.15**	**< 0.001**	**-0.36** (0.33 – 0.54)	**-3.54**	**< 0.001**
Free / Reduced Lunch Eligibility		**0.22** (0.14 – 0.30)	**5.52**	**< 0.001**	**0.33** (0.14 – 0.51)	**3.46**	**< 0.001**	**0.30** (0.10 – 0.50)	**3.01**	**0.003**	**0.28** (0.33 – 0.54)	**2.68**	**0.008**
Race/Ethnicity	Asian	**0.37** (0.21 – 0.52)	**4.61**	**< 0.001**	**0.44** (0.09 – 0.80)	**2.44**	**0.015**	**0.59** (0.24 – 0.94)	**3.28**	**0.001**	**0.50** (0.33 – 0.54)	**2.53**	**0.012**
	Black	**0.12** (0.04 – 0.20)	**2.92**	**0.004**	**0.34** (0.15 – 0.53)	**3.56**	**< 0.001**	**0.27** (0.08 – 0.46)	**2.82**	**0.005**	**0.38** (0.33 – 0.54)	**3.56**	**< 0.001**
	Latinx	**0.15** (0.07 – 0.23)	**3.85**	**< 0.001**	**0.24** (0.05 – 0.43)	**2.50**	**0.013**	0.17 (-0.02 – 0.36)	1.72	0.087	**0.26** (0.33 – 0.54)	**2.44**	**0.015**
	White	0.11 (-0.01 – 0.22)	1.84	0.066	0.17 (-0.10 – 0.45)	1.25	0.211	0.10 (-0.17 – 0.37)	0.72	0.475	0.10 (0.33 – 0.54)	0.68	0.499
BRIEF	GEC				**-0.29** (-0.39 – -0.19)	**-5.52**	**< 0.001**				**-0.24** (0.33 – 0.54)	**-3.95**	**< 0.001**
	GEC × Age				0.09 (-0.02 – 0.20)	1.59	0.113				**-0.14** (0.33 – 0.54)	**-3.04**	**0.003**
BRIEF-SR	GEC							**-0.17** (-0.26 – -0.09)	**-4.08**	**< 0.001**	**0.16** (0.33 – 0.54)	**2.35**	**0.019**
	GEC × Age							0.03 (-0.07 – 0.12)	0.55	0.581	-0.09 (0.33 – 0.54)	-1.63	0.104

Table 4. Summary of General Linear Regression Model Predicting **Language Arts** Scantron© Scaled Score. Bold-faced terms are significant at α = .05.

		Base Model			**Model with BRIEF GEC Terms**			**Model with BRIEF-SR GEC Terms**			**Model with *both* BRIEF *and* BRIEF-SR GEC Terms**		
Model Term		**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***	**β** (95% CI)	***t***	***p***
Age		**0.09** (0.04 – 0.13)	**3.87**	**< 0.001**	**0.21** (0.08 – 0.34)	**3.12**	**0.002**	**0.24** (0.11 – 0.38)	**3.58**	**< 0.001**	**0.22** (0.07 – 0.36)	**2.90**	**0.004**
Gender		**0.37** (0.31 – 0.43)	**11.52**	**< 0.001**	**0.23** (0.05 – 0.41)	**2.55**	**0.011**	**0.40** (0.21 – 0.58)	**4.28**	**< 0.001**	**0.26** (0.07 – 0.36)	**2.71**	**0.007**
IEP Status		**-0.48** (-0.56 – -0.41)	**-13.12**	**< 0.001**	**-0.24** (-0.46 – -0.02)	**-2.17**	**0.031**	**-0.56** (-0.78 – -0.33)	**-4.93**	**< 0.001**	**-0.26** (0.07 – 0.36)	**-2.12**	**0.035**
Free / Reduced Lunch Eligibility		**0.13** (0.05 – 0.21)	**3.13**	**0.002**	0.21 (-0.01 – 0.42)	1.91	0.057	0.21 (-0.03 – 0.45)	1.76	0.080	0.20 (0.07 – 0.36)	1.66	0.099
Race/Ethnicity	Asian	**0.35** (0.18 – 0.51)	**4.07**	**< 0.001**	0.42 (-0.02 – 0.85)	1.90	0.059	**0.66** (0.24 – 1.08)	**3.06**	**0.002**	0.42 (0.07 – 0.36)	1.78	0.077
	Black	0.06 (-0.02 – 0.15)	1.50	0.135	**0.24** (0.01 – 0.46)	**2.07**	**0.039**	0.15 (-0.08 – 0.39)	1.28	0.202	0.18 (0.07 – 0.36)	1.47	0.143
	Latinx	0.05 (-0.03 – 0.14)	1.26	0.208	0.13 (-0.09 – 0.36)	1.15	0.252	0.08 (-0.16 – 0.31)	0.63	0.527	0.10 (0.07 – 0.36)	0.84	0.404
	White	**0.15** (0.03 – 0.28)	**2.40**	**0.016**	0.08 (-0.26 – 0.42)	0.45	0.650	0.21 (-0.13 – 0.56)	1.23	0.219	0.02 (0.07 – 0.36)	0.13	0.896
BRIEF	GEC				**-0.42** (-0.54 – -0.30)	**-6.78**	**< 0.001**				**-0.39** (0.07 – 0.36)	**-5.30**	**< 0.001**
	GEC × Age				0.04 (-0.12 – 0.19)	0.47	0.640				**-0.14** (0.07 – 0.36)	**-2.48**	**0.014**
BRIEF-SR	GEC							**-0.23** (-0.34 – -0.13)	**-4.43**	**< 0.001**	0.08 (0.07 – 0.36)	0.85	0.395
	GEC × Age							-0.06 (-0.19 – 0.08)	-0.83	0.408	-0.09 (0.07 – 0.36)	-1.23	0.221

Adding in both the main effect for BRIEF GEC scores and GEC × Age interaction terms to the model greatly improved the fit of the model to these data ($BIC_{Base\ Model}$ = 4474.09, $BIC_{BRIEF\ Model}$ = 602.19, χ^2_2 = 3871.89, $p < .001$). Adding those two BRIEF terms also changed the significances (and of course effect sizes) of those terms. Gender and free / reduced lunch status were no longer significant, representing the associations both have with executive functioning often found by others (e.g., Martoni et al., 2015; Noble et al., 2015).

Teacher-rated overall EFs had a small to medium, significant effect on Math Scaled Scores (β = -0.20, p = .019). The interaction with students' ages, however, was not significant (β = 0.17, p = .179). This lack of an interaction suggests that the magnitude of the effect of EFs on math performance does not appreciably change as adolescents age.

The effect of student self-rated overall EFs showed a similar pattern. The main effect of EFs was significant (β = -0.20, p = .031) while the interaction with age was not. The size of main effect for student-rated EFs was thus not appreciably different than for rated by the teachers (95% confidence interval for BRIEF GEC, β = -0.36 – -0.03; for BRIEF-SR GEC, β = -0.39 – -0.02).

All demographic terms remained significant in the model with student self-rated executive functioning. This included gender.

3.2.3 Reading Scaled Scores

All terms (student age, gender, IEP status, and race/ethnicity) displayed small-to-medium effect sizes and significant relationships with Reading Scaled Scores (Table 3). Adding either BRIEF or BRIEF-SR GEC scores to the model led both to significant improvements in the model (χ^2_2s = 4913.56 and 4876.03, respectively; ps < .001) and to significantly main effects for executive functioning (β_{BRIEF} = -0.29, $p < .001$; $\beta_{BRIEF\text{-}SR}$ = -0.17, $p < .001$)—but not significant GEC × Age interactions (β = 0.09, p = .113; $\beta_{BRIEF\text{-}SR}$ = -0.17, p = .581).

The size of the main effects for overall executive function relationships with Reading Scaled Scores (β_{BRIEF} = -0.29, $\beta_{BRIEF\text{-}SR}$ = -0.17) were similar to the size of the effects for Math Scaled Scores (β_{BRIEF} = -0.20, $\beta_{BRIEF\text{-}SR}$ = -0.20). However, the effects for Reading scores were more clearly significant, indicating more reliable relationships to reading ability than to math ability.

3.2.4 Language Arts Scaled Scores

Not all race/ethnicity categories significantly predicted Language Arts Scaled Scores in the base model, but the effects for student age, gender, IEP status, and free / reduced lunch status all were (Table 4).

Adding terms for teacher-rated overall executive functioning significantly improved the model ($BIC_{Base\ Model}$ = 5617.82, BIC_{BRIEF} = 769.58, χ^2_2 = 4848.24, $p < .001$). It led to a large and significant main effect for executive functioning (β = -0.41, $p < .001$) but no significant interaction with age (β = 0.04, $p < .640$)—and to a loss of significance for the free / reduced lunch, change in significance of race/ethnicities and generally smaller effects sizes for the demographic terms. This hints towards more complex relationships between demographics, executive functioning, and language arts abilities.

The pattern was similar, but slightly less pronounced, for student self-rated executive functioning. Adding executive functioning terms also significantly improved the model's fit (BIC_{BRIEF} = 797.91, χ^2_2 = 4819.91, $p < .001$), and generated a significant main effect for executive functioning (β = -0.23, $p < .001$) but no significant interaction with age (β = -0.06, p = .410). It also led to changes in the size—and sometimes significance—of the other model terms; free / reduced lunch status was no longer significant (β = 0.21, p = .080), and the pattern of significance among race/ethnicities changed slightly.

The size of the main effect for student-rated executive functioning (95% CI = -0.33 – -0.13) was significantly smaller than that for the teacher-rated executive functioning (95% CI = -0.54 – -0.30). The association between executive functioning and language arts ability (when accounting for effects of relevant demographics) is weaker when measured by the student themself and when measured by their teachers.

3.2.5 Effects of Individual Executive Functions

The BRIEF and BRIEF-SR sub-scores showed relatively similar patterns in predicting the three Scaled Scores. The models used to test the effects of individual EFs are available in Appendix 2, available in supplemental online materials. Those tables show that BRI and MCI sub-scores were also never appreciably different from each other. This remained true for the individual executive functions that comprise those sub-scores (and the overall GEC score). For example, the coefficients for the BRIEF BRI and MCI main effects predicting Math Scaled Scores were -0.18 (95% CI = -0.33 – -0.01) and -0.19 (95% CI = -0.36 – -0.3), respectively. For BRIEF-SR, they were -0.15 (95% CI = -0.34 – -0.03) and -0.24 (95% CI = -0.44 – -0.05), respectively. The model coefficients and p-values for the various sub-scores and individual executive function scores are provided in the Appendix 2, available online.

For Language Arts, the BRIEF BRI and MCI main effects were both comparably large (βs = -0.42 and -0.41, respectively), and the individual executive functioning scores that comprise each were within each other's confidence intervals. The sub-scores for the BRIEF-SR were also very similar (both βs ≈ -0.22), and their individual executive functioning scores also did not significantly differ.

There was more variability of the associations with Scaled Scores among the individual EFs rated by the students themselves through the BRIEF-SR. EFs were again more predictive of Reading and Language Arts scores than Math; in fact, all individual BRIEF-SR EFs were significant predictors of Reading and Language Arts (except Organization of Materials, $\beta = -0.084$, $t = -1.96$, $p = .051$). However, only BRIEF-SR Task Completion, Plan/Organize, and Monitor significantly predicted Math Scaled Scores.

3.2.6 Comparison of Teacher- Versus Student-Rated Executive Functioning

In addition to investigating whether teacher-rated and student self-rated EFs predicted Scaled Scores individually, we also analyzed if they predicted the scores relative to each other. The farthest-right set of columns in Tables 2 – 4 present the results of these analyses.

Teacher-rated BRIEF GEC scores significantly predicted Math Scaled Scores when both they and student self-rated BRIEF-SR scores were added together in the model. Although the BRIEF-SR GEC main effect had been significant when it was along in the model, it was no longer significant when BRIEF scores were also added (Table 2). We therefore conclude that teacher and student ratings do not contribute sufficiently unique information to the prediction of Language Arts scores; the magnitude of the association with BRIEF scores may also have been stronger than that with BRIEF-SR scores.

The pattern of significance among the other predictors (age, gender, etc.) was similar when only BRIEF or BRIEF-SR scores were added to the Math model. Age, IEP status, free / reduced lunch eligibility, and all race/ethnicities were significant, but gender was not.

For the model predicting Reading Scaled Scores, both the main effects for BRIEF and BRIEF-SR scores were significant; teacher and student ratings thus each contributed significantly unique information about that student's reading performance. The BRIEF × Age interaction was also significant; teacher-rated executive functioning became more strongly associated with Reading scores as the students aged. Once again, the pattern of significances of the other predictors resembled those for the models with only BRIEF or BRIEF-SR terms (Table 3).

Finally, for Language Arts Scale Scores, both the main effect and age interaction terms for the BRIEF remained significant, but neither BRIEF-SR term was. The BRIEF and BRIEF-SR main effects had each been significant when they were alone in the model; as with Math Scaled Scores, teachers and students do not contribute sufficiently unique information about the student's performance (Table 4).

The BRIEF × Age interaction had not been significant when BRIEF terms were alone in the model, but this term became significant when BRIEF-SR terms were added. It is not immediately clear why the changing effect of teacher-rated executive functioning would become clearer when student self-rated executive functioning—and its non-significant changes over time—were also considered.

The pattern of significances of the other terms was also similar to those when the BRIEF or BRIEF-SR scores were alone in the model, although effects of self-identifying with any of the races/ethnicities was not significant. Age, gender, and IEP status were still significant, and free / reduced lunch eligibility still was not.

3.3 Summary of Results

Table 5 summarizes which terms in the various models were reliably significant predictors of the three Scaled Scores, which were sometimes significant, and which were never significant. Teacher-rated BRIEF GEC scores reliably predicted performance in math, reading, and English language arts across the middle and high school grades. Age was also reliably associated with performance on these standardized exams: Older students tended to do better on these exams even though the exam scores are standardized by grade level. BRIEF scores themselves also tended to become more strongly associated with Reading and Language Arts (but not Math) scores as the students aged.

Table 5. Summary of Significant Predictors in All General Linear Regression Models

		Math	Reading	Language Arts
Age		◆	◆	◆
Gender			◆	◆
IEP Status		◆	◆	◆
Free / Reduced Lunch Eligibility		∘	◆	◆
Race/Ethnicity		◆	∘	∘
BRIEF	Main Effect	◆	◆	◆
	Changes over Time		∘	∘
BRIEF-SR	Main Effect	∘	◆	∘
	Changes over Time			

◆ = Reliably significant.

