Executive Functioning Predicts Academic Achievement in Middle School: A 4-Year Longitudinal Study

William Ellery Samuels

College of Staten Island/The City University of New York

Nelly Tournaki

College of Staten Island/The City University of New York

Sheldon Blackman

John W. Lavelle Preparatory Charter School

Christopher Zilinski

John W. Lavelle Preparatory Charter School

This is an Accepted Manuscript of an article published by Taylor & Francis Group in

The Journal of Educational Research, available online:

http://www.tandfonline.com/doi/abs/10.1080/00220671.2014.979913?journalCode=vjer20.

Acknowledgments: The authors would like to acknowledge the great contributions of Ken Byalin, Ph.D. and Evelyn Finn, Lavelle Preparatory School; JoAnn Sacks, Ph.D. and Stan Sacks, Ph.D., National Development & Research Institutes, Inc.; Vicki Sudhalter, Ph.D., New York State Office of People with Developmental Disabilities; and the faculty, staff, and students of Lavelle Preparatory School.

Submitted on July 22, 2014 to *The Journal of Educational Research*

Accepted on October 13, 2014

Published online on May 24, 2016

Abstract

Background

Executive functioning (EF) is a strong predictor of children's and adolescents' academic performance. Although research indicates that EF can increase during childhood and adolescence, few studies have tracked the effect of EF on academic performance throughout the middle school grades.

Method

EF was measured at the end of grades 6 – 9 through 21 teachers' and 22 teacher assistants' assessments of 322 adolescents from disadvantaged backgrounds who attended an urban, chartered middle/high school. Assessment of EF was done through the completion of the Behavior Rating Inventory of Executive Function (BRIEF).

Results

BRIEF global executive composite scores (GEC) predicted both current and future English/language arts, mathematics, science, social studies, and Spanish annual grade point averages (GPAs). The effect of BRIEF GEC scores often overshadowed the effects of gender, poverty, and having an Individual Education Plan; the other, non-BRIEF-related effects retained slightly more impact among teacher assistant-derived data than teacher-derived data. The strong relationships between BRIEF GEC scores and these GPAs also remained constant over these four years: There was little evidence that EF changed over the measured grades or that the relationship between EF and grades itself regularly changed.

Discussion

The findings indicate that EF scores during early middle grades can well predict academic performance in subsequent secondary-school grades. Although methodological constraints may have impeded the abilities of other factors (viz., poverty) to be significantly related to GPAs, the effects of EF were strong and robust enough to prompt us to recommend its use to guide long-term, academic interventions.

Key words: Executive functioning, academic achievement, Individual Education Plans (IEPs), middle school, longitudinal research

Executive Functioning Predicts Academic Achievement in Middle School:

A 4-Year Longitudinal Study

Executive functioning (EF) is definable as a set of control processes that allow individuals to manage and direct their attention, thoughts, and actions to meet adaptive goals (Best & Miller, 2010; Blair & Raver, 2012; Diamond & Lee, 2011). Individual differences in EF have been found to have consistent and substantial implications in everyday life (e.g., Monette, Bigras, & Guay, 2011), including in school. Students are expected to perform many, complex, cognitive activities and to exhibit frequent, overt, goal-directed behaviors such as concentrating on a task, attending to a teacher, following rules, and suppressing counter-productive impulses; when students are successful in these tasks, they are considered to be exhibiting high EF (Anderson, 2002; Blair, 2002; Blair & Razza, 2007; Diamond & Lee, 2011; Isquith, Gioia, & Espy, 2004). Although some researchers believe that EF represents a unifying latent psychological construct, most researchers agree that effective EF entails the coordination of several component skills (Garon, Bryson, & Smith, 2008; Kimberg, D'Esposito, & Farah, 1997) such as working memory, inhibitory control, and attentional set shifting (Brocki & Bohlin, 1999; Huizinga, Dolan, & van der Molan, 2006; Miyake et al., 2000; Pennington, 1998; Pennington & Ozonhofff, 1996; van der Sluis, de Jong, & van der Leij, 2007; Welsh, 1991); some (e.g., Baughman & Cooper, 2007; Miyake et al., 2000) have also added complex planning and metacognitive tasks to what comprise EF. Indeed, Roebers, Cimeli, Rothlisberger and Neuenschwander (2012) examined EF and metacognition through separate tasks and found that EF was significantly related to metacognitive control, both cross-sectionally and longitudinally.

Of course, the identification of early, modifiable predictors of achievement—as well as the identification of processes underlying individual variation in performance that are distinct from other known factors, such as psychometric intelligence (or IQ)—can help guide efforts to improve the long-term success of many children and adolescents (Ginsburg, 1997; Rourke & Conway, 1997). Such identification is particularly relevant for improving the success of educational strategies aimed at assisting children who struggle academically (Rack, Snowling, & Olson, 1992; Savage, Pillay, & Melidona, 2007). We believe that EF may be such a predictor.

EF and Academic Achievement

EF contributes to children's abilities to learn (Bull, Johnson, & Roy, 1999; Bull & Scerif, 2001; Duncan et al. 2007; Lehto, 1996; Lorsbach, Wilson, & Reimer, 1996; McLean & Hitch, 1999; Russell, Jarrold, & Henry, 1996; Swanson, 1993, 1999; Swanson, Ashbaker, & Lee, 1996) and to succeed in school (Blair, 2002; Blair & Razza, 2007; Diamond & Lee, 20011). Studies continue to provide evidence that EF is associated with various aspects of academic success among both children (e.g., Checa & Rueda, 2011; Clark, Prior, & Kinsella, 2002; Hughes & Ensor, 2011; Lan et al., 2011; Masten et al., 2012) and adolescents (e.g., Bierman et al., 2009; Kotsopoulos & Lee, 2012; Latzman, Elkovitch, Young, & Clark, 2010; Waber et al., 2006). Although direct effects seem to be especially strong for mathematics (Bull, Espy, Wiebe, Sheffield, & Nelson, 2011; van der Ven, Kroesbergen, Boom, & Leseman, 2012; Lee, Ng, & Ng, 2009), they have also been found to be substantially related to reading, writing, and science achievement (e.g., Monette et al., 2011; St. Clair-Thompson & Gathercole, 2006).

EF and IQ

Related to effortful-control (Liew, 2012), working memory (Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003), and self-regulated learning (Garner, 2009), EF is at least partially independent from intelligence. It fosters academic success beyond that accounted for by IQ (Bucker, Mezzacappa, Elkovitch, Young, & Clark, 2010), and studies have suggested that EF shows a distinct association with concurrent mathematics performance beyond that made by measures of children's general cognitive abilities (Bull & Scerif, 2001; Espy et al., 2004; St. Clair- Thompson &

Gathercole, 2006). Clark, Prichard, and Woodward (2010), for example, found that different measures of EF (including the Stroop task, Wisconsin Naming Task, and counting span) each predict unique aspects of the variance in mathematical ability beyond that accounted for by IQ or reading achievement. Clark, Prichard, and Woodward also predicted early mathematics achievement of six year old children based on the children's EFs when they were four years old: Even when general IQ and concurrent reading achievement were accounted for, EF continued to show a unique relationship with later performance on the Woodcock-Johnson Math Fluency subtest. In addition, Blair (2006), Garon et al. (2008), and Miyake et al. (2000) provided evidence of discriminant validity of EF and IQ tests in adolescents and adults; Masten et al. (2012) studied a sample of preschoolers and reported that a confirmatory factor analysis supported the construct validity of their EF battery as distinct from the general factor in IQ tests. Masten et al. also found that older children and children with higher IQ scores tended to have higher EF scores and that higher levels of EF predicted school adjustment independent of IQ.

EF and Individualized Education Plans

EF has been found to be related to many of the conditions that commonly lead to the creation of an Individual Educational Plan (IEP), which prescribes specific, additional, academic support(s) based on a professional diagnosis. Alloway, Gathercole, Adams, and Willis (2005), for example, found lower levels of EF among older children with disabilities than those without disabilities. Similarly, Semrud-Clikeman, Fine, and Bledsoe (2013) found that children with non-verbal learning disorders or Asperger's Syndrome demonstrated greater EF-related cognitive deficits compared to matched controls, although which areas were deficient depended on the diagnosis, not all areas were deficient, and EF deficits covaried with IQ. After factoring out fine motor skills that may confound performance on EF tasks, Hartman, Houwen, Scherder, and Visscher (2010) still found that older children with Intellectual Disabilities had lower EFs than matched controls. Controlling for IQ, Diamantopoulou, Rydell, Thorell, and Bohlin (2007) found that EF and Attention-Deficit/Hyperactivity Disorder (ADHD) independently predicted academic performance about one year later, but they also found that ADHD and EF were related and that they interacted with special education status: Children with both high levels of ADHD and low levels of EF were most likely to receive special educational supports. Jansen, De Lange, and Van der Molen (2013) reported that adolescents with mild to borderline Intellectual Disabilities demonstrated depressed EF; lower EFs were also related to poorer performances in math. After a five-week intervention, many of the adolescents' math performance improved, but their EFs did not.

As well as fostering school readiness, EF plays an important role in cognitive and social development throughout childhood (Cartwright, 2012; Dilworth-Bart, 2012; Hostinar, Stellern, Schaefer, Carlson, & Gunnar, 2012), and EF is found to be related to at least some non-cognitive challenges that may lead to students being given extra support through an IEP. For example, Hintermair (2013) found that children who are deaf or hard-of-hearing tended to display lower levels of EF as measured by a version of the BRIEF (described in the Methods section, below) than controls displayed; these deficits were less marked among children in inclusive settings than in schools for deaf children, but in both settings predicted communicative competence and a range of behavioral problems including hyperactivity, conduct problems, and even prosociality. Finally, Schuchardt, Bockmann, Bornemann, and Maehler (2013) reported lower EF and working memory among children with Dyslexia and among children with serious deficits in language production and/or reception.

The Development of EF over Time

EF is observable and measurable early in human development and continues to increase into adolescence (Best & Miller, 2010). Pronounced EF improvements in early childhood are observed with respect to the accuracy of performance, likely reflecting children's growing ability to consciously select among different responses (including the ability to inhibit a prepotent response) by reasoning about available options and by switching between task demands while updating the tasks' goals and specifics. Further improvement in EF performance is generally observed as children advance through elementary school; these gains tend to be mirrored by the emergence of a speed-accuracy trade-off that corresponds to children's growing awareness of a discrepancy between task demands, on the one hand, and their own performance, on the other. Children become increasingly capable of integrating different mental representations (e.g., with respect to changing rules) allowing a more accurate awareness of their performance as well as increasingly flexible adjustments of responses (Lyons & Zelazo, 2011). In fact, Friedman, Miyake, Robinson, and Hewitt (2011) examined the developmental trajectories in toddlers' self-restraint and found that individual differences predict their EF 14 years later.

In the present study, the EF of students throughout middle school and into the ninth grade was measured. There is evidence that EF develops rapidly in the preschool years (Zelazo & Bauer, 2013) but increases more slowly during adolescence (Best & Miller, 2010; Davidson, Amso, Anderson, & Diamond, 2006; Huizinga, Dolan, & van der Molen, 2006; Huizinga & van der Molen, 2007; Luciana, Conklin, Hooper, & Yarger, 2005; Romine & Reynolds, 2005; Somsen, 2007). For example, investigating a large sample of 2,036 children aged 5 to 17 years, Best, Miller, and Naglieri (2011) found that performance on EF-related tasks improved at least until age 15, although improvements in the later years were less pronounced. Therefore EF—and its relationship to academic success—may continue to change significantly throughout the middle school grades, although deleterious factors such as Autism Spectrum Disorder (Rosenthal et al., 2013), early life trauma (Hostinar et al., 2012), poverty, adversity, toxins, and neglect (Blair &

Raver, 2012; Shonkoff, 2011) may impede its growth while enrichment may encourage it (Davis et al., 2011; Diamond & Lee, 2011).