∘ = Sometimes significant.

= Not significant (i.e., blank cells).

Student self-rated BRIEF-SR scores were also significantly associated with performance in all three content areas—but not always when teacher-rated BRIEF scores were also added. BRIEF-SR scores only remained significant—i.e., only continued to provide unique information—when predicting Reading scores. The predictiveness of the BRIEF-SR scores did not change over time (their interaction with age was never significant). What information students' self-rating provided did not change.

Students' gender identification and eligibility for free / reduced lunches reliably predicted performance on the Reading and Language Arts exams, but less so performance on the Math exam. Gender was never a significant predictor of Math scores when any EF terms were considered. free / reduced lunch eligibility was significant whenever BRIEF-SR terms were included. Students' race/ethnicity was sometimes associated with their performance on the three exams. Consistent patterns here are difficult to detect and anyway would be even harder to interpret. Students with IEPs tended to perform more poorly on all three exams; this association remained after either teacher- or student-self-rated executive functioning scores were added.

4. Discussion

4.1 Relationship Between EFs and Academic Performance on Standardized Tests

This longitudinal study supports the relationship between EFs and academic performance on standardized tests among adolescents in schools that integrates students with and without disabilities. Our findings indicate that the overall levels of adolescent student EFs measured by both teachers (via the BRIEF) and by the students themselves (via the BRIEF-SR) significantly predicted those students' performance on standardized reading, math, and English language arts exams. These results support the importance of EFs for academic performance and build upon prior research that used GPAs to predict EFs that were also assessed by teachers and by the students themselves (Samuels et al., 2016, 2019). Such predictions have been better established for younger children (Waber et al., 2006; Wagner et al., 2019) but are not as well established for adolescents. Only Best et al. (2011) investigated the relationships between EFs and academic achievement among both children and adolescents and found that EFs were moderately correlated with success in both math and reading achievement therefore, consistent with our findings. Below are some of our specific findings.

4.1.1 Comparison of BRIEF with BRIEF-SR

One of the goals of our study was to compare the contributions of the BRIEF with those of the BRIEF-SR for their uses as experimental tools. We found that BRIEF scores (reported by teachers) out-performed BRIEF-SR scores (self-reported by students) as predictors of academic outcomes. The effect sizes (β-weights) for BRIEF terms showed some tendency to be larger than those for the BRIEF-SR; those BRIEF terms were also more reliably significant—even when both BRIEF and BRIEF-SR terms were both added to the models. BRIEF GEC ratings successfully predicted student standardized reading, math and English language arts scores across the middle and high school grades; BRIEF-SR scores were predictive but less so. However, when teacher-rated BRIEF scores were also added, BRIEF-SR scores only remained significant—i.e., only continued to provide unique information—when predicting Reading scores. These results suggest that BRIEF and BRIEF-SR scores cannot be used interchangeably to make significant predictions and are somewhat different from previous findings (Samuels et al., 2016, 2019) in which the scores of the BRIEF or BRIEF-SR could be used alone to make significant predictions about how students perform in middle and high school courses. The current study suggests that using both may be unnecessary, although—of course—more research must be conducted to better support that.

4.2 The Behavioral Regulation Index (BRI) and the Metacognitive Index (MCI)

It is of considerable significance that both EF overall and its two components—the Metacognitive Index (MCI) and the Behavioral Regulation Index (BRI)—correlated with academic performance as measured by standardized testing. It supports prior factor analyses, enhances the overall findings of this study, and can influence academic practice. That is, it suggests that classroom work includes a good deal of emphasis on behavioral regulation, i.e., the "ability to shift cognitive set and modulate emotions and behavior via appropriate inhibitory control" while allowing "metacognitive processes to successfully guide active, systematic problem solving (and supports) appropriate self-regulation" (Gioia et al., 2000, p. 20). The current study sample were students in a school implementing an established, research-based "wellness" curriculum designed for adolescence and focused on social-emotional skills training and problem solving.

4.3 Students with Disabilities

Adolescents' EFs—especially when measured by their current teachers—predicts performance on standardized academic assessments throughout the middle and high school grades. This effect existed among a rather diverse sample of students both with and without diagnosed disabilities (the latter, scoring lower on measures of EF and Academic Performance, as expected).

4.4 Demographic Factors

4.4.1 Age

Age was reliably associated with performance on these standardized exams: older students tended to do better even though the exam scores were standardized by grade level. BRIEF scores themselves also tended to become more strongly associated with Reading and Language Arts (but not Math) scores as the students aged. EFs typically improve throughout childhood, aligning well with the maturation of the frontal lobes (Anderson, 2002) and with cortical areas that continue to develop throughout adolescence (Faridi et al., 2015; Yakovlev & Lecours, 1967). Theodoraki et al. (2020) investigated whether EFs continue to develop during the later stages of adolescence (aged 15 – 18 years) by examining three behavioral components of EFs: inhibition, shifting, and working memory. After adjusting for covariates, age was found to be a significant predictor of students' performance on the inhibition but not the shifting or working memory tasks, suggesting different developmental trajectories for the three EF components. However, in a meta-analysis of existing studies, Cortés Pascual et al. (2018) did not find age to be a significant moderator of the effects of EFs on adolescents' academic performance.

The effect (and non-effect) of age here deserves further consideration—because this is among the few studies to track adolescents so consistently over so many years. Similar to what Cortés Pascual et al. (2018) did, the effect of age we are measuring here is whether it moderates the effect of EF on academic performance. Cortés Pascual et al.'s work benefits from testing the effect of age across several (19) studies but could not consistently track the development of EFs within individuals; their meta-analysis can thus be seen as providing a strong, generalizable, but cross-sectional investigation of the moderating effect of age, with most of the studies included therein sampling adolescents near the lower range of those sampled here (Note 4). The current study, however, represents a closer, more controlled investigation of a smaller sample of adolescents who tend to be older and less diverse (Note 5) than those studied by Cortés Pascual et al.

We benefited from having obtained both teacher- and self-reports on EFs every year for the same adolescents. A different teacher rates a given adolescent every year. Although we have not tested this proposition, it may well be then that each year's teacher is using a different standard to evaluate EFs, with, e.g., a six-grade teacher basing their ratings on sixth graders and a twelfth-grade teachers basing their ratings on twelfth graders. If so, this would suggest that it is not the improvement in EFs per se that is leading to an effect of age (on reading and language arts), but that EFs are increasingly important as an adolescent age—whatever the EF's level of maturation at those ages.

We did not find a significant effect of age on the effects of self-reported EFs on standardized exam performance. We generally found fewer and weaker effects of EFs measured through the students themselves; this is at least partially due to the greater variability within the students' self-reported scores. However, it may also be that adolescents view their own behaviors quite differently than their teachers (BRIEF GEC and BRIEF-SR GEC scores correlated weakly here, $r = .19$)—and that the criteria used by teachers measures EFs in ways more relevant to academic performance. This conjecture remains untested but gains some support from the fact that the effect of age on teacher-reported EFs was only significant here when both teacher- and student-reported scores were both added to the model; adding them both together isolates the effects of each from each other (Note 6), making any differences in what those scores measure more acute.

4.4.2 Gender

Identifying as female reliably predicted better performance on the Reading and Language Arts exams, but less so for performance on the Math exam. Gender was never a significant predictor of Math scores when any executive functioning terms were considered suggesting that the effect of gender is sufficiently mediated by girls' stronger development of EFs here. Our findings do not confirm those reported by Grissom and Reyes (2019) who reviewed several studies and concluded that while individual factors may show a tendency towards gender differences in EFs (e.g., increased impulsive action in males, reduced reaction time in males, avoidance of frequent punishment in females, improved working memory in females), those differences are not overwhelming. Within-gender variability often far exceeds between-gender variability, and in few cases could one look at a given person's data and be able to classify them by their responses as male or female.

4.4.3 Race/Ethnicity

Students' race/ethnicity was sometimes associated with their performance on the standardized exams. Consistent patterns here are difficult to detect but our findings seem to support results by Reid and Ready (2022) who report on the heterogeneity of EFs' development. More specifically, they stated that low-SES and Hispanic dual-language-learning children with immigrant parents entered kindergarten with the lowest average EFs skills but then made remarkable EFs gains. However, low-SES, non-dual-language-learning Black and Hispanic children had similarly low initial EFs skills, but did not exhibit the same pattern of catch-up, in part due to their reduced likelihood of enjoying positive relationships with their teachers.

4.5 Conclusions and Implications: The Importance of EFs in Academic Performance

Adolescents' EFs—especially when measured by their current teachers—well predict performance on standardized academic assessments throughout the middle and high school grades. The current study therefore further supports the body of research underlining the importance of EFs in academics; it also advances that understanding, as follows.

First, we found this effect among a rather diverse sample of students with and without diagnosed disabilities. Although all of these students attended one school, they represent the general population better than most studies on adolescent's EFs where those with particular disabilities (e.g., ADHD) are the only ones studied. These findings should therefore support the roles of EFs among equally general populations of adolescents.

Second, EFs rated through teachers were more reliably predictive of standardized exam performances than EFs rated by the students themselves. Prior studies (e.g., Samuels et al., 2019) found that adolescents' EFs predicted their GPAs; although such studies lend support to the EF- academics association, the teachers who rated the EFs were among those assigning grades. The current study removes that potential bias, demonstrating the importance of EFs on academic outcomes created and assessed well outside of the school. It is worth noting that we do not believe these results mean that students' self-assessments are either wrong or uninformative. It may be that students self-assess their EFs accurately, but if they do, it's in ways that are less related to their academic performance.

Students' self-assessments also tended to vary more than those made by the teachers. A different teacher assessed the students each year, so the variability in teacher assessments here arises from some different sources than the variability in a student assessing themselves over different years. If one's goal is simply to measure EFs in ways more closely associated with academic performance, then, asking teachers—and perhaps multiple teachers—is advisable.

Third, we found that the predictiveness of EFs tended to improve across these grades—when measured by teachers, but not those measured by students themselves. EFs were somewhat more strongly associated with performance on Reading and Language Arts exams than on Math exams. Although we cannot say why this is so, it implies that EFs may even be more important among older adolescents, perhaps as they become increasingly entrusted (even expected) to be responsible for their own academic performance.

Fourth, as expected, students with IEPs tended to display lower EFs scores and to obtain lower scores on these standardized assessments than those without IEPs.

Fifth, students who identified as female tended to show stronger EFs. Nonetheless, gender per se continued to significantly predict exam performance even after EFs terms were added to the models (thus partialing out the effects of EFs on the gender-exam association). Therefore, those who identify as female appear to benefit from generally stronger EFs and from other gender-related factors not directly measured by instruments like the BRIEF and BRIEF-SR.

Sixth, we did not find that any one individual EF was markedly more predictive than any other individual EF. This was especially true among those rated by the teachers. This was somewhat less true among those rated by the students themselves where Task Completion, Plan/Organize, and Monitor significantly were the only individual EFs that significantly predicted Math scores (all individual student-rated EFs significantly predicted Reading and Language Arts scores). It may be that we could not detect real differences between the EFs, but even where they exist, these individual differences appear to matter less than a consideration of EFs more broadly. We therefore recommend at least measuring a wider range of EFs and expecting that general interventions may be more effective than ones tailored to specific EFs.

Finally, an initial analysis of those with a diagnosed disability, indicated that a one-year change in age was associated with essentially the same level of academic performance as the prior age, when students were placed in the inclusive classrooms of this study. In contrast, Tormanen and Roebers (2018) found that after two years students with disabilities placed in self-contained classrooms displayed *significantly lower scores* in academic achievement and EFs. It is clear that more research is needed on this subject.

Together, these findings suggest that EFs offer unique and important insights into adolescents' performance on standardized exams. Researchers can further investigate: 1) The relationship between EF and performance for those with diagnosed disabilities; and 2) the mechanisms through which EFs affect performance—especially on high-stakes exams—and whether interventions can help students further strengthen their EFs. Teachers and administrators can consider a holistic development of their students while knowing that addressing broad competencies like EFs can help students with important, specific tasks like performance on standardized exams.

4.6 Limitations

The primary limitations of this study are that the sample, though large, contains students who all attended the same middle and high school. In addition, the study design was single group observation in an important setting (student with and without disabilities) and the design did not manipulate EFs to investigate more systematic effects on academic performance. More controlled trials are needed in a systematic series of studies.

Following students at only two schools affects the generalizability of the results (their internal validity should not be affected by this). The schools that these students attend are entirely inclusive, at least in the sense that students with and without disabilities share all classes and extracurricular activities. These rare inclusive environments may affect the roles of EFs in students' performance, although we cannot at this time say how.

The importance of EFs may differ among other student bodies. EFs may also develop differently among other populations, so longitudinal studies conducted elsewhere may find stronger, weaker, or simply different effects of time and age. The mixed results found elsewhere (e.g., Cortés Pascual et al., 2018) also suggest that this is a particularly worthwhile area for further study. These limitations notwithstanding, the results of this study extend the contribution of executive functioning to standardized assessments, adolescents with and without diagnosed disabilities and for both its cognitive and behavioral components.

References

Ahmed, S. F., Tang, S., Nicholas, E., Waters, N. E., & Davis-Kean, P. (2018). Executive function and academic achievement: Longitudinal relations from early childhood to adolescence. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000296.

Alathea, L. (2015). *captioner: Numbers Figures and Creates Simple Captions*. R package version 2.2.3, https://CRAN.R-project.org/package=captioner

Albert, W. D., Hanson, J. L., Skinner, A. T., Dodge, K. A., Steinberg, L., Deater-Deckard, K., … & Lansford, J. E. (2020). Individual differences in executive function partially explain the socioeconomic gradient in middle-school academic achievement. *Developmental Science*, *23*(5), e12937-n/a. https://doi.org/10.1111/desc.12937

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher, 49*(3), 198-211. https://doi.org/10.3102/0013189X20902110.

Al-Yagon, M., & Borenstein, T. (2022). Adolescents' executive functions: Links to inattention, hyperactivity-impulsivity, trait mindfulness, and attachment relationships with fathers and mothers. *Research in Developmental Disabilities*, *124*, 104212-104212. https://doi.org/10.1016/j.ridd.2022.104212

Anderson, P. (2002). Assessment and development of executive function during childhood. *Child Neuropsychology*, *8*, 71-82.

Andersson, U. (2008). Working memory as a predictor of written arithmetical skills in children: The importance of central executive functions. *British Journal of Educational Psychology*, *78*(2), 181-204. https://doi.org/10.1348/000709907X209854.

Aronen, E. T., Vuontela, V., Steenari, M.R., Salmi, J., & Carlson, S. (2005). Working memory, psychiatric symptoms, and academic performance at school. *Neurobiology of Learning and Memory, 83*(1), 33-42. https://doi.org/10.1016/j.nlm.2004.06.010.

Artuso, C., Carretti, B., & Palladino, P. (2019). Short-term training on working memory updating and metacognition in primary school: The effect on reading comprehension. *School Psychology International*, *40*(6), 641-657. https://doi.org/10.1177/0143034319881671

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641-60. https://doi.org/10.1111/j.1467-8624.2010.01499.x.

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180-200.

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.

Bierman, K. L., Torres Marcela, M., Domitrovich Celene, E., Welsh, J. A., & Gest Scott, D. (2009). Behavioral and cognitive readiness for school: Cross-domain associations for children attending head start. *Social Development*, *18*(2), 305-323. https://doi.org/10.1080/87565641.2011.591857.

Blair, C., & Cybele, R. C. (2012). Individual development and evolution: Experiential canalization of self-regulation. *Developmental Psychology*, *48*(3), 647-657. https://doi.org/10.1037/a0026472.

Boschloo, A., Krabbendam, L., Aben, A., de Groot, R., & Jolles, J. (2014). Sorting test, tower test, and BRIEF-SR do not predict school performance of healthy adolescents in preuniversity education. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00287.

Brijbooms, E., Groen, M. A., & Verhoeven, L. (2017). How Executive Functions predict development in syntactic complexity of narrative writing in the upper elementary grades. *Reading and Writing*, *30*, 209-231.

Buckner, J. C., Mezzacappa, E., & Beardslee, W. R. (2009). Self-regulation and its relations to adaptive functioning in low income youths. *American Journal of Orthopsychiatry*, *79*(1), 19-30. https://doi.org/10.1037/a0014796

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1), 36-41. https://doi.org/10.1111/cdep.12059.

Burton, N. W., & Ramist, L. (2001). Predicting success in college: SAT® studies of classes graduating since1980. Research Report No. 2001-2. *College Entrance Examination Board*. College Entrance Examination Board.

Enwefa, C. (2015). Grading practice as valid measures of academic achievement of secondary schools students for national development. *Journal of Education and Practice*, *6*(26), 24.

Clark, C. A. C., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology*, *46*(5), 1176-1191. https://doi.org/10.1037/a0019672.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology*, *31*(1), 103-128.

Cortés Pascual, A., Muñoz, N. M., & Robres, A. Q. (2019). The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Frontiers in Psychology*, *10*(1582). https://doi.org/10.3389/fpsyg.2019.01582.

Danielsson, H., Henry, L., Ronnberg, J., & Nilsson, L. G. (2010). Executive functions in individuals with intellectual disability. *Research in Developmental Disabilities: A Multidisciplinary Journal*, *31*(6), 1299-1304.

Deer, L. K., Hastings, P. D., & Hostinar, C. E. (2020). The role of childhood executive function in explaining income disparities in long-term academic achievement. *Child Development*, *91*(5), e1046-e1063. https://doi.org/10.1111/cdev.13383.

DePasquale, C. E., Tyrell, F. A., Kalstabakken, A. W., Labella, M. H., Thibodeau, E. L., Masten, A. S., & Barnes, A. J. (2021). Lifetime stressors, hair cortisol, and executive function: Age-related associations in childhood. *Developmental Psychobiology*, *63*(5), 1043-1052. https://doi.org/10.1002/dev.22076

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-68. https://doi.org/10.1146/annurev-psych-113011-143750.

Dubuc, M. M., Aubertin-Leheudre, M., & Karelis, A. D. (2020). Relationship between interference control and working memory with academic performance in high school students: The Adolescent Student Academic Performance Longitudinal Study (ASAP). *Journal of Adolescence (London, England.) 80*(1), 204-213. https://doi.org/10.1016/j.adolescence.2020.03.001.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83-87. https://doi.org/10.1111/1467-8721.01235.

Dutra, N. B., Chen, L., Anum, A., Burger, O., Davis, H. E., Dzokoto, V. A., … & Legare, C. H. (2022). Examining relations between performance on non-verbal executive function and verbal self-regulation tasks in demographically-diverse populations. *Developmental Science*, *25*(5), e13228-n/a. https://doi.org/10.1111/desc.13228

Faridi, N., Karama, S., Burgaleta, M., White, M. T., Evans, A. C., Fonov, V., … & Waber, D. P. (2015). Neuroanatomical correlates of behavioral rating versus performance measures of working memory in typically developing children and adolescents. *Neuropsychology*, *29*(1), 82-91.

Foley, J., Garcia, J., Shaw, L., & Golden, C. (2009). IQ predicts neuropsychological performance in children. *International Journal of Neuroscience*, *119*(10), 1830-1847. https://proxy.library.csi.cuny.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=44192628&site=ehost-live.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*(6), 373-378. https://doi.org/10.1111/j.0956-7976.2004.00687.x.

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, *50*(6), 1698-1709. Epub 2014 Apr 21. PMID: 24749550. https://doi.org/10.1037/a0036633

Galla, B. M., Shulman, E. P, Plummer, B. D, Gardner, M., Hutt, S. J., Goyer, J. P., … & Duckworth, A. L. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal*, *56*(6), 2077-2115. https://doi.org/10.3102/0002831219843292.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function: Professional Manual.* Lutz, Florida: Psychological Assessment Resources.

Gioia, G. A., Isquith, P. K., & Kenworthy, L. (2015). *Behavior Rating Inventory of Executive Function®*. 2nd ed. Lutz, Florida: PAR, Inc.

Grissom, N. M., & Reyes, T. M. (2019). Let's call the whole thing off: Evaluating gender and sex differences in executive function. *Neuropsychopharmacology*, *44*, 86-96.

Guy, S. C., Isquith, P. K., & Gioia, S. C. (2004). *Behavior Rating Inventory of Executive Function—Self-Report Version: Professional Manual.* Lutz, Florida: Psychological Assessment Resources.

Hall, D., Jarrold, C., Towse, J. N., & Zarandi, A. L. (2015). The developmental influence of primary memory capacity on working memory and academic achievement. *Developmental Psychology*, *51*(8), 1131-1147. https://doi.org/10.1037/a0039464.

Hannon, B. (2014). Predicting college success: The relative contributions of five social/personality factors, five cognitive/learning factors and SAT scores. *Journal of Education and Training Studies*, *2*(4), 46.

Hannon, B., & McNaughton-Cassill, M. (2011). SAT performance: Understanding the contributions of cognitive/learning and social/personality factors. *Applied Cognitive Psychology*, *25*(4), 528-535. https://doi.org/10.1002/acp.1725.

Harrell, F. (2022). *Hmisc: Harrell Miscellaneous*. R package version 4.7-0, https://CRAN.R-project.org/package=Hmisc.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017-2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Huizinga, M., & van der Molen, M. W. (2007). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Developmental Neuropsychology*, *31*(2), 193-215. https://doi.org/10.1080/87565640701190817.

Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child*, *2*(2), 125-32. https://doi.org/10.1080/21622965.2013.748389.

Karr, J. E., Rodriguez, J. E., Goh, P. K., Martel, M. M., & Rast, P. (2022). The unity and diversity of executive functions: A network approach to life span development. *Developmental Psychology*, *58*(4), 751-767. https://doi.org/10.1037/dev0001313

Keller, A., Sydnor, V. J., Pines, A., Fair, D. A., Bassett, D. S., & Satterthwaite, T. D. (2023). Hierarchical functional system development supports executive function. *Trends in Cognitive Sciences*, *27*(2), 160-174. https://doi.org/10.1016/j.tics.2022.11.005

Kobrin, J., & Patterson, B. F. (2010). *Exploring the Variability in the Validity of SAT Scores and High School GPA for Predicting First-Year College Grades at Different Colleges and Universities*. College Board.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT® for Predicting First-Year College Grade Point Average. Research Report No. 2008-5*. College Board.

Kotsopoulos, D., & Lee, J. (2012). A naturalistic study of executive function and mathematical problem-solving. *The Journal of Mathematical Behavior*, *31*(2), 196-208. https://doi.org/10.1016/j.jmathb.2011.12.005.

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241-53. https://doi.org/10.3102/0013189X20912798.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13.

Langberg, J. M., Dvorsky, M. R., & Evans, S. W. (2013). What specific facets of executive function are associated with academic functioning in youth with attention-deficit/hyperactivity disorder? *Journal of Abnormal Child Psychology*, *41*(7), 1145-59. https://doi.org/10.1007/s10802-013-9750-z.

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, *84*(6), 1933-53. https://doi.org/10.1111/cdev.12096.

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, *101*(2), 373-87. https://doi.org/10.1037/a0013843.

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59-80. doi:10.1348/026151003321164627

Locascio, G., Mahone, E. M., Eason, S. H., & Cutting, L. E. (2010). Executive dysfunction among Children with reading comprehension deficits. *Journal of Learning Disabilities*, *43*(5), 441-454. https://doi.org/10.1177/0022219409355476.

López, M. (2013). Rendimiento académico: Su relación con la memoria de trabajo. revista electronic*a. Actualidades Investigativas En Educación*, *13*(3), 1-19.

Martoni, R. M., Salgari, G., Galimberti, E., Cavallini, M. C., & O'Neill, J. (2015). Effects of gender and executive function on visuospatial working memory in adult obsessive–compulsive disorder. *European Archives of Psychiatry and Clinical Neuroscience*, *265*(8), 707-718. https://doi.org/10.1007/s00406-015-0604-2.

McNeish, D. M., Radunzel, J., & Sanchez, E. (2015). A multidimensional perspective of college readiness: Relating student and school characteristics to performance on the ACT®. *ACT Research Report Series 2015 (6)*. ACT, Inc.

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

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), 158–73. https://doi.org/10.1016/j.jecp.2011.01.008.

New York State Education Department. (2015–2016). *Student Assessments and Associated Growth Models for Teacher and Principal Evaluation: Publicly Available Services Summary*. Scantron. New York State Education Department.

Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., ... & Sowell, E. R. (2015). Family income, parental education and brain structure in children and adolescents. *Nature neuroscience*, *18*(5), 773-778. https://doi.org/10.1038/nn.3983.

Nutley, S. B., & Soderqvist, S. (2017). How is working memory training likely to influence academic performance? Current evidence and methodological considerations. *Frontiers in Psychology*, *8*, 69-69. https://doi.org/10.3389/fpsyg.2017.00069.

Ober, T. M., Brooks, P. J., Homer, B. D., & Rindskopf, D. (2020). Executive functions and decoding in children and adolescents: a meta-analytic investigation. *Educational Psychology Review*, *32*(3), 735-763. https://doi.org/10.1007/s10648-020-09526-0

Ostrosky-Solís, F., Gómez-Pérez, M. E., Matute, E., Rosselli, M., Ardila, A., & Pineda, D. (2007). Neuropsi attention and memory: A neuropsychological test battery in Spanish with norms by age and educational level. *Applied Neuropsychology*, *14*(3), 156-70. https://doi.org/10.1080/09084280701508655.

Reddy, L. A., Hale, L. E., & Brodzinsky, L. K. (2011). Discriminant validity of the Behavior Rating Inventory of Executive Function Parent Form for children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly*, *26*(1), 45-55. https://doi.org/10.1037/a0022585.

Reid, J. L., & Ready, D. D. (2022). Heterogeneity in the development of executive function: Evidence from nationally representative data. *Early Education and Development, (33)*2, 243-267. https://doi.org/10.1080/10409289.2020.1866949

Ren, X., Schweizer, K., Wang, Y., & Xu, F. (2015). The prediction of students' academic performance with fluid intelligence in giving special consideration to the contribution of learning. *Advances in Cognitive Psychology*, *11*(3), 97-105. https://doi.org/10.5709/acp-0175-z.

Revelle, W. (2022) *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version 2.2.5.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353-387. https://doi.org/10.1037/a0026838.

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2014). *Handbook of Executive Functioning*. New York, NY: Springer.

Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2012). The role of socioeconomic status in SAT-grade relationships and in college admissions decisions. *Psychological Science*, *23*(9), 1000-1007. https://doi.org/10.1177/0956797612438732.

Samuels, W. E., Tournaki, N., Blackman, S., & Zilinski, C. (2016). Executive functioning predicts academic achievement in middle school: A 4-year longitudinal study. *The Journal of Educational Research*, *109*(5), 478-490. https://doi.org/10.1080/00220671.2014.979913.

Samuels, W. E., Tournaki, N., Sacks, S., Blackman, S., Sacks, J., Byalin, K., & Zilinski, C. (2019). Predicting GPAs with executive functioning assessed by teachers and by the students themselves. *The European Educational Researcher*, *2*(3), 173-194. https://doi.org/10.31757/euer.232.

Scantron. (2015). *Performance Series Scores*. Scantron Corporation.

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*(3), 232-246. https://doi.org/10.1080/09297040802220029.

Snowman, J., & Biehler, R. F. (2003). *Psychology Applied to Teaching* (10th ed.). Boston: Houghton Mifflin.

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

Theodoraki, T. E., McGeown, S. P., Rhodes, S. M., & MacPherson, S. E. (2020). Developmental changes in executive function during adolescence: A study of inhibition, shifting, and working memory. *British Journal of Developmental Psychology, 38,* 74-89.

Toplak, M. E., Connors, L., Shuster, J., Knezevic, B., & Parks, S. (2008). Review of cognitive, cognitive-behavioral, and neural-based interventions for attention-deficit/hyperactivity disorder (ADHD). *Clinical Psychology Review*, *28*(5), 801-823.

Tormanen, M. R. K., & Roebers, C. M. (2018). Developmental outcomes of children in classes for special education needs: Results from a longitudinal study. *Journal of Research in Special Educational Needs, 18*(2), 83-93.

Torske, T., Nærland, T., Øie, M. E., Stenberg, N., & Andreassen, O. A. (2017). Metacognitive aspects of executive function are highly associated with social functioning on parent-rated measures in children with Autism Spectrum Disorder. *Frontiers in Behavioral Neuroscience*, *11*, 258-270.