Assessment of EF

There are two ways by which EF is assessed: performance-based tests and rating scales. Historically, the assessment of EF in clinical and research settings relied on the former, the construct-related validity of which is backed by a substantial body of research (e.g., Anderson, Anderson, Jacobs, & Spencer-Smith, 2008). However, a number of authors have argued that performance-based measures have limited ecological-related validity—that is, there is little relationship between a neuropsychological test and the participant's behavior in real-world settings (Isquith, Roth, & Gioia, 2013; Sbordone, 1996). In addition, several researchers have argued that rating scales, by their very nature, capture well observations of every-day, real-world behaviors (Gioia & Isquith, 2004; Silver, 2000).

In the present study, one of the existing rating scales, the Behavior Rating Inventory of Executive Function (BRIEF, Gioia, Andrews, Epsy, & Isquith, 2003) was chosen to investigate the relationship of EF with academic achievement. The validity of the BRIEF has been well supported, and it has been used in several studies. Waber et al. (2006), for example, examined relationships between ratings of EF and academic performance and found that teacher ratings on the BRIEF were the best predictors of performance on statewide academic testing relative to other EF performance measures or broadband rating scales. Other researchers have also reported associations between BRIEF scores and both reading (Locascio, Mahone, Eason, & Cutting, 2010) and mathematics skills (Clark, Pritchard, & Woodward, 2010).

The Present Study

Most existing studies have employed cross-sectional, rather than longitudinal, research designs, making it difficult to ascertain changes over time. Non-longitudinal designs may lead to identification of academic problems late in the students' career because they have already been failing for quite some time and interventions may become more difficult (Lyon, Fletcher, & Barnes, 2003). Even researchers who incorporate a longitudinal approach usually collect data on only twice in the participants' life (e.g., Roebers, Cimeli, Rothlisberger, & Neuenschwander, 2012; Clark, Prichard, & Woodward, 2010). For the present study, data were collected over four consecutive academic years. More specifically, the present study examined teachers' and teacher assistants' ratings of EF of middle and high school students (grades 6 – 9) attending an urban charter school in one of New York City's boroughs. Each student's EF was measured with the BRIEF by both a teacher and teacher assistant; their academic performance in a number of courses was measured every year for four consecutive academic years.

The goals of the present study were to (1) assess how much EF can add to the prediction of students' GPAs in English/language arts (ELA), mathematics, social studies, science, and Spanish courses beyond that made by IEPs, gender, and poverty, and (2) examine the development of the relationship between EF and GPAs over time, i.e., whether the development of higher EF comes to affect grades more strongly. These relationships were assessed for when EF was measured by both teachers and teacher assistants.

The first hypothesis was that EF makes a unique contribution to the predictions of students' academic performance made by whether or not a student has an IEP, their gender, and poverty level. The second hypothesis was that EF improves over the years and perhaps has an increasingly greater effect on academics. Given the rarity of being able to study all of these relationships over time with a general, disadvantaged sample of students, recommendations for policy and practice will be made based on our findings.

Method

Context

The study was conducted at an urban charter school in New York City. The mission of the school—as given in the school's charter—is to provide "a rigorous college preparatory education that equips and empowers . . . all students, including those living with emotional challenges." All classrooms include students with and without disabilities in a 40%:60% ratio; the maximum number of students in each class is 17 (i.e., about 7 students with and 10 without disabilities). Students are fully integrated "to break down barriers through the power of [students'] 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."

Participants

The students, whose EFs were rated through the BRIEF, were aged 12 to 15 years in grades 6 through 9. The ethnicity of the student population is 32% Hispanic; among the non-Hispanic students, 42% are African-American, 8% are Asian-American, and 17% are European-American. The school first opened for the 2009-2010 academic year with a single sixth grade class. The school added a new grade each year as this cohort progressed; this growth is reflected in Table 1.

Twenty-one teachers and 22 teacher assistants participated in this study by completing the BRIEF for students in their classes. All teachers held teaching certifications by the New York State Education Department; all teaching assistants had four years of college but were not necessarily certified.

Materials

Executive functioning: BRIEF. The BRIEF (Gioia, Andrews Epsy, & Isquith, 2003) is probably the most well-known paper-and-pencil measure of EF. It is an 86-item questionnaire which includes five different sub-scales assessing a subject's inhibition, shifting, emotional control, working memory, and planning/organization. Examples of items include how often the subject "needs help from an adult to stay on task," "becomes overwhelmed by large assignments," "makes careless errors," "does not take initiative," and "interrupts others." It has demonstrated good test-retest reliability (Gioia, Isquith, Guy, & Kenworthy, 2000), and has been shown to differentiate between comparison controls on the one hand and children with various disabilities such as ADHD, Reading Disorder, Autism Spectrum Disorder, and Traumatic Brain Injury on the other hand (Gioia & Isquith, 2004; Gioia, Isquith, Kenworthy, & Baron, 2002; Mahone et al., 2002).

The BRIEF is published by Psychological Assessment Resources, Inc. (PAR), from whom the right to use it was obtained. The BRIEF was completed by 21 teachers and 22 teacher assistants who knew the students well. Although the BRIEF produces sub-domain scores, only results concerning the overall General Executive Composite (GEC) score are presented here. The GEC is composed of both behavioral and cognitive functions; separate analyses of these two sub-domains found complementary results.

Academic performance: GPA. Academic performance was measured through students' annual grade point averages (GPAs) for the following courses: English/language arts (ELA), mathematics, science, social studies, and Spanish. Grades in other areas such as movement (a part of physical education) and music were also available but not analyzed here since they are less generalizable. In addition, overall, inter-course GPA was not analyzed since it would either contain these other, less-easily generalized grades or be redundant with those included here. Middle school GPAs are among the best predictors of not only high school GPAs but also whether students ultimately graduate high school (Lohmeier & Raad, 2012); they have also been found to be strongly related to performance on standardized exams, such as the Stanford Test of Basic Skills (Wentzel, 1993). Although related to scores on standardized tests, high school grades have been found to be the best predictor of short- and long-term college grades—not only out-predicting, e.g., SAT scores, but also being less related to family income or parental education than SAT scores (Geiser & Santelices, 2007).

Procedure

The BRIEF was distributed to the teachers and teacher assistants by the school administration near the end of every academic year. Each student's name was given once, after randomization within constraints, to one of their teachers and to one of their teacher assistants. Teachers and teacher assistants rated these students using the BRIEF and returned all completed forms within one week of initial administration. With institutional and school IRB permission to conduct the study, the BRIEF and academic data were linked, anonymized, and then furnished to the authors for analysis.

Analyses

Zero-order analyses, e.g., of how many boys versus girls had IEPs, were conducted with standard chi-square tests, Pearson product-moment correlations, *t*-tests, etc. The results of these analyses will be presented first to establish the context for the main analyses testing the study's two main hypotheses.

These hypotheses were tested through taxonomies of multilevel models of change using full maximum likelihood estimations. Separate models were run with each of the various course GPAs as the only criterion variable. Different models were also run for BRIEF GEC scores computed from teacher ratings and from teacher assistant ratings. Finally, to better show the effect of BRIEF GEC scores relative to the other variables (i.e., gender, lunch status, IEPs, and time), we first ran models with all terms *except* those related to EF. We then added BRIEF-related terms to the models, noting not only the effects of EF as measured by BRIEF GEC scores on the various GPAs and any changes in the effect over time, but also how the pattern of significant effects changed, changes in model fit, and how predictive were students' initial scores on academic performance in later years.

Conduct multilevel models of change were used because they handle data well when not all (but nonetheless enough) participants produce data at every wave and when the lengths of time at each wave vary (Singer & Willett, 2004), as were the cases here. Using full maximum likelihood estimations—as opposed to restricted maximum likelihood estimations—allows tests to be conducted on all factors in the model, but require greater assumptions be made about the data that are unlikely to be violated by the data analyzed here.

In addition to analyses related to fixed factors like gender and IEP status, these multilevel models of change provide other relevant statistics: the covariance and time residuals and the deviance statistic. Covariance residuals indicate how much information remains unexplained at different levels of the model; a significant covariance residual suggests that considerably more information remains about what affects the given GPA than is covered by the factors in the respective model. Covariance residuals also indicate the extent to which the initial status of the predictor terms is related to later academic performance, i.e., how well sixth grade levels can be used to estimate GPAs in the seventh, eighth, and ninth grades. The time residuals test whether there is more to the changes over time than is explained by the model's factors. The deviance statistic is a measure of overall model fit; smaller numbers indicate that the model's parameters better accounts for the data than larger numbers. The deviance statistic for each model represents a benchmark against which to test for any improvements made by additions made to the successive model; significantly large changes in the deviance statistic indicates that the successive model better fits the data.

Dummy codes were used for gender (where males = 0 and females = 1) and whether or not a student had an IEP (referred to as "IEP status" where having an IEP = 1 and having no diagnosed need = 0). “Lunch status” was coded such that receiving a free lunch = −1, a reduced lunch = 0, and a paid lunch = 1. All other variables (BRIEF GEC scores, time, and GPAs) were standardized across the entire dataset. Time was measured as standardized scores of students’ Julian age on the day the BRIEF was completed, centered on the students’ tenth birthdays. The α-levels used to test significance were set at .01 to reduce the chance of experiment-wise Type I (false positive) errors. Data were removed case-wise from each set of models, resulting in slightly different results for the non-final models that would ultimately include teacher BRIEF GEC scores and teaching assistant BRIEF GEC scores.

All analyses were conducted with R, version 3.0.2 (R Core Team, 2013), except for the multilevel change models, which were conducted with SPSS, version 20, since SPSS’s MIXED command uses Satterthwaite’s approximation (Satterthwaite, 1946; Welch, 1947) to compute denominator degrees of freedom. R packages used included psych (Revelle, 2014) and nlme (Pinheiro et al., 2013).

Results

Descriptive and Overall Statistics

Student information. Most students had data collected about them each year they attended the school. However, data were not always available for each student every year: attrition, absences, incomplete instruments, lack of matching academic records, etc. precluded collecting sufficient data on every student; the analyses here include only those students for whom data existed on all respective measures.

Data were collected on 386 students of which 210 were boys and 176 were girls. Table 2 presents the means and standard deviations for girls’ and boys’ BRIEF GEC scores as reported by their teachers and teacher assistants; ages; and annual GPAs in ELA, mathematics, science, social studies, and Spanish. Scores are given for each grade; note that grade levels here does not represent a cohort, but summaries of all students' scores when they were in that grade level. Also note that, per convention, lower BRIEF GEC scores denote *higher* levels of EF. GPAs were computed on an 11-point scale where A+ = 10, A = 9, ... C- = 2, D = 1, and F = 0.

Poverty was measured by whether the students were eligible for free or reduced-price school lunches. Most of the students attending the school live in poor families: 83% of the students receive free (68%) or reduced-price (15%) school lunches.