Tsubomi, H., & Watanabe, K. (2017). Development of visual working memory and distractor resistance in relation to academic performance. *Journal of Experimental Child Psychology*, *154*, 98-112. https://doi.org/10.1016/j.jecp.2016.10.005.

van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *The British Journal of Educational Psychology*, *82*(Pt 1), 100-119. Retrieved from https://proxy.library.csi.cuny.edu/login?url= http://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=22429060&site=ehost-live

Vogan, V. M., Leung, R. C., Safar, K., Martinussen, R., Smith, M. L., & Taylor, M. J. (2018). Longitudinal examination of everyday executive functioning in children with ASD: Relations with social, emotional, and behavioral functioning over time. *Frontiers in Psychology*, *9*, 1774.

Vuontela, V., Carlson, S., Troberg, A.-N., Fontell, T., Simola, P., Saarinen, S., & Aronen, E. T. (2013). Working memory, attention, inhibition, and their relation to adaptive functioning and behavioral/emotional symptoms in school-aged children. *Child Psychiatry and Human Development*, *44*(1), 105-122. https://doi.org/10.1007/s10578-012-0313-2

Waber, D. P., Gerber, E. B., Turcios, V. T., Wagner, E. R., & Forbes, P. W. (2006). Executive functions and performance on high-stakes testing in children from urban schools. *Developmental Neuropsychology*, *29*(3), 459-477. https://doi.org/10.1037/t15174-000

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.9. https://CRAN.R-project.org/package=dplyr

Wichstrøm, L. (1995). Harter's Self-Perception Profile for Adolescents: Reliability, validity, and evaluation of the question format. *Journal of Personality Assessment*, *65*(1), 100. https://proxy.library.csi.cuny.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=6381096&site=ehost-live

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, *39*(1), 1-37. https://doi.org/10.1111/j.1745-3984.2002.tb01133.x.

Xie, Y. (2022). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.39.

York, T. T., Gibson III, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research & Evaluation*, *20*(5), 1-20.

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354-360. https://doi.org/10.1111/j.1750-8606.2012.00246.x.

Zwick, R., & Green, G. J. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement*, *44*(1), 23-45. https://doi.org/10.1111/j.1745-3984.2007.00025.x

Notes

Note 1. Note that, following convention, all BRIEF and BRIEF-SR are scored such that lower scores denote greater executive functioning, so these negative correlations indicate that students whose EFs were rated as stronger tended to have higher Scaled Scores

Note 2. The correlations of these sub-scores and individual scores are presented in the online supplement (Appendix 1); the highlights of those investigations include the following.

Note 3. Kraft (2020) found that effect sizes for education interventions tend to be somewhat lower than what Cohen (1988) proposed for "small," "medium," and "large" effects—even for large, randomized control trials and in light of potential publication bias. Kraft's analyses led him to suggest that, among other things, a Cohen's *d*s of .05 should be considered "small," between .05 – .2 "medium," and above .2 "large." The effect sizes generated by general linear models—including the multilevel models used here—are β-weights (equivalent to Cohen's *f*s; these are half the size of Cohen's *d*s, so here .025 may be tentatively considered "small," .025 – .1 "medium," and > .1 "large."

Note 4. Of the 19 studies included in Cortés Pascual et al.'s (2016) analyses of age effects, one had maximum participant ages of 15, and one had maximum participant ages of 17; all others had maximum ages of 12 or less, below the range of ages for the adolescents studied here.

Note 5. Although ethnically/racially and economically diverse, our participants all attend the same schools. The studies included in Cortés Pascual et al.'s (2016) analyses include adolescents not only from the US (k = 5) and Canada (k = 1) but also from Western Europe (k = 11), Brazil (k = 1), and Japan (k = 1).

Note 6. As well as forcing any covariance between the BRIEF and BRIEF-SR scores to be placed within one or the other of the variables—usually more into the stronger predictor (here, BRIEF scores).

Appendices

Appendix 1. Correlations Between All Variables

		Scantron Scaled Scores		
		Language Arts	**Math**	**Reading**
Scantron Scaled Scores	**Language Arts**		.717	.752
	Math	.717		.695
	Reading	.752	.695	
BRIEF Scores	**Emotional Control**	-.479	-.343	-.386
	Inhibit	-.440	-.327	-.370
	Initiate	-.447	-.323	-.331
	Monitor	-.462	-.380	-.404
	Organization of Materials	-.447	-.333	-.351
	Plan/Organize	-.461	-.349	-.397
	Shift	-.450	-.332	-.374
	Working Memory	-.477	-.390	-.409
	Behavioral Regulation Index	-.480	-.353	-.395
	Metacognitive Index	-.485	-.377	-.403
	Global Executive Composite	-.487	-.370	-.403
BRIEF-SR Scores	**Emotional Control**	-.063	-.006	-.065
	Inhibit	-.186	-.056	-.087
	Monitor	-.178	-.110	-.129
	Organization of Materials	-.143	-.111	-.037
	Plan/Organize	-.163	-.126	-.070
	Shift	-.088	-.056	-.108
	Task Completion	-.234	-.227	-.176
	Working Memory	-.086	-.067	-.028
	Behavioral Regulation Index	-.149	-.061	-.109
	Metacognitive Index	-.176	-.151	-.090
	Global Executive Composite	-.171	-.114	-.104
	Special Education Status	-.351	-.349	-.375
	Age	.090	.221	.210
	Gender	.122	-.008	.070
Ethnicity/Race	**American Indian**	.053	.062	.044
	Asian	.083	.114	.071
	African	-.080	-.081	-.088
	Latin	-.128	-.095	-.098
	Multiracial	.025	.024	-.002
	European	.030	-.005	-.011

		BRIEF Scores										
		Emotional Control	Inhibit	Initiate	Monitor	Organization of Materials	Plan/Organize	Shift	Working Memory	Behavioral Regulation Index	Metacognitive Index	Global Executive Composite
Scantron Scaled Scores	Language Arts	-.479	-.440	-.447	-.462	-.447	-.461	-.450	-.477	-.480	-.485	-.487
	Math	-.343	-.327	-.323	-.380	-.333	-.349	-.332	-.390	-.353	-.377	-.370
	Reading	-.386	-.370	-.331	-.404	-.351	-.397	-.374	-.409	-.395	-.403	-.403
BRIEF Scores	Emotional Control		.906	.924	.919	.920	.918	.908	.932	.973	.954	.966
	Inhibit	.906		.910	.919	.886	.934	.875	.909	.965	.946	.958
	Initiate	.924	.910		.906	.909	.914	.889	.913	.942	.959	.957
	Monitor	.919	.919	.906		.893	.939	.915	.938	.951	.970	.967
	Organization of Materials	.920	.886	.909	.893		.890	.873	.905	.926	.944	.941
	Plan/Organize	.918	.934	.914	.939	.890		.917	.936	.957	.976	.973
	Shift	.908	.875	.889	.915	.873	.917		.919	.956	.937	.949
	Working Memory	.932	.909	.913	.938	.905	.936	.919		.953	.974	.971
	Behavioral Regulation Index	.973	.965	.942	.951	.926	.957	.956	.953		.980	.993
	Metacognitive Index	.954	.946	.959	.970	.944	.976	.937	.974	.980		.997
	Global Executive Composite	.966	.958	.957	.967	.941	.973	.949	.971	.993	.997	
BRIEF-SR Scores	Emotional Control	.124	.088	.083	.094	.082	.098	.098	.092	.107	.094	.100
	Inhibit	.251	.249	.240	.228	.210	.235	.224	.223	.251	.236	.243
	Monitor	.223	.216	.208	.202	.181	.210	.207	.199	.224	.208	.215
	Organization of Materials	.097	.110	.099	.104	.073	.106	.108	.093	.109	.100	.104
	Plan/Organize	.147	.148	.130	.149	.103	.161	.158	.142	.156	.146	.151
	Shift	.094	.088	.073	.091	.060	.100	.102	.084	.098	.087	.092

	Task Completion	.190	.181	.170	.192	.146	.200	.201	.193	.197	.190	.194
	Working Memory	.117	.112	.087	.107	.082	.118	.115	.109	.119	.107	.112
	Behavioral Regulation Index	.206	.190	.179	.182	.158	.191	.187	.177	.202	.186	.193
	Metacognitive Index	.161	.160	.140	.161	.118	.171	.170	.157	.169	.158	.163
	Global Executive Composite	.192	.184	.167	.180	.144	.190	.187	.176	.195	.181	.187
	Special Education Status	.254	.228	.237	.275	.234	.265	.266	.273	.257	.268	.265
	Age	-.013	-.023	-.006	-.022	-.033	-.019	-.005	-.016	-.015	-.019	-.018
	Gender	-.170	-.218	-.212	-.198	-.178	-.211	-.188	-.203	-.200	-.209	-.207
Ethnicity/Race	American Indian	-.031	-.043	-.033	-.046	-.019	-.044	-.037	-.044	-.038	-.040	-.040
	Asian	-.113	-.099	-.099	-.101	-.115	-.114	-.112	-.107	-.111	-.111	-.112
	African	.131	.138	.118	.120	.131	.125	.125	.133	.137	.130	.133
	Latin	-.072	-.085	-.078	-.066	-.074	-.070	-.065	-.073	-.077	-.074	-.076
	Multiracial	.021	.016	.019	.023	.022	.020	.024	.020	.021	.022	.021
	European	-.100	-.092	-.089	-.093	-.086	-.096	-.097	-.091	-.100	-.095	-.097

		BRIEF-SR Scores										
		Emotional Control	Inhibit	Monitor	Organization Materials	of Plan/Organiz e	Shift	Task Completion	Working Memory	Behavioral Regulation Index	Metacognitive Index	Global Executive Composite
Scantron Scaled Scores	Language Arts	-.063	-.186	-.178	-.143	-.163	-.088	-.234	-.086	-.149	-.176	-.171
	Math	-.006	-.056	-.110	-.111	-.126	-.056	-.227	-.067	-.061	-.151	-.114
	Reading	-.065	-.087	-.129	-.037	-.070	-.108	-.176	-.028	-.109	-.090	-.104
BRIEF Scores	Emotional Control	.124	.251	.223	.097	.147	.094	.190	.117	.206	.161	.192
	Inhibit	.088	.249	.216	.110	.148	.088	.181	.112	.190	.160	.184
	Initiate	.083	.240	.208	.099	.130	.073	.170	.087	.179	.140	.167
	Monitor	.094	.228	.202	.104	.149	.091	.192	.107	.182	.161	.180
	Organization of Materials	.082	.210	.181	.073	.103	.060	.146	.082	.158	.118	.144
	Plan/Organize	.098	.235	.210	.106	.161	.100	.200	.118	.191	.171	.190
	Shift	.098	.224	.207	.108	.158	.102	.201	.115	.187	.170	.187
	Working Memory	.092	.223	.199	.093	.142	.084	.193	.109	.177	.157	.176
	Behavioral Regulation Index	.107	.251	.224	.109	.156	.098	.197	.119	.202	.169	.195
	Metacognitive Index	.094	.236	.208	.100	.146	.087	.190	.107	.186	.158	.181
	Global Executive Composite	.100	.243	.215	.104	.151	.092	.194	.112	.193	.163	.187
BRIEF-SR Scores	Emotional Control		.644	.507	.389	.543	.588	.464	.587	.835	.582	.742
	Inhibit	.644		.639	.486	.634	.593	.535	.639	.888	.669	.816
	Monitor	.507	.639		.412	.620	.585	.511	.569	.765	.621	.728
	Organization of Materials	.389	.486	.412		.619	.506	.543	.581	.541	.741	.682
	Plan/Organize	.543	.634	.620	.619		.752	.768	.771	.761	.930	.897
	Shift	.588	.593	.585	.506	.752		.668	.712	.830	.776	.846

	Task Completion	.464	.535	.511	.543	.768	.668		.693	.652	.878	.813
	Working Memory	.587	.639	.569	.581	.771	.712	.693		.754	.896	.875
	Behavioral Regulation Index	.835	.888	.765	.541	.761	.830	.652	.754		.794	.942
	Metacognitive Index	.582	.669	.621	.741	.930	.776	.878	.896	.794		.952
	Global Executive Composite	.742	.816	.728	.682	.897	.846	.813	.875	.942	.952	
	Special Education Status	.000	.023	.072	.020	.091	.069	.194	.082	.043	.118	.087
	Age	-.078	-.092	-.039	.013	.029	-.018	.075	.028	-.072	.043	-.012
	Gender	.231	.013	-.007	-.025	.006	.061	-.008	.063	.095	.016	.056
Ethnicity/Race	American Indian	-.012	-.013	-.051	-.004	-.022	-.019	-.044	-.012	-.024	-.025	-.026
	Asian	-.036	-.044	-.050	-.046	-.047	-.038	-.096	-.056	-.049	-.070	-.063
	African	.051	.104	.051	.067	.032	.055	.010	.069	.083	.048	.068
	Latin	-.023	-.038	.014	-.044	.002	-.015	.064	-.021	-.024	.005	-.010
	Multiracial	.037	.014	.003	.029	.006	.013	.023	.016	.022	.019	.021
	European	-.073	-.122	-.121	-.041	-.055	-.059	-.067	-.058	-.110	-.064	-.091

				Ethnicity/Race						
		Special Education Status	**Age**	**Gender**	**American Indian**	**Asian**	**African**	**Latin**	**Multiracial**	**European**
Scantron Scaled Scores	**Language Arts**	-.351	.090	.122	.053	.083	-.080	-.128	.025	.030
	Math	-.349	.221	-.008	.062	.114	-.081	-.095	.024	-.005
	Reading	-.375	.210	.070	.044	.071	-.088	-.098	-.002	-.011
BRIEF Scores	**Emotional Control**	.254	-.013	-.170	-.031	-.113	.131	-.072	.021	-.100
	Inhibit	.228	-.023	-.218	-.043	-.099	.138	-.085	.016	-.092
	Initiate	.237	-.006	-.212	-.033	-.099	.118	-.078	.019	-.089
	Monitor	.275	-.022	-.198	-.046	-.101	.120	-.066	.023	-.093
	Organization of Materials	.234	-.033	-.178	-.019	-.115	.131	-.074	.022	-.086
	Plan/Organize	.265	-.019	-.211	-.044	-.114	.125	-.070	.020	-.096
	Shift	.266	-.005	-.188	-.037	-.112	.125	-.065	.024	-.097
	Working Memory	.273	-.016	-.203	-.044	-.107	.133	-.073	.020	-.091
	Behavioral Regulation Index	.257	-.015	-.200	-.038	-.111	.137	-.077	.021	-.100
	Metacognitive Index	.268	-.019	-.209	-.040	-.111	.130	-.074	.022	-.095
	Global Executive Composite	.265	-.018	-.207	-.040	-.112	.133	-.076	.021	-.097
BRIEF-SR Scores	**Emotional Control**	.000	-.078	.231	-.012	-.036	.051	-.023	.037	-.073
	Inhibit	.023	-.092	.013	-.013	-.044	.104	-.038	.014	-.122
	Monitor	.072	-.039	-.007	-.051	-.050	.051	.014	.003	-.121
	Organization of Materials	.020	.013	-.025	-.004	-.046	.067	-.044	.029	-.041
	Plan/Organize	.091	.029	.006	-.022	-.047	.032	.002	.006	-.055
	Shift	.069	-.018	.061	-.019	-.038	.055	-.015	.013	-.059
	Task Completion	.194	.075	-.008	-.044	-.096	.010	.064	.023	-.067
	Working Memory	.082	.028	.063	-.012	-.056	.069	-.021	.016	-.058