Correlations between overall BRIEF scores and overall GPAs. A first bird's-eye view of the relationship between EF and academic success can be made by looking at overall BRIEF GEC scores and overall GPAs. The correlation between the mean teacher BRIEF scores (averaged across all available waves for each student) and the mean teacher assistant BRIEF scores was .68 (t_{230} = 14.0, p< .001). Both mean teacher and mean teacher assistant BRIEF scores also correlated highly with the GPAs averaged across all years. Mean teacher BRIEF correlations ranged from -.58 (Spanish, $t_{207} = -10.32$, p< .001) to -.68 (social studies, $t_{206} = -13.03$, p< .001); mean teaching assistant BRIEF correlations ranged from -.45 (Spanish, $t_{206} = -7.33$, p< .001) to -.53 (science, $t_{206} = -9.04$, p< .001). GPAs also correlated highly with each other, ranging from .73 (Spanish and social studies, t_{237} = 16.24, p< .001) to .82 (ELA and math, t_{239} = 22.19, p< .001).

IEP status and student characteristics. The IEP status was known for 322 of the 386 students. Seventy-eight (23.6%) of these 332 students had IEPs. It is not known how many of the remaining 64 students had IEPs, so these students were not included in further analyses. Collection of IEP data began three years after the school had opened, therefore nearly all (62) of the 64 students for whom the IEP status was not known were students who left the school during the first two years. Although we do not know the IEP status of those students, the admission policy of the school has not changed, so the ratio is likely similar for those who left. In any case, all analyses only included students without missing data.

IEP status and overall GPAs. Table 3 presents the overall mean BRIEF GEC scores and GPAs for each student across all waves based on their IEP status. A series of one-way ANOVAs on these mean scores tested overall differences in BRIEF scores and GPAs for those students who did and did not have IEPs. All seven of these tests were significant (smallest $F_{1, 199} = 11.59$ for teacher assistant BRIEF & IEP status; all per-comparison *p*s < .007 using a Bonferroni correction since the independence of the tests could not be ensured) supporting the findings that students with IEPs were more likely to have higher mean BRIEF scores (i.e., be rated as having lower EFs) and lower mean grades than those without IEPs.

IEP status and student gender. Table 4 presents the genders of students with and without IEPs for those who remained in school long enough for us to know their IEP status. A Pearson's chi-square with Yate's correction (which helps the test approximate a continuous distribution; Everitt & Skrondal, 2010) found that boys were more likely to have IEPs than girls ($\chi_2^2 = 4.38$, $p = .036$).

IEP status and poverty. Table 5 displays the numbers of students receiving lunch aid, also as a function of IEP status. A Pearson's chi-square test found no relationship between the lunch and IEP statuses ($\chi_2^2 = 2.49$, $p = .289$).

Effects of Student Characteristics, IEP Status, and BRIEF GEC Scores on Course GPAs

The study's goals are primarily addressed through series of multilevel models of change in course GPAs. Initial models' predictors included gender, lunch status, IEP status, and change in GPAs over time. To each of these models was then added a term for BRIEF GEC scores (derived from either teachers' or teacher assistants' ratings of the students) and an interaction term for BRIEF scores x time. Significant BRIEF terms indicated that EF adds to the prediction of course GPAs above that made by gender, lunch status, IEP status, and time (hypothesis 1); significant BRIEF scores x time terms indicate that the relationship between BRIEF scores and GPAs changed over time (hypothesis 2). Analyzing the goals through pairs of models—instead of one omnibus model with all terms included—allows a closer analysis of both how the pattern of significant terms changes with the introduction of BRIEF-related terms and how much the addition of those terms allows the model to better fit the data.

Summary of effects. The general patterns of significant terms are presented in Figure 1 where the pairs of models are shown broken down by teacher and teacher-assistant models and then by course. Gender was significant in several of the initial models, as was the effect of time on GPAs. Lunch status was never significant.

The pattern of significant terms changed markedly when BRIEF GEC-related factors were added. With one exception (ELA GPAs with teacher-assistant-derived BRIEF scores), BRIEF-related terms accounted for all of the relationships with GPAs among these terms. In most cases, too, the relationship between BRIEF scores and GPAs was stable: Only when predicting social studies and Spanish GPAs among the teacher assistant models were the BRIEF x time interaction terms significant. The details and implications of these patterns are discussed next.

Initial models: gender, lunch status, IEP status, and time. The backdrop against which EF was examined were models comprised of gender, lunch status, IEP status, and time, presented in Table 6. We found that girls earned higher grades than boys in ELA and science courses. There was also a tendency for girls to have higher GPAs in math and Spanish courses, which only reached significance among teacher-assistant models. The effect of lunch status was never significant. The effect of time was significant, indicating that ELA, math, and—for the teacher-assistant models—Spanish GPAs decreased over time. Finally, IEP status was significant in all models: Students with IEPs tended to have lower GPAs in all courses than students without IEPs.

The time residuals for all of the initial models were not significant, reflecting that little remains unsaid about the change over time in the GPAs; indeed there was little change in them. The lack of time effects—supported by the significant covariance residuals—indicates that the initial status of the predictor terms is related to later academic performance. This is not surprising since gender and poverty were both time-invariant.

Final models: adding BRIEF GEC and BRIEF GEC x time interactions to the initial models. Table 7 details changes in the pattern of significant effects that happened when teacher- or teacher assistant-derived BRIEF GEC scores and their interactions with time were both added to the models. The BRIEF terms were all highly significant—across courses and for both teachers and teacher assistants. EF, as measured by BRIEF scores, is a strong, consistent predictor of academic success. These strong effects outshined most of the other effects from the initial models.

The BRIEF GEC score x time interaction terms were significant in models predicting social studies and Spanish GPAs when BRIEF scores were generated by teacher assistants. For both of these courses, students with stronger EF (lower BRIEF GEC scores) tended to have GPAs that improve over these years.