	Behavioral Regulation Index	.043	-.072	.095	-.024	-.049	.083	-.024	.022	-.110
	Metacognitive Index	.118	.043	.016	-.025	-.070	.048	.005	.019	-.064
	Global Executive Composite	.087	-.012	.056	-.026	-.063	.068	-.010	.021	-.091
	Special Education Status		.088	-.166	-.035	-.095	-.059	.038	.051	.070
	Age	.088		-.044	-.046	-.032	-.063	.039	.006	-.021
	Gender	-.166	-.044		.007	-.028	.009	-.010	.071	-.016
Ethnicity/Race	**American Indian**	-.035	-.046	.007		-.016	-.066	-.064	-.009	-.030
	Asian	-.095	-.032	-.028	-.016		-.125	-.121	-.017	-.057
	African	-.059	-.063	.009	-.066	-.125		-.508	-.072	-.237
	Latin	.038	.039	-.010	-.064	-.121	-.508		-.070	-.230
	Multiracial	.051	.006	.071	-.009	-.017	-.072	-.070		-.033
	European	.070	-.021	-.016	-.030	-.057	-.237	-.230	-.033	

Appendix 2. MLMs for BRIEF and BRIEF-SR Sub-score

Math

BRIEF

GEC

		β (95% CI)	*t*	*p*
Age		0.32 (0.09 – 0.55)	2.76	.006
Gender		0.08 (-0.14 – 0.29)	0.71	.478
IEP Status		-0.29 (-0.53 – -0.05)	-2.37	.018
Free/Reduced Lunch Eligibility		0.19 (-0.06 – 0.43)	1.50	.134
Race/Ethnicity	Asian	0.87 (0.34 – 1.39)	3.27	.001
	Black	0.58 (0.34 – 0.83)	4.73	< .001
	Latinx	0.38 (0.12 – 0.65)	2.84	.005
	White	0.47 (0.11 – 0.82)	2.59	.010
EF	GEC	-0.20 (-0.36 – -0.03)	-2.35	.019
	GEC × Age	0.18 (-0.08 – 0.44)	1.35	.179

Sub-scores

BRI

		β (95% CI)	*t*	*p*
Age		0.33 (0.11 – 0.56)	2.89	.004
Gender		0.08 (-0.14 – 0.30)	0.72	.472
IEP Status		-0.31 (-0.54 – -0.07)	-2.53	.012
Free/Reduced Lunch Eligibility		0.18 (-0.06 – 0.43)	1.48	.140
Race/Ethnicity	Asian	0.88 (0.36 – 1.40)	3.35	.001
	Black	0.59 (0.35 – 0.84)	4.81	< .001
	Latinx	0.39 (0.13 – 0.66)	2.92	.004
	White	0.48 (0.12 – 0.83)	2.66	.008
EF	BRI	-0.20 (-0.36 – -0.03)	-2.37	.019
	BRI × Age	0.17 (-0.09 – 0.43)	1.30	.197

Emotional Control

		β (95% CI)	*t*	*p*
Age		0.33 (0.09 – 0.56)	2.77	.006
Gender		0.10 (-0.12 – 0.32)	0.92	.357
IEP Status		-0.33 (-0.57 – -0.09)	-2.70	.007
Free/Reduced Lunch Eligibility		0.17 (-0.08 – 0.41)	1.35	.179
Race/Ethnicity	Asian	0.85 (0.32 – 1.38)	3.16	.002
	Black	0.58 (0.33 – 0.82)	4.64	< .001
	Latinx	0.38 (0.11 – 0.65)	2.78	.006
	White	0.50 (0.14 – 0.85)	2.73	.007
EF	Emotional Control	-0.18 (-0.33 – -0.02)	-2.18	.030
	Emotional Control × Age	0.17 (-0.08 – 0.43)	1.33	.185

Inhibit

		β (95% CI)	*t*	*p*
Age		0.35 (0.13 – 0.57)	3.11	.002
Gender		0.08 (-0.14 – 0.29)	0.71	.479
IEP Status		-0.32 (-0.56 – -0.09)	-2.71	.007
Free/Reduced Lunch Eligibility		0.20 (-0.05 – 0.44)	1.59	.113
Race/Ethnicity	Asian	0.96 (0.44 – 1.47)	3.67	< .001
	Black	0.60 (0.36 – 0.85)	4.89	< .001
	Latinx	0.42 (0.15 – 0.68)	3.10	.002
	White	0.48 (0.13 – 0.84)	2.71	.007
EF	Inhibit	-0.18 (-0.33 – -0.02)	-2.23	.027
	Inhibit × Age	0.19 (-0.04 – 0.42)	1.65	.101

Shift

		β (95% CI)	*t*	*p*
Age		0.37 (0.13 – 0.61)	3.07	.002
Gender		0.09 (-0.12 – 0.31)	0.86	.388
IEP Status		-0.30 (-0.54 – -0.06)	-2.50	.013
Free/Reduced Lunch Eligibility		0.17 (-0.07 – 0.42)	1.40	.163
Race/Ethnicity	Asian	0.92 (0.40 – 1.45)	3.45	.001
	Black	0.62 (0.37 – 0.87)	4.85	< .001
	Latinx	0.41 (0.15 – 0.68)	3.04	.003
	White	0.52 (0.16 – 0.88)	2.86	.005
EF	Shift	-0.23 (-0.38 – -0.08)	-2.98	.003
	Shift × Age	0.05 (-0.17 – 0.28)	0.48	.635

MCI

		β (95% CI)	*t*	*p*
Age		0.31 (0.08 – 0.54)	2.68	.008
Gender		0.08 (-0.14 – 0.30)	0.73	.467
IEP Status		-0.28 (-0.52 – -0.04)	-2.31	.022
Free/Reduced Lunch Eligibility		0.19 (-0.06 – 0.43)	1.52	.131
Race/Ethnicity	Asian	0.86 (0.34 – 1.38)	3.24	.001
	Black	0.58 (0.33 – 0.82)	4.67	< .001
	Latinx	0.38 (0.11 – 0.64)	2.79	.006
	White	0.46 (0.11 – 0.82)	2.56	.011
EF	MCI	-0.19 (-0.36 – -0.03)	-2.32	.022
	MCI × Age	0.18 (-0.08 – 0.44)	1.35	.179

Initiate

		β (95% CI)	*t*	*p*
Age		0.32 (0.09 – 0.55)	2.71	.007
Gender		0.10 (-0.12 – 0.32)	0.90	.369
IEP Status		-0.33 (-0.58 – -0.09)	-2.71	.007
Free/Reduced Lunch Eligibility		0.18 (-0.07 – 0.43)	1.45	.149
Race/Ethnicity	Asian	0.91 (0.38 – 1.44)	3.38	.001
	Black	0.57 (0.32 – 0.82)	4.54	< .001
	Latinx	0.39 (0.12 – 0.66)	2.84	.005
	White	0.51 (0.15 – 0.87)	2.78	.006
EF	Initiate	-0.12 (-0.28 – 0.03)	-1.54	.124
	Initiate × Age	0.19 (-0.07 – 0.45)	1.47	.144

Monitor

		β (95% CI)	*t*	*p*
Age		0.33 (0.10 – 0.55)	2.84	.005
Gender		0.10 (-0.11 – 0.32)	0.95	.345
IEP Status		-0.29 (-0.54 – -0.05)	-2.40	.017
Free/Reduced Lunch Eligibility		0.18 (-0.06 – 0.43)	1.48	.141
Race/Ethnicity	Asian	0.90 (0.38 – 1.42)	3.42	.001
	Black	0.57 (0.33 – 0.81)	4.61	< .001
	Latinx	0.38 (0.11 – 0.65)	2.80	.006
	White	0.48 (0.12 – 0.84)	2.64	.009
EF	Monitor	-0.21 (-0.37 – -0.05)	-2.54	.012
	Monitor × Age	0.11 (-0.15 – 0.36)	0.84	.405

Organization of Materials

		β (95% CI)	*t*	*p*
Age		0.31 (0.08 – 0.54)	2.63	.009
Gender		0.09 (-0.13 – 0.31)	0.82	.411
IEP Status		-0.30 (-0.54 – -0.06)	-2.42	.016
Free/Reduced Lunch Eligibility		0.20 (-0.05 – 0.44)	1.59	.114
Race/Ethnicity	Asian	0.87 (0.35 – 1.40)	3.28	.001
	Black	0.58 (0.33 – 0.82)	4.66	< .001
	Latinx	0.38 (0.11 – 0.65)	2.81	.005
	White	0.47 (0.11 – 0.83)	2.59	.010
EF	Organization of Materials	-0.17 (-0.32 – -0.02)	-2.18	.031
	Organization of Materials × Age	0.17 (-0.08 – 0.42)	1.33	.185

Plan/Organize

		β (95% CI)	*t*	*p*
Age		0.33 (0.10 – 0.55)	2.82	.005
Gender		0.09 (-0.13 – 0.30)	0.78	.439
IEP Status		-0.31 (-0.55 – -0.06)	-2.50	.013
Free/Reduced Lunch Eligibility		0.19 (-0.05 – 0.44)	1.55	.122
Race/Ethnicity	Asian	0.87 (0.35 – 1.40)	3.29	.001
	Black	0.58 (0.33 – 0.82)	4.64	< .001
	Latinx	0.38 (0.12 – 0.65)	2.84	.005
	White	0.47 (0.11 – 0.82)	2.58	.011
EF	Plan/Organize	-0.15 (-0.32 – 0.02)	-1.73	.085
	Plan/Organize × Age	0.22 (-0.05 – 0.49)	1.62	.107

Working Memory

		β (95% CI)	*t*	*p*
Age		0.33 (0.10 – 0.57)	2.79	.006
Gender		0.08 (-0.13 – 0.30)	0.75	.453
IEP Status		-0.27 (-0.51 – -0.03)	-2.18	.030
Free/Reduced Lunch Eligibility		0.17 (-0.08 – 0.41)	1.36	.176
Race/Ethnicity	Asian	0.85 (0.33 – 1.38)	3.20	.002
	Black	0.59 (0.34 – 0.83)	4.76	< .001
	Latinx	0.38 (0.11 – 0.64)	2.81	.005
	White	0.48 (0.13 – 0.83)	2.68	.008
EF	Working Memory	-0.26 (-0.42 – -0.10)	-3.23	.001
	Working Memory × Age	0.11 (-0.13 – 0.34)	0.90	.370

BRIEF-SR

GEC

		β (95% CI)	*t*	*p*
Age		0.49 (0.22 – 0.75)	3.59	< .001
Gender		0.19 (-0.03 – 0.40)	1.70	.091
IEP Status		-0.57 (-0.82 – -0.32)	-4.54	< .001
Free/Reduced Lunch Eligibility		0.30 (0.03 – 0.56)	2.19	.029
Race/Ethnicity	Asian	1.17 (0.66 – 1.69)	4.48	< .001
	Black	0.56 (0.29 – 0.83)	4.05	< .001
	Latinx	0.49 (0.21 – 0.78)	3.41	.001
	White	0.49 (0.11 – 0.87)	2.52	.012
EF	GEC	-0.21 (-0.39 – -0.02)	-2.18	.031
	GEC × Age	0.04 (-0.23 – 0.32)	0.31	.754

Sub-scores

BRI

		β (95% CI)	*t*	*p*
Age		0.46 (0.19 – 0.72)	3.40	.001
Gender		0.19 (-0.03 – 0.41)	1.73	.085
IEP Status		-0.59 (-0.84 – -0.33)	-4.60	< .001
Free/Reduced Lunch Eligibility		0.27 (0.00 – 0.54)	1.98	.049
Race/Ethnicity	Asian	1.20 (0.68 – 1.73)	4.55	< .001
	Black	0.55 (0.28 – 0.83)	3.99	< .001
	Latinx	0.51 (0.23 – 0.80)	3.52	.001
	White	0.52 (0.13 – 0.90)	2.66	.009
EF	BRI	-0.15 (-0.34 – 0.04)	-1.60	.110
	BRI × Age	0.05 (-0.23 – 0.33)	0.36	.720

Emotional Control

		β (95% CI)	*t*	*p*
Age		0.44 (0.17 – 0.71)	3.22	.001
Gender		0.19 (-0.04 – 0.42)	1.62	.107
IEP Status		-0.61 (-0.86 – -0.36)	-4.73	< .001
Free/Reduced Lunch Eligibility		0.22 (-0.05 – 0.49)	1.62	.107
Race/Ethnicity	Asian	1.23 (0.70 – 1.76)	4.56	< .001
	Black	0.54 (0.26 – 0.82)	3.83	< .001
	Latinx	0.55 (0.26 – 0.85)	3.75	< .001
	White	0.60 (0.21 – 0.98)	3.03	.003
EF	Emotional Control	-0.05 (-0.26 – 0.16)	-0.46	.646
	Emotional Control × Age	0.09 (-0.23 – 0.41)	0.58	.565

Inhibit

		β (95% CI)	*t*	*p*
Age		0.44 (0.18 – 0.71)	3.26	.001
Gender		0.18 (-0.04 – 0.40)	1.63	.105
IEP Status		-0.59 (-0.84 – -0.34)	-4.65	< .001
Free/Reduced Lunch Eligibility		0.26 (-0.01 – 0.53)	1.92	.056
Race/Ethnicity	Asian	1.21 (0.68 – 1.73)	4.55	< .001
	Black	0.57 (0.29 – 0.84)	4.06	< .001
	Latinx	0.52 (0.23 – 0.81)	3.55	< .001
	White	0.54 (0.16 – 0.93)	2.79	.006
EF	Inhibit	-0.12 (-0.29 – 0.05)	-1.41	.159
	Inhibit × Age	0.07 (-0.18 – 0.32)	0.53	.595

Monitor

		β (95% CI)	*t*	*p*
Age		0.45 (0.19 – 0.71)	3.38	.001

		β (95% CI)	*t*	*p*
Gender		0.17 (-0.05 – 0.38)	1.49	.137
IEP Status		-0.58 (-0.83 – -0.33)	-4.55	< .001
Free/Reduced Lunch Eligibility		0.27 (0.01 – 0.54)	2.03	.043
Race/Ethnicity	Asian	1.20 (0.68 – 1.73)	4.56	< .001
	Black	0.56 (0.29 – 0.83)	4.03	< .001
	Latinx	0.53 (0.24 – 0.81)	3.65	< .001
	White	0.54 (0.16 – 0.92)	2.80	.006
EF	Monitor	-0.21 (-0.38 – -0.03)	-2.26	.025
	Monitor × Age	0.00 (-0.28 – 0.28)	0.01	.996

Shift

		β (95% CI)	*t*	*p*
Age		0.46 (0.19 – 0.73)	3.38	.001
Gender		0.19 (-0.03 – 0.41)	1.70	.091
IEP Status		-0.57 (-0.82 – -0.32)	-4.44	< .001
Free/Reduced Lunch Eligibility		0.24 (-0.02 – 0.51)	1.79	.075
Race/Ethnicity	Asian	1.25 (0.73 – 1.77)	4.76	< .001
	Black	0.55 (0.28 – 0.83)	3.98	< .001
	Latinx	0.53 (0.24 – 0.82)	3.61	< .001
	White	0.53 (0.14 – 0.92)	2.71	.007
EF	Shift	-0.15 (-0.33 – 0.04)	-1.54	.124
	Shift × Age	0.03 (-0.24 – 0.30)	0.24	.808

MCI

		β (95% CI)	*t*	*p*
Age		0.50 (0.23 – 0.77)	3.64	< .001
Gender		0.17 (-0.04 – 0.39)	1.60	.111
IEP Status		-0.56 (-0.81 – -0.32)	-4.50	< .001
Free/Reduced Lunch Eligibility		0.30 (0.03 – 0.56)	2.22	.028
Race/Ethnicity	Asian	1.16 (0.65 – 1.67)	4.48	< .001
	Black	0.56 (0.29 – 0.82)	4.08	< .001
	Latinx	0.49 (0.20 – 0.77)	3.39	.001
	White	0.49 (0.11 – 0.87)	2.56	.011
EF	MCI	-0.24 (-0.42 – -0.05)	-2.47	.014
	MCI × Age	0.02 (-0.25 – 0.30)	0.15	.878

Plan/Organize

		β (95% CI)	*t*	*p*
Age		0.48 (0.21 – 0.75)	3.52	.001
Gender		0.16 (-0.05 – 0.38)	1.49	.138
IEP Status		-0.57 (-0.82 – -0.32)	-4.54	< .001
Free/Reduced Lunch Eligibility		0.30 (0.03 – 0.56)	2.21	.028
Race/Ethnicity	Asian	1.21 (0.70 – 1.72)	4.67	< .001
	Black	0.56 (0.29 – 0.83)	4.10	< .001
	Latinx	0.51 (0.23 – 0.80)	3.56	< .001
	White	0.51 (0.13 – 0.89)	2.66	.009
EF	Plan/Organize	-0.21 (-0.40 – -0.03)	-2.26	.025
	x Age	0.04 (-0.23 – 0.31)	0.32	.749

Organization of Materials

		β (95% CI)	*t*	*p*
Age		0.46 (0.19 – 0.72)	3.36	.001
Gender		0.19 (-0.02 – 0.41)	1.75	.082
IEP Status		-0.59 (-0.84 – -0.34)	-4.62	< .001
Free/Reduced Lunch Eligibility		0.27 (0.00 – 0.53)	1.97	.050
Race/Ethnicity	Asian	1.18 (0.65 – 1.70)	4.45	< .001
	Black	0.55 (0.28 – 0.82)	3.97	< .001
	Latinx	0.50 (0.20 – 0.79)	3.36	.001
	White	0.53 (0.15 – 0.92)	2.74	.007
EF	Organization of Materials	-0.14 (-0.33 – 0.05)	-1.46	.147
	Organization of Materials × Age	0.07 (-0.20 – 0.35)	0.52	.603

Task Completion

		β (95% CI)	*t*	*p*
Age		0.53 (0.27 – 0.79)	3.97	< .001
Gender		0.15 (-0.06 – 0.37)	1.42	.158
IEP Status		-0.53 (-0.78 – -0.28)	-4.25	< .001
Free/Reduced Lunch Eligibility		0.26 (0.00 – 0.52)	1.99	.048
Race/Ethnicity	Asian	1.15 (0.64 – 1.66)	4.47	< .001
	Black	0.58 (0.31 – 0.85)	4.30	< .001
	Latinx	0.51 (0.23 – 0.79)	3.61	< .001
	White	0.51 (0.14 – 0.88)	2.72	.007
EF	Task Completion	-0.28 (-0.45 – -0.10)	-3.18	.002
	Task Completion × Age	0.01 (-0.25 – 0.28)	0.11	.916

Working Memory

		β (95% CI)	*t*	*p*
Age		0.48 (0.20 – 0.76)	3.42	.001
Gender		0.18 (-0.03 – 0.40)	1.66	.099
IEP Status		-0.59 (-0.84 – -0.34)	-4.66	< .001
Free/Reduced Lunch Eligibility		0.26 (0.00 – 0.53)	1.94	.054
Race/Ethnicity	Asian	1.20 (0.68 – 1.72)	4.55	< .001
	Black	0.54 (0.27 – 0.81)	3.89	< .001
	Latinx	0.52 (0.23 – 0.81)	3.53	.001
	White	0.55 (0.16 – 0.93)	2.80	.006
EF	Working Memory	-0.15 (-0.35 – 0.05)	-1.49	.138
	Working Memory × Age	0.06 (-0.23 – 0.35)	0.41	.686

Reading

BRIEF

GEC

		β (95% CI)	*t*	*p*
Age		0.36 (0.26 – 0.45)	7.45	< .001
Gender		0.20 (0.05 – 0.36)	2.57	.011
IEP Status		-0.29 (-0.47 – -0.11)	-3.18	.002
Free/Reduced Lunch Eligibility		0.33 (0.14 – 0.51)	3.46	.001
Race/Ethnicity	Asian	0.45 (0.09 – 0.80)	2.44	.015
	Black	0.34 (0.15 – 0.53)	3.56	< .001
	Latinx	0.24 (0.05 – 0.43)	2.50	.013
	White	0.17 (-0.10 – 0.45)	1.25	.211
EF	GEC	-0.29 (-0.39 – -0.19)	-5.52	< .001
	GEC × Age	0.09 (-0.02 – 0.20)	1.59	.113

Sub-scores

BRI

		β (95% CI)	*t*	*p*
Age		0.36 (0.27 – 0.45)	7.57	< .001
Gender		0.20 (0.05 – 0.36)	2.56	.011
IEP Status		-0.30 (-0.48 – -0.12)	-3.33	.001
Free/Reduced Lunch Eligibility		0.33 (0.14 – 0.51)	3.45	.001
Race/Ethnicity	Asian	0.45 (0.09 – 0.80)	2.47	.014
	Black	0.35 (0.16 – 0.54)	3.62	< .001
	Latinx	0.25 (0.06 – 0.43)	2.56	.011
	White	0.17 (-0.10 – 0.45)	1.26	.210
EF	BRI	-0.30 (-0.40 – -0.20)	-5.79	< .001
	BRI × Age	0.09 (-0.02 – 0.20)	1.53	.128

Emotional Control

		β (95% CI)	*t*	*p*
Age		0.36 (0.27 – 0.45)	7.58	< .001
Gender		0.22 (0.06 – 0.37)	2.78	.006
IEP Status		-0.31 (-0.49 – -0.13)	-3.43	.001
Free/Reduced Lunch Eligibility		0.31 (0.13 – 0.50)	3.31	.001
Race/Ethnicity	Asian	0.42 (0.06 – 0.78)	2.30	.022
	Black	0.33 (0.14 – 0.52)	3.47	.001
	Latinx	0.23 (0.04 – 0.42)	2.37	.018
	White	0.18 (-0.10 – 0.45)	1.28	.201
EF	Emotional Control	-0.31 (-0.41 – -0.20)	-5.83	< .001
	Emotional Control × Age	0.08 (-0.03 – 0.20)	1.51	.133

Inhibit

		β (95% CI)	*t*	*p*
Age		0.36 (0.26 – 0.45)	7.49	< .001
Gender		0.21 (0.06 – 0.37)	2.66	.008
IEP Status		-0.34 (-0.51 – -0.16)	-3.72	< .001
Free/Reduced Lunch Eligibility		0.34 (0.15 – 0.53)	3.57	< .001
Race/Ethnicity	Asian	0.50 (0.15 – 0.86)	2.78	.006
	Black	0.34 (0.15 – 0.53)	3.49	.001
	Latinx	0.25 (0.06 – 0.44)	2.61	.009
	White	0.18 (-0.09 – 0.45)	1.30	.196
EF	Inhibit	-0.27 (-0.37 – -0.17)	-5.39	< .001
	Inhibit × Age	0.08 (-0.03 – 0.19)	1.42	.156

Shift

		β (95% CI)	*t*	*p*
Age		0.36 (0.27 – 0.46)	7.63	< .001
Gender		0.21 (0.05 – 0.36)	2.60	.010
IEP Status		-0.31 (-0.49 – -0.13)	-3.37	.001
Free/Reduced Lunch Eligibility		0.32 (0.13 – 0.51)	3.34	.001
Race/Ethnicity	Asian	0.45 (0.09 – 0.81)	2.48	.014
	Black	0.37 (0.17 – 0.56)	3.77	< .001
	Latinx	0.26 (0.07 – 0.45)	2.74	.006
	White	0.20 (-0.08 – 0.47)	1.42	.157
EF	Shift	-0.26 (-0.36 – -0.16)	-5.14	< .001
	Shift × Age	0.09 (-0.02 – 0.19)	1.56	.121

MCI

		β (95% CI)	*t*	*p*
Age		0.35 (0.26 – 0.45)	7.38	< .001
Gender		0.21 (0.05 – 0.36)	2.61	.009
IEP Status		-0.29 (-0.48 – -0.11)	-3.14	.002
Free/Reduced Lunch Eligibility		0.33 (0.14 – 0.52)	3.46	.001
Race/Ethnicity	Asian	0.45 (0.09 – 0.81)	2.45	.015
	Black	0.34 (0.15 – 0.53)	3.50	.001
	Latinx	0.24 (0.05 – 0.43)	2.47	.014
	White	0.18 (-0.10 – 0.45)	1.27	.205
EF	MCI	-0.28 (-0.38 – -0.18)	-5.27	< .001
	MCI × Age	0.09 (-0.02 – 0.20)	1.61	.108

Initiate

		β (95% CI)	*t*	*p*
Age		0.36 (0.26 – 0.46)	7.40	< .001
Gender		0.23 (0.07 – 0.39)	2.84	.005
IEP Status		-0.35 (-0.53 – -0.16)	-3.73	< .001
Free/Reduced Lunch Eligibility		0.33 (0.14 – 0.52)	3.36	.001
Race/Ethnicity	Asian	0.50 (0.14 – 0.86)	2.70	.007
	Black	0.33 (0.14 – 0.53)	3.38	.001
	Latinx	0.25 (0.06 – 0.44)	2.55	.011
	White	0.21 (-0.06 – 0.49)	1.51	.132
EF	Initiate	-0.22 (-0.33 – -0.12)	-4.29	< .001
	Initiate × Age	0.07 (-0.04 – 0.17)	1.26	.209

Monitor

		β (95% CI)	*t*	*p*
Age		0.35 (0.26 – 0.45)	7.31	< .001
Gender		0.22 (0.07 – 0.38)	2.83	.005
IEP Status		-0.30 (-0.49 – -0.12)	-3.26	.001
Free/Reduced Lunch Eligibility		0.33 (0.14 – 0.52)	3.43	.001
Race/Ethnicity	Asian	0.47 (0.11 – 0.83)	2.58	.010
	Black	0.33 (0.14 – 0.52)	3.37	.001
	Latinx	0.24 (0.05 – 0.43)	2.46	.014
	White	0.19 (-0.09 – 0.46)	1.32	.187
EF	Monitor	-0.26 (-0.36 – -0.16)	-4.99	< .001
	Monitor × Age	0.08 (-0.02 – 0.18)	1.53	.128

Organization of Materials

		β (95% CI)	*t*	*p*
Age		0.35 (0.26 – 0.45)	7.28	< .001
Gender		0.22 (0.06 – 0.38)	2.78	.006
IEP Status		-0.32 (-0.50 – -0.14)	-3.44	.001
Free/Reduced Lunch Eligibility		0.34 (0.15 – 0.53)	3.56	< .001
Race/Ethnicity	Asian	0.47 (0.10 – 0.83)	2.54	.012
	Black	0.34 (0.15 – 0.54)	3.54	< .001
	Latinx	0.25 (0.06 – 0.44)	2.59	.010
	White	0.19 (-0.08 – 0.47)	1.36	.174
EF	Organization of Materials	-0.24 (-0.34 – -0.15)	-4.92	< .001
	Organization of Materials × Age	0.11 (0.00 – 0.21)	2.04	.042

Plan/Organize

		β (95% CI)	*t*	*p*
Age		0.36 (0.26 – 0.45)	7.43	< .001
Gender		0.22 (0.06 – 0.37)	2.72	.007
IEP Status		-0.31 (-0.49 – -0.12)	-3.30	.001
Free/Reduced Lunch Eligibility		0.33 (0.14 – 0.52)	3.44	.001
Race/Ethnicity	Asian	0.46 (0.10 – 0.82)	2.49	.013
	Black	0.34 (0.15 – 0.53)	3.48	.001
	Latinx	0.23 (0.04 – 0.43)	2.39	.017
	White	0.18 (-0.10 – 0.45)	1.25	.213
EF	Plan/Organize	-0.26 (-0.37 – -0.16)	-4.94	< .001
	Plan/Organize × Age	0.08 (-0.04 – 0.19)	1.32	.189