All models—even those with significant BRIEF x time interactions—had strongly significant covariance residuals. One thing these significant covariance residuals indicate is that initial BRIEF scores well predict GPAs in later years for all courses and for both teacher- and teacher assistant-derived scores. In other words, knowing a student's BRIEF score—whether obtained from the student's teacher or teacher assistant—strongly predicted her/his GPA both for a broad range of both current and future courses, at least up to the ninth grade.

The reductions in deviance statistics from the initial models (the Δ*D*s in Table 7) are all highly significant (critical χ_3^2 = 11.34): The addition of BRIEF-related terms to the models greatly improved the fit of the models to the data. Nonetheless, even though the other terms lose their significance when BRIEF-related factors are added to the models, the other terms help model fit; the deviance statistics are significantly worse when those other terms are removed from the models with BRIEF-related factors (Δ*D*s ≈ −200). Even the removal of just IEP status from the final models significantly worsened their fits (Δ*D*s ≈ −150); even though BRIEF scores here eclipsed the effects of IEP status, not considering IEPs would noticeably weakens our ability to predict GPAs.

Effects of BRIEF GEC-Related Terms among Students With and Without IEPs

That last point about the importance of IEPs raises another consideration. It is possible that EF acts differently among students with and without IEPs. We tested this by dividing the data into those who had IEPs and those who did not. On each of these sub-sets of the data, we ran models that contained gender, lunch status, time, BRIEF GEC scores, and BRIEF x time interaction terms as the predictors and each of the course GPAs as outcomes. The models were run for both the teacher- and teacher-assistant-derived BRIEF scores.

There were very few differences between the sub-sets of students with and without IEPs. Looking at only those students who had IEPs, BRIEF GEC scores were still significant predictors of GPAs for all courses among both the teacher (smallest absolute β_{Spanish} = -.374, p = .0023) and teacher assistant (smallest absolute β_{ELA} = -.511, p = .0033) models. There was not sufficient data for the models to converge when analyzing either the teacher science model or the teacher assistant math, social studies, or Spanish models. The BRIEF x time effect was not significant among any of the models that converged (largest $\beta_{\text{Teacher social studies}}$ = .201, p = .0249).

BRIEF GEC scores were also significant predictors among those students who had not been diagnosed with the need for an IEP. For the teacher-derived data, the smallest absolute $\beta_{Science}$ = -.425, p < .0001; for the teacher assistant-derived data, the smallest absolute β_{ELA} = -.287, p = .0039.The model for Spanish GPAs among teacher assistant data did not converge. The BRIEF x time effect was only significant among the teacher assistant data when predicting ELA GPAs (β = .325, p = .0039) and Spanish GPAs (β = .297, p = .0078).

Another way to look at the relationship between BRIEF GEC scores and IEPs is to assess whether BRIEF scores themselves change over these years—and whether rates of change differ for those who have IEPs from those who do not. For both teacher- and teacher-assistant-derived BRIEF scores, the results were roughly the same: BRIEF scores were relatively stable and similar among students with and without IEPS. In these analyses, BRIEF scores were the outcomes in multilevel change models in which time and IEP status were the predictors. For the model on teacher-derived BRIEF scores, the effect of time did not quite achieve significance (β_{243} = 0.10, *S.E.* = 0.043, p = .018), the IEP status x time effect was not significant ($\beta_{243} \approx 0$, *S.E.* = 0.075, p = .99), and—not surprisingly—the IEP status "main effect" term was significant (β_{329} = 0.63, *S.E.* = 0.11, p< .0001). For the model on teacher-assistant-derived BRIEF scores, the time effect (β_{51} = 0.076, *S.E.* = 0.069, p = .27) and the IEP status x time effect were both not significant (β_{51} = −0.22, *S.E.* = 0.15, p = .14); the IEP status effect was (β_{199} = 0.49, *S.E.* = 0.15, p = .0012).

Discussion

EF offers a broad indicator of risk or resilience in early childhood (Masten et al., 2012); in the current study prediction of academic achievement in middle and early high school students was examined. The findings were (1) EF, as measured by BRIEF GEC scores, significantly adds to predictions of students' GPAs made by IEPs, gender, and poverty; and that (2) the development of EF over time leads to higher GPAs in some courses, but in general the effect of

EF on GPAs was quite stable over time. In fact, the effect was stable enough that initial assessments of EF remained predictive of GPAs to ninth grade. The addition of further waves of data and the inclusion of other relevant factors may well reveal changes in EF over time; they will, however, be unlikely to efface the robust relationships found here.

Students' rated EFs were strongly associated with their GPAs in all courses measured here: ELA, mathematics, social studies, science, and Spanish. There does not appear to be much difference between the courses in the magnitude of the effect of EF on GPA.

Not only was this effect of EF beyond that made by IEP, gender, and poverty, but it was considerably stronger than that of the other terms. These results support those found by Waber et al. (2006), Bierman et al. (2009), and Kotsopoulos and Lee (2012) that EF has a powerful effect on disadvantaged, adolescent students' academic performance.

Girls and those students without IEPs tended to have higher GPAs. Lunch status—a measure of family poverty—was unrelated to GPAs. In addition, even though the effect of BRIEF GEC scores eclipsed most other effects, the inclusion of IEP status at least was important for overall model fit. IEP status and BRIEF scores were highly correlated, but further analyses indicate that they may act through different mechanisms.

Investigating EF in the field over several years adds further insights into the role of EF in academic performance, and not all of these insights correspond to the findings of others. GPAs among these students tended to decrease over the years, but changes in GPAs were insignificant when BRIEF GEC-related factors were taken into consideration (i.e., the time term was not significant when BRIEF-related terms were added). In addition, any changes in BRIEF scores were only associated with changes the GPAs of a few courses—not surprising given that neither BRIEF scores nor GPAs changed greatly. This helped initial EF ratings to be strong predictors of GPAs in later grades, both when the assessments are made by teachers or teacher assistants. The relationship between EF and academic success is consistent with other research (e.g., Bierman et al., 2009; Kotsopoulos & Lee, 2012; Latzman, Elkovitch, Young, & Clark, 2010; Waber et al., 2006), but our inability to detect reliable changes in EF over several adolescent years is not. In a very large, field-based study, Best, Miller, and Naglieri (2011) found increases in EF continued at least into middle adolescents. Environmental stressors such as poverty and adversity can impede its development (Blair & Raver, 2012; Buckner, Mezzacappa, & Beardslee, 2003; 2009; Shonkoff, 2011) and may account for at least some of the negative findings here.

Teachers' ratings of students' EF were highly correlated with teacher assistants' EF ratings. The patterns of which terms were significant were also quite similar. Nonetheless, when differences did occur, it was that models incorporating teacher assistant ratings found a greater number of significant terms, viz., gender, time-related terms, and—once—IEP status. The fits (e.g., the deviance statistics, $-2LL$s) of teacher and teacher assistant models were similar, so the models that incorporated teacher assistant ratings do not appear to be better models. Although in general, then, teachers and teacher assistants are roughly equally useful sources of ratings of students' EF-related behaviors, ratings given by teacher assistants may be slightly more independent of (and thus possibly less biased by) students' genders and academic performance. The differences in the effects between teacher and teacher assistant ratings, however, were quite small.

Whether or not a student had an IEP was related to the student's GPA in all subjects measured here. In addition, even though IEP status was related to gender and BRIEF GEC scores, including IEPs in predictions of GPAs significantly improved the fits of both the initial and final models. IEP status and lunch status, a measure of family poverty, were not found to be related among these students. There are many reasons that students are diagnosed with benefitting from an IEP—not all of which pertain to intellectual needs. In addition, students with disabilities may have never been diagnosed and thus not given IEPs. The effect of IEP status here, then, represents a general but robust effect of the disabilities that typically comprise IEPs; more advanced considerations of their role will require further, more precise investigation.

Lunch status was never found to be significantly related to GPAs, a result we did not expect to find. Most (~80%) of these students lived in poverty, and the lunch status scale also has a very limited range of values; both of these limit the variability of the lunch status term. Given the strong effect that poverty is known to have on academic success (e.g., Haveman & Wolfe, 1995), the lack of an effect here may further support the need for more precise measurements of poverty and the factors that precipitated these students' IEPs.

The effect of time, however, did not suffer from a lack of precision. Knowing the students' birth dates and the dates on which the teachers and teacher assistants completed the BRIEF allowed us to measure time to the day. Its effect on grades was nonetheless never omnipresent and was finally eclipsed by the presence of BRIEF-related factors. The persistently weaker effect of time may simply indicate that students' learning kept pace with their grade level—i.e., that teachers in the various grade levels all tended to give similar grades. It also argues that EF can be used as a long-term predictor of academic success throughout middle school and into 9th grade.

Recommendations for Policy and Practice

Even though further research is needed to elucidate several of the areas explored through this study, the current findings have important educational significance. Although the BRIEF is long, it is relatively easy to administer, and both teachers and teacher assistants produce useful, informative ratings that qualified professionals can be trained to compute and interpret. Therefore, BRIEF GEC scores represent a potentially cost-effective diagnostic tool that can be used to prioritize both short- and long-term academic and behavioral interventions among disadvantaged, middle school and early high school students.

There were some differences in models using teacher ratings versus teacher assistant ratings; the latter may be slightly more independent of gender, time, and possibly IEP status. If an administrator were to choose only one of these two groups to complete the BRIEF, then, the latter may be preferred. Any differences may come from these teacher assistants being able to work one-on-one with more students more often, so a safe recommendation would be that the BRIEF should be completed by professionals who are very familiar with the students.

In addition, the exceptionally strong relationship between grades and BRIEF GEC scores is stable enough that an initial diagnosis made in early middle school grades well predicts academic performance at least into the first year of high school. Finding that the relationships between BRIEF GEC scores and GPAs changed little over these four years further underlines their use to make longer-term decisions about interventions among adolescents. EF ratings offer an additional, highly-predictive piece of information about students and their needs.

More generally, the strong relationship between EF and GPAs suggests the development of EF-related behaviors and capacities are important components of academic success across many disciplines. EF proved rather stable across the time ages studied here; nonetheless, students would likely strongly benefit from efforts to nurture the growth of their EF. Kaufman (2010) as well as Dawson and Guare (2010) and others offer recommendations for ways to address EF in the classroom

In addition, the BRIEF used as an early diagnostic tool appears to provide useful, cost-effective, *complimentary* information to that contained within most IEPs. Administrators, special education coordinators, counselors, teachers, and teacher assistants can use EF ratings along with any IEP directives to plan both short- and long-term supports.

Given the stability of BRIEF scores over these years, the BRIEF need not to be administered every year in middle school. For students with disabilities, for example, it seems appropriate to follow the precedent and infrastructure that exists in special education and administers the BRIEF on the tri-annual cycle on which standardized tests are scheduled to be administered. Then, the Committee on Special Education can use this information when reviewing and revising IEPs. Schools may choose to follow the same timeline for students without disabilities. EF was an important predictor of GPAs among all of these students. The scores may help promote broader conversations about interventions than do IEPs meetings, and assessing all students may help reduce perceived differences between students with and without IEPs.