Working Memory

		β (95% CI)	*t*	*p*
Age		0.36 (0.26 – 0.45)	7.49	< .001
Gender		0.20 (0.05 – 0.36)	2.56	.011
IEP Status		-0.29 (-0.47 – -0.11)	-3.14	.002
Free/Reduced Lunch Eligibility		0.32 (0.13 – 0.50)	3.36	.001
Race/Ethnicity	Asian	0.43 (0.07 – 0.79)	2.36	.019
	Black	0.34 (0.15 – 0.53)	3.54	< .001
	Latinx	0.24 (0.05 – 0.43)	2.51	.012
	White	0.18 (-0.09 – 0.46)	1.32	.189
EF	Working Memory	-0.29 (-0.39 – -0.19)	-5.53	< .001
	Working Memory × Age	0.10 (0.00 – 0.20)	1.88	.061

BRIEF-SR

GEC

		β (95% CI)	*t*	*p*
Age		0.41 (0.32 – 0.51)	8.58	< .001
Gender		0.39 (0.24 – 0.54)	5.05	< .001
IEP Status		-0.55 (-0.73 – -0.38)	-6.15	< .001
Free/Reduced Lunch Eligibility		0.30 (0.10 – 0.50)	3.01	.003
Race/Ethnicity	Asian	0.59 (0.24 – 0.94)	3.28	.001
	Black	0.27 (0.08 – 0.46)	2.82	.005
	Latinx	0.17 (-0.02 – 0.36)	1.72	.087
	White	0.10 (-0.17 – 0.37)	0.72	.475
EF	GEC	-0.17 (-0.26 – -0.09)	-4.08	< .001
	GEC × Age	0.03 (-0.07 – 0.12)	0.55	.581

Sub-scores

BRI

		β (95% CI)	*t*	*p*
Age		0.40 (0.31 – 0.50)	8.34	< .001
Gender		0.40 (0.25 – 0.56)	5.18	< .001
IEP Status		-0.55 (-0.73 – -0.38)	-6.15	< .001
Free/Reduced Lunch Eligibility		0.30 (0.10 – 0.49)	2.97	.003
Race/Ethnicity	Asian	0.59 (0.24 – 0.94)	3.31	.001
	Black	0.27 (0.08 – 0.46)	2.82	.005
	Latinx	0.16 (-0.03 – 0.35)	1.65	.100
	White	0.10 (-0.18 – 0.37)	0.70	.485
EF	BRI	-0.18 (-0.27 – -0.10)	-4.20	< .001
	BRI × Age	0.03 (-0.06 – 0.12)	0.64	.525

Emotional Control

		β (95% CI)	*t*	*p*
Age		0.40 (0.31 – 0.50)	8.27	< .001
Gender		0.44 (0.28 – 0.60)	5.46	< .001
IEP Status		-0.58 (-0.75 – -0.40)	-6.40	< .001
Free/Reduced Lunch Eligibility		0.27 (0.08 – 0.47)	2.72	.007
Race/Ethnicity	Asian	0.57 (0.21 – 0.92)	3.15	.002
	Black	0.25 (0.06 – 0.44)	2.57	.011
	Latinx	0.16 (-0.03 – 0.35)	1.70	.091
	White	0.13 (-0.14 – 0.40)	0.92	.359
EF	Emotional Control	-0.19 (-0.28 – -0.09)	-3.99	< .001
	Emotional Control × Age	0.02 (-0.08 – 0.11)	0.38	.707

Inhibit

		β (95% CI)	t	p
Age		0.40 (0.31 – 0.50)	8.25	< .001
Gender		0.38 (0.22 – 0.53)	4.84	< .001
IEP Status		-0.57 (-0.75 – -0.39)	-6.23	< .001
Free/Reduced Lunch Eligibility		0.29 (0.09 – 0.49)	2.84	.005
Race/Ethnicity	Asian	0.63 (0.27 – 0.99)	3.47	.001
	Black	0.29 (0.100 – 0.48)	3.02	.003
	Latinx	0.19 (-0.01 – 0.38)	1.90	.058
	White	0.15 (-0.12 – 0.43)	1.10	.270
EF	Inhibit	-0.12 (-0.21 – -0.04)	-2.87	.004
	Inhibit × Age	0.03 (-0.06 – 0.13)	0.72	.470

Monitor

		β (95% CI)	*t*	*p*
Age		0.41 (0.31 – 0.50)	8.50	< .001
Gender		0.37 (0.22 – 0.52)	4.75	< .001
IEP Status		-0.55 (-0.73 – -0.37)	-6.04	< .001
Free/Reduced Lunch Eligibility		0.31 (0.11 – 0.50)	3.05	.002
Race/Ethnicity	Asian	0.61 (0.25 – 0.96)	3.39	.001
	Black	0.27 (0.08 – 0.46)	2.78	.006
	Latinx	0.18 (-0.01 – 0.37)	1.90	.058
	White	0.13 (-0.14 – 0.40)	0.92	.357
EF	Monitor	-0.17 (-0.26 – -0.09)	-3.89	< .001
	Monitor × Age	0.07 (-0.03 – 0.16)	1.44	.151

Shift

		β (95% CI)	t	p
Age		0.40 (0.30 – 0.49)	8.28	< .001
Gender		0.40 (0.24 – 0.55)	5.11	< .001
IEP Status		-0.55 (-0.73 – -0.37)	-6.09	< .001
Free/Reduced Lunch Eligibility		0.28 (0.09 – 0.48)	2.83	.005
Race/Ethnicity	Asian	0.64 (0.29 – 0.98)	3.59	< .001
	Black	0.27 (0.08 – 0.46)	2.85	.005
	Latinx	0.17 (-0.01 – 0.36)	1.82	.070
	White	0.11 (-0.16 – 0.39)	0.82	.414
EF	Shift	-0.18 (-0.26 – -0.09)	-4.04	< .001
	Shift × Age	0.01 (-0.08 – 0.10)	0.16	.871

MCI

		β (95% CI)	*t*	*p*
Age		0.42 (0.32 – 0.51)	8.66	< .001
Gender		0.38 (0.23 – 0.53)	4.88	< .001
IEP Status		-0.56 (-0.74 – -0.38)	-6.18	< .001
Free/Reduced Lunch Eligibility		0.30 (0.10 – 0.49)	2.95	.003
Race/Ethnicity	Asian	0.60 (0.25 – 0.96)	3.36	.001
	Black	0.27 (0.08 – 0.46)	2.84	.005
	Latinx	0.18 (-0.01 – 0.37)	1.89	.060
	White	0.12 (-0.15 – 0.40)	0.90	.372
EF	MCI	-0.15 (-0.24 – -0.07)	-3.61	< .001
	MCI × Age	0.02 (-0.07 – 0.11)	0.40	.690

Plan/Organize

		β (95% CI)	*t*	*p*
Age		0.42 (0.32 – 0.51)	8.60	< .001
Gender		0.37 (0.22 – 0.53)	4.79	< .001
IEP Status		-0.57 (-0.75 – -0.39)	-6.24	< .001
Free/Reduced Lunch Eligibility		0.31 (0.11 – 0.51)	3.03	.003
Race/Ethnicity	Asian	0.62 (0.27 – 0.98)	3.48	.001
	Black	0.28 (0.09 – 0.47)	2.88	.004
	Latinx	0.19 (0.00 – 0.38)	1.95	.052
	White	0.13 (-0.14 – 0.41)	0.97	.335
EF	Plan/Organize	-0.14 (-0.23 – -0.06)	-3.24	.001
	Plan/Organize × Age	0.06 (-0.03 – 0.15)	1.24	.218

Organization of Materials

		β (95% CI)	*t*	*p*
Age		0.41 (0.31 – 0.50)	8.36	< .001
Gender		0.38 (0.22 – 0.53)	4.78	< .001
IEP Status		-0.58 (-0.76 – -0.40)	-6.34	< .001
Free/Reduced Lunch Eligibility		0.28 (0.08 – 0.48)	2.76	.006
Race/Ethnicity	Asian	0.65 (0.29 – 1.00)	3.56	< .001
	Black	0.28 (0.09 – 0.48)	2.89	.004
	Latinx	0.20 (0.01 – 0.39)	2.05	.042
	White	0.19 (-0.09 – 0.46)	1.35	.179
EF	Organization of Materials	-0.08 (-0.17 – 0.00)	-1.96	.051
	Organization of Materials × Age	0.02 (-0.07 – 0.11)	0.47	.641

Task Completion

		β (95% CI)	*t*	*p*
Age		0.42 (0.33 – 0.52)	8.84	< .001
Gender		0.36 (0.21 – 0.51)	4.63	< .001
IEP Status		-0.54 (-0.71 – -0.36)	-5.94	< .001
Free/Reduced Lunch Eligibility		0.29 (0.09 – 0.48)	2.88	.004
Race/Ethnicity	Asian	0.58 (0.22 – 0.93)	3.22	.001
	Black	0.28 (0.09 – 0.46)	2.90	.004
	Latinx	0.20 (0.01 – 0.38)	2.04	.043
	White	0.11 (-0.16 – 0.38)	0.79	.430
EF	Task Completion	-0.19 (-0.28 – -0.10)	-4.32	< .001
	Task Completion × Age	0.01 (-0.08 – 0.10)	0.25	.806

Working Memory

		β (95% CI)	*t*	*p*
Age		0.41 (0.31 – 0.51)	8.42	< .001
Gender		0.40 (0.24 – 0.55)	5.03	< .001
IEP Status		-0.57 (-0.75 – -0.39)	-6.26	< .001
Free/Reduced Lunch Eligibility		0.26 (0.06 – 0.46)	2.59	.010
Race/Ethnicity	Asian	0.65 (0.29 – 1.00)	3.58	< .001
	Black	0.27 (0.08 – 0.46)	2.82	.005
	Latinx	0.19 (0.00 – 0.38)	1.98	.048
	White	0.16 (-0.12 – 0.43)	1.13	.258
EF	Working Memory	-0.13 (-0.21 – -0.04)	-2.99	.003
	Working Memory × Age	-0.02 (-0.12 – 0.07)	-0.47	.637

Language

BRIEF

GEC

		β (95% CI)	*t*	*p*
Age		0.21 (0.08 – 0.34)	3.12	.002
Gender		0.23 (0.05 – 0.41)	2.55	.011
IEP Status		-0.24 (-0.46 – -0.02)	-2.17	.031
Free/Reduced Lunch Eligibility		0.21 (-0.01 – 0.42)	1.91	.057
Race/Ethnicity	Asian	0.42 (-0.02 – 0.85)	1.90	.059
	Black	0.24 (0.01 – 0.46)	2.07	.039
	Latinx	0.13 (-0.09 – 0.36)	1.15	.253
	White	0.08 (-0.26 – 0.42)	0.46	.650
EF	GEC	-0.42 (-0.54 – -0.30)	-6.78	< .001
	GEC × Age	0.04 (-0.12 – 0.19)	0.47	.640

Sub-scores

BRI

		β (95% CI)	*t*	*p*
Age		0.22 (0.09 – 0.35)	3.25	.001
Gender		0.23 (0.05 – 0.41)	2.55	.011
IEP Status		-0.26 (-0.48 – -0.05)	-2.39	.018
Free/Reduced Lunch Eligibility		0.21 (-0.01 – 0.42)	1.88	.061
Race/Ethnicity	Asian	0.43 (0.00 – 0.86)	1.96	.051
	Black	0.24 (0.02 – 0.47)	2.13	.034
	Latinx	0.14 (-0.08 – 0.37)	1.25	.213
	White	0.09 (-0.25 – 0.43)	0.51	.614
EF	BRI	-0.42 (-0.54 – -0.30)	-6.93	< .001
	BRI × Age	0.02 (-0.13 – 0.17)	0.23	.815

Emotional Control

		β (95% CI)	*t*	*p*
Age		0.21 (0.08 – 0.34)	3.17	.002
Gender		0.24 (0.07 – 0.42)	2.72	.007
IEP Status		-0.26 (-0.47 – -0.05)	-2.40	.017
Free/Reduced Lunch Eligibility		0.20 (-0.02 – 0.41)	1.81	.071
Race/Ethnicity	Asian	0.37 (-0.06 – 0.79)	1.70	.091
	Black	0.22 (0.00 – 0.44)	1.99	.048
	Latinx	0.12 (-0.10 – 0.34)	1.06	.291
	White	0.08 (-0.26 – 0.41)	0.45	.656
EF	Emotional Control	-0.46 (-0.58 – -0.34)	-7.52	< .001
	Emotional Control × Age	0.01 (-0.14 – 0.16)	0.13	.897

Inhibit

		β (95% CI)	*t*	*p*
Age		0.23 (0.09 – 0.36)	3.30	.001
Gender		0.26 (0.07 – 0.44)	2.76	.006
IEP Status		-0.32 (-0.54 – -0.11)	-2.93	.004
Free/Reduced Lunch Eligibility		0.22 (0.00 – 0.44)	2.00	.047
Race/Ethnicity	Asian	0.52 (0.08 – 0.95)	2.34	.020
	Black	0.23 (0.01 – 0.46)	2.02	.045
	Latinx	0.16 (-0.07 – 0.39)	1.34	.182
	White	0.10 (-0.24 – 0.45)	0.59	.555
EF	Inhibit	-0.35 (-0.47 – -0.23)	-5.92	< .001
	Inhibit × Age	0.06 (-0.09 – 0.21)	0.78	.436

Shift

		β (95% CI)	*t*	*p*
Age		0.23 (0.10 – 0.37)	3.42	.001
Gender		0.25 (0.07 – 0.43)	2.68	.008
IEP Status		-0.28 (-0.50 – -0.05)	-2.45	.015
Free/Reduced Lunch Eligibility		0.19 (-0.03 – 0.41)	1.68	.095
Race/Ethnicity	Asian	0.43 (-0.01 – 0.87)	1.94	.054
	Black	0.26 (0.03 – 0.49)	2.26	.024
	Latinx	0.16 (-0.07 – 0.39)	1.37	.171
	White	0.14 (-0.20 – 0.48)	0.80	.426
EF	Shift	-0.37 (-0.49 – -0.26)	-6.16	< .001
	Shift × Age	-0.01 (-0.16 – 0.14)	-0.16	.876

MCI

		β (95% CI)	*t*	*p*
Age		0.21 (0.07 – 0.34)	3.04	.003
Gender		0.24 (0.06 – 0.42)	2.58	.010
IEP Status		-0.24 (-0.46 – -0.01)	-2.10	.037
Free/Reduced Lunch Eligibility		0.21 (-0.01 – 0.43)	1.92	.056
Race/Ethnicity	Asian	0.41 (-0.02 – 0.84)	1.88	.062
	Black	0.23 (0.01 – 0.46)	2.02	.044
	Latinx	0.13 (-0.10 – 0.35)	1.09	.277
	White	0.08 (-0.26 – 0.42)	0.45	.655
EF	MCI	-0.41 (-0.53 – -0.29)	-6.58	< .001
	MCI × Age	0.05 (-0.10 – 0.20)	0.62	.534

Initiate

		β (95% CI)	*t*	*p*
Age		0.22 (0.08 – 0.35)	3.22	.001
Gender		0.26 (0.08 – 0.44)	2.80	.005
IEP Status		-0.29 (-0.51 – -0.07)	-2.60	.010
Free/Reduced Lunch Eligibility		0.21 (-0.01 – 0.43)	1.92	.057
Race/Ethnicity	Asian	0.46 (0.03 – 0.90)	2.10	.036
	Black	0.22 (-0.01 – 0.45)	1.91	.058
	Latinx	0.14 (-0.09 – 0.37)	1.21	.228
	White	0.12 (-0.22 – 0.46)	0.69	.490
EF	Initiate	-0.39 (-0.51 – -0.27)	-6.54	< .001
	Initiate × Age	-0.01 (-0.15 – 0.13)	-0.15	.881

Monitor

		β (95% CI)	*t*	*p*
Age		0.20 (0.07 – 0.34)	2.95	.003
Gender		0.28 (0.09 – 0.46)	2.99	.003
IEP Status		-0.27 (-0.49 – -0.05)	-2.39	.018
Free/Reduced Lunch Eligibility		0.21 (-0.01 – 0.43)	1.87	.063
Race/Ethnicity	Asian	0.44 (0.00 – 0.88)	1.98	.049
	Black	0.21 (-0.02 – 0.44)	1.82	.070
	Latinx	0.11 (-0.12 – 0.34)	0.98	.329
	White	0.09 (-0.25 – 0.44)	0.53	.600
EF	Monitor	-0.36 (-0.48 – -0.23)	-5.77	< .001
	Monitor × Age	0.06 (-0.09 – 0.21)	0.82	.412

Organization of Materials

		β (95% CI)	*t*	*p*
Age		0.20 (0.07 – 0.33)	2.96	.003
Gender		0.25 (0.07 – 0.43)	2.73	.007
IEP Status		-0.25 (-0.47 – -0.03)	-2.26	.024
Free/Reduced Lunch Eligibility		0.23 (0.01 – 0.45)	2.10	.037
Race/Ethnicity	Asian	0.41 (-0.02 – 0.84)	1.87	.062
	Black	0.24 (0.01 – 0.46)	2.06	.040
	Latinx	0.14 (-0.09 – 0.36)	1.22	.225
	White	0.09 (-0.25 – 0.43)	0.52	.604
EF	Organization of Materials	-0.39 (-0.51 – -0.28)	-6.68	< .001
	Organization of Materials × Age	0.04 (-0.11 – 0.18)	0.51	.612

Plan/Organize

		β (95% CI)	*t*	*p*
Age		0.22 (0.09 – 0.35)	3.22	.001
Gender		0.25 (0.07 – 0.44)	2.72	.007
IEP Status		-0.27 (-0.49 – -0.05)	-2.38	.018
Free/Reduced Lunch Eligibility		0.21 (-0.01 – 0.43)	1.92	.056
Race/Ethnicity	Asian	0.44 (0.00 – 0.87)	1.97	.050
	Black	0.23 (0.01 – 0.46)	2.02	.044
	Latinx	0.13 (-0.10 – 0.36)	1.10	.274
	White	0.09 (-0.26 – 0.43)	0.50	.618
EF	Plan/Organize	-0.37 (-0.49 – -0.25)	-5.93	< .001
	Plan/Organize × Age	0.06 (-0.10 – 0.21)	0.74	.459

Working Memory

		β (95% CI)	*t*	*p*
Age		0.21 (0.08 – 0.35)	3.14	.002
Gender		0.24 (0.06 – 0.42)	2.61	.010
IEP Status		-0.25 (-0.48 – -0.03)	-2.25	.025
Free/Reduced Lunch Eligibility		0.19 (-0.02 – 0.41)	1.76	.080
Race/Ethnicity	Asian	0.40 (-0.04 – 0.84)	1.81	.072
	Black	0.23 (0.01 – 0.46)	2.04	.042
	Latinx	0.14 (-0.09 – 0.36)	1.19	.237
	White	0.10 (-0.25 – 0.44)	0.56	.576
EF	Working Memory	-0.38 (-0.50 – -0.26)	-6.08	< .001
	Working Memory × Age	0.10 (-0.05 – 0.24)	1.29	.197

BRIEF-SR

GEC

		β (95% CI)	*t*	*p*
Age		0.24 (0.11 – 0.38)	3.58	< .001
Gender		0.40 (0.21 – 0.58)	4.28	< .001
IEP Status		-0.56 (-0.78 – -0.33)	-4.93	< .001
Free/Reduced Lunch Eligibility		0.21 (-0.03 – 0.45)	1.76	.080
Race/Ethnicity	Asian	0.66 (0.24 – 1.08)	3.06	.002
	Black	0.15 (-0.08 – 0.39)	1.28	.202
	Latinx	0.08 (-0.16 – 0.31)	0.63	.527
	White	0.21 (-0.13 – 0.56)	1.23	.219
EF	GEC	-0.23 (-0.34 – -0.13)	-4.43	< .001
	GEC × Age	-0.06 (-0.19 – 0.08)	-0.83	.408

Sub-scores

BRI

		β (95% CI)	*t*	*p*
Age		0.23 (0.09 – 0.36)	3.32	.001
Gender		0.41 (0.22 – 0.59)	4.37	< .001
IEP Status		-0.56 (-0.78 – -0.34)	-4.98	< .001
Free/Reduced Lunch Eligibility		0.20 (-0.04 – 0.44)	1.62	.107
Race/Ethnicity	Asian	0.67 (0.25 – 1.10)	3.12	.002
	Black	0.15 (-0.08 – 0.39)	1.27	.204
	Latinx	0.08 (-0.16 – 0.31)	0.63	.532
	White	0.22 (-0.12 – 0.56)	1.26	.209
EF	BRI	-0.22 (-0.32 – -0.12)	-4.14	< .001
	BRI × Age	-0.05 (-0.19 – 0.08)	-0.78	.436

Emotional Control

		β (95% CI)	*t*	*p*
Age		0.22 (0.09 – 0.36)	3.24	.001
Gender		0.45 (0.26 – 0.64)	4.70	< .001
IEP Status		-0.60 (-0.82 – -0.38)	-5.35	< .001
Free/Reduced Lunch Eligibility		0.14 (-0.10 – 0.38)	1.17	.244
Race/Ethnicity	Asian	0.68 (0.26 – 1.11)	3.15	.002
	Black	0.13 (-0.11 – 0.37)	1.06	.291
	Latinx	0.11 (-0.13 – 0.35)	0.92	.361
	White	0.28 (-0.06 – 0.62)	1.60	.110
EF	Emotional Control	-0.19 (-0.30 – -0.08)	-3.34	.001
	Emotional Control × Age	-0.11 (-0.25 – 0.02)	-1.67	.096

Inhibit

		β (95% CI)	*t*	*p*
Age		0.21 (0.08 – 0.35)	3.12	.002
Gender		0.38 (0.20 – 0.56)	4.07	< .001
IEP Status		-0.57 (-0.79 – -0.35)	-5.04	< .001
Free/Reduced Lunch Eligibility		0.20 (-0.04 – 0.44)	1.67	.096
Race/Ethnicity	Asian	0.68 (0.25 – 1.10)	3.14	.002
	Black	0.18 (-0.06 – 0.42)	1.49	.138
	Latinx	0.09 (-0.15 – 0.33)	0.72	.472
	White	0.24 (-0.11 – 0.58)	1.36	.177
EF	Inhibit	-0.21 (-0.31 – -0.11)	-3.98	< .001
	Inhibit × Age	-0.05 (-0.18 – 0.09)	-0.69	.490

Monitor

		β (95% CI)	*t*	*p*
Age		0.24 (0.10 – 0.37)	3.44	.001
Gender		0.37 (0.19 – 0.56)	3.98	< .001
IEP Status		-0.57 (-0.80 – -0.34)	-4.97	< .001
Free/Reduced Lunch Eligibility		0.21 (-0.03 – 0.45)	1.73	.085
Race/Ethnicity	Asian	0.67 (0.25 – 1.10)	3.11	.002
	Black	0.15 (-0.09 – 0.39)	1.26	.209
	Latinx	0.10 (-0.14 – 0.34)	0.79	.433
	White	0.24 (-0.11 – 0.58)	1.37	.173
EF	Monitor	-0.20 (-0.30 – -0.09)	-3.60	< .001
	Monitor × Age	0.03 (-0.10 – 0.16)	0.48	.633

Shift

		β (95% CI)	*t*	*p*
Age		0.24 (0.10 – 0.37)	3.44	.001
Gender		0.38 (0.20 – 0.57)	4.08	< .001
IEP Status		-0.58 (-0.80 – -0.35)	-5.03	< .001
Free/Reduced Lunch Eligibility		0.18 (-0.06 – 0.42)	1.49	.137
Race/Ethnicity	Asian	0.72 (0.29 – 1.14)	3.32	.001
	Black	0.16 (-0.08 – 0.40)	1.29	.198
	Latinx	0.12 (-0.12 – 0.36)	0.97	.334
	White	0.26 (-0.09 – 0.61)	1.47	.143
EF	Shift	-0.16 (-0.27 – -0.05)	-2.96	.003
	Shift × Age	0.01 (-0.12 – 0.13)	0.14	.890

MCI

		β (95% CI)	*t*	*p*
Age		0.26 (0.12 – 0.39)	3.75	< .001
Gender		0.38 (0.20 – 0.56)	4.12	< .001
IEP Status		-0.56 (-0.78 – -0.34)	-4.96	< .001
Free/Reduced Lunch Eligibility		0.21 (-0.02 – 0.45)	1.77	.077
Race/Ethnicity	Asian	0.66 (0.24 – 1.09)	3.09	.002
	Black	0.16 (-0.08 – 0.39)	1.30	.195
	Latinx	0.09 (-0.14 – 0.33)	0.78	.434
	White	0.24 (-0.11 – 0.58)	1.36	.174
EF	MCI	-0.23 (-0.33 – -0.12)	-4.26	< .001
	MCI × Age	-0.05 (-0.19 – 0.08)	-0.83	.408

Plan/Organize

		β (95% CI)	*t*	*p*
Age		0.25 (0.12 – 0.38)	3.68	< .001
Gender		0.38 (0.20 – 0.56)	4.08	< .001
IEP Status		-0.57 (-0.79 – -0.35)	-5.06	< .001
Free/Reduced Lunch Eligibility		0.24 (0.01 – 0.48)	2.02	.045
Race/Ethnicity	Asian	0.68 (0.26 – 1.09)	3.18	.002
	Black	0.16 (-0.08 – 0.39)	1.31	.191
	Latinx	0.09 (-0.14 – 0.33)	0.76	.446
	White	0.22 (-0.12 – 0.56)	1.26	.208
EF	Plan/Organize	-0.24 (-0.35 – -0.14)	-4.52	< .001
	Plan/Organize × Age	-0.03 (-0.16 – 0.10)	-0.43	.670

Organization of Materials

		β (95% CI)	*t*	*p*
Age		0.25 (0.12 – 0.39)	3.69	< .001
Gender		0.40 (0.21 – 0.58)	4.21	< .001
IEP Status		-0.58 (-0.81 – -0.36)	-5.13	< .001
Free/Reduced Lunch Eligibility		0.22 (-0.02 – 0.46)	1.80	.073
Race/Ethnicity	Asian	0.67 (0.25 – 1.10)	3.12	.002
	Black	0.15 (-0.09 – 0.39)	1.21	.228
	Latinx	0.11 (-0.13 – 0.35)	0.91	.362
	White	0.29 (-0.06 – 0.63)	1.63	.104
EF	Organization of Materials	-0.19 (-0.29 – -0.08)	-3.46	.001
	Organization of Materials × Age	-0.06 (-0.19 – 0.07)	-0.89	.375

Task Completion

		β (95% CI)	*t*	*p*
Age		0.25 (0.12 – 0.39)	3.71	< .001
Gender		0.35 (0.16 – 0.53)	3.72	< .001
IEP Status		-0.55 (-0.77 – -0.33)	-4.84	< .001
Free/Reduced Lunch Eligibility		0.19 (-0.05 – 0.42)	1.55	.122
Race/Ethnicity	Asian	0.66 (0.24 – 1.09)	3.06	.002
	Black	0.17 (-0.07 – 0.41)	1.41	.159
	Latinx	0.13 (-0.11 – 0.36)	1.05	.296
	White	0.25 (-0.09 – 0.60)	1.47	.144
EF	Task Completion	-0.21 (-0.32 – -0.10)	-3.89	< .001
	Task Completion × Age	-0.04 (-0.17 – 0.08)	-0.68	.499

Working Memory

		β (95% CI)	*t*	*p*
Age		0.24 (0.11 – 0.38)	3.54	< .001
Gender		0.39 (0.21 – 0.58)	4.17	< .001
IEP Status		-0.59 (-0.81 – -0.36)	-5.13	< .001
Free/Reduced Lunch Eligibility		0.16 (-0.09 – 0.40)	1.28	.203
Race/Ethnicity	Asian	0.72 (0.28 – 1.15)	3.27	.001
	Black	0.16 (-0.08 – 0.40)	1.28	.200
	Latinx	0.12 (-0.12 – 0.36)	0.99	.326
	White	0.30 (-0.05 – 0.64)	1.71	.089
EF	Working Memory	-0.16 (-0.27 – -0.06)	-3.02	.003
	Working Memory × Age	-0.06 (-0.20 – 0.08)	-0.85	.399