Of course, GEC scores are still general measures. EF is a broad and not always consistently defined construct; different measures of it produce different relationships (e.g., Best, Miller, & Naglieri, 2011; Bull & Scerif, 2001). Interventions that address it are equally diverse (Riccio & Gomes, 2013), and more effective interventions may arise from targeting specific aspects of EF (Cantin, Mann, & Hund, 2012; Dawson & Guare, 2010), even though commonalities exist (e.g., many include stable, warm relationships with adults that promote self-regulation and stress management; Liew, Chen, & Hughes, 2010). BRIEF GEC scores, therefore, may be best used as a component of an initial battery administered earlier during a student's adolescence to make general, long-term decisions; these initial assessments can then be tailored if and when additional information and resources become available.

Limitations

In addition to the somewhat limited range of time available for analysis, all students haled from the same school. Admittance to the school is by lottery among the many applicants, so they at least initially represent well a population of diverse students from disadvantaged backgrounds; in their first year at the school, students are a more truly random selection of the general population than is achieved in most studies. However, the development EF among the students and their continued enrollment in the school may not be independent; school-based factors may affect their development and thus reduce the generalizability to other settings. The mission of the school itself centers on providing disadvantaged students, especially those with special needs, with an enriched and supportive academic environment; it is therefore possible that the developmental trajectories of EF among these students would differ from peers on other education settings, e.g., that the rates would be accelerated at the current school.

The IEP statuses of students who left the school during the first two academic years that the school was open are not known, nor are the reasons these students left. Anecdotal reports from the school's administrators indicate that familial changes (e.g., moving out of the area) as well as poor academic performance likely contributed to most of the reasons for leaving. Those who left for academic reasons would bias our sample, perhaps weakening the effect of IEPs on academics. In addition, among the students who did not have IEPs, there is no information on how many were evaluated but found not to need an IEP or how many were simply never evaluated. Finally, the IEPs that students had could be for non-intellective reasons such as socio-emotional or physical, possibly weakening the strength of the IEP effect on GPAs.

References

Alloway, T. P., Gathercole, S. E., Adams, A.-M., & Willis, C. (2005). Working memory abilities in children with special educational needs. *Educational and Child Psychology*, *22*(4), 56 – 67.

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

Anderson, V., Anderson, P. J., Jacobs, R., & Spencer Smith, M. (2008). Development and assessment of executive function: From preschool to adolescence. In V. Anderson, R. Jacobs, & P. J. Anderson (Eds.), *Executive Function and the Frontal Lobes: A Lifespan Perspective* (pp.123 – 154). New York, NY: Psychology Press.

Baughman, F. D. & Cooper, R. P. (2007). Inhibition and young children's performance on the Tower of London task. *Cognitive Systems Research*, *8*, 216 – 226.

Best, J. R. & Miller, P. H. (2010). A Developmental perspective on executive function. *Child Development*, *81*, 1641 – 1660. doi:10.1111/j.1467-8624.2010.01499.x.

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*, 327 – 336.

Bierman, K. L., Torres, M. M., Domitrovich, C. E., Welsh, J. A., & Gest, S. D. (2009). Behavioral and cognitive readiness for school: Cross-domain associations for children attending Head Start. *Social Development*, *18*, 305 – 323.

Blair, C. (2002). School readiness: Integrating cognition and *emotion* in a neurobiological conceptualization of children's functioning at school entry. *American Psychology, 57*(2), 111-127.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences, 29*(2), 109 – 125.

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

Blair, C., & Razza, R. P. (2007).Relating effortful control, executive functioning, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78,* 647 – 663.

Buckner, J. C., Mezzacappa, E., & Beardslee, W. R. (2003). Characteristics of resilient youths living in poverty: The role of self-regulatory processes. *Development and Psychopathology, 15*, 139 – 162.

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*, 19-30.

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: Sources of variation in emergent mathematic achievement. *Developmental Science, 14*, 679 – 692. doi:10.1111/j.1467-7687.2010.01012.x.

Bull, R., Johnson, R. S., & Roy, J. A. (1999). Exploring the roles of the visuo-spatial sketchpad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15*, 421 – 442

Bull, R. & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*, 273 – 293.

Cantin, R. H., Mann, T. D., & Hund, A. M. (2012). Executive functioning predicts school readiness and success: Implications for assessment and intervention. *Communique*, *41*(4), 1.

Cartwright, K. B. (2012). Insights from cognitive neuroscience: The importance of executive function for early reading development and education. *Early Education and Development*, *23*, 24 – 36.

Checa, P. & Rueda, M. R. (2011). Behavioral and brain measures of executive attention and school competence in late childhood. *Developmental Neuropsychology*, *36*, 1018 – 1032.

Clark, C., Prior, M., & Kinsella, G. (2002). The relationship between executive function abilities, adaptive behaviour, and academic achievement in children with externalizing behaviour problems. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *43*, 785 – 796.

Clark, C. A. C., Pritchard, V. E., & Woodward, L. Z. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology*, *46*, 1176-1191.

Davidson, M.C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037 – 2078.

Davis, C. L., Tomporowski, P. D., McDowell, J. E., Austin, B. P., Miller, P. H., Yanasak, N. E., Allison, J. D., & Naglieri, J. A. (2011). Exercise improves executive function and achievement and alters brain activation in overweight children: A randomized, controlled trial. *Health Psychology*, *30*, 91 – 98.

Dawson, P. & Guare, R. (2010). *Executive Skills in Children and Adolescents: A Practical Guide to Assessment and Intervention* (2nd Ed.). New York, NY: Guilford Press.

Diamantopoulou, S., Rydell, A.-M., Thorell, L. B. & Bohlin, G. (2007). Impact of executive functioning and symptoms of attention deficit hyperactivity disorder on children's peer relations and school performance. *Developmental Neuropsychology*, *32*(1), 521 – 542.

Diamond, A. & Lee, K. (2011).Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*, 959-964.

Dilworth-Bart, J. E. (2012). Does executive function mediate SES and home quality associations with academic readiness? *Early Childhood Research Quarterly*, *27*, 416 – 425.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428 – 1446. doi:10.1037/0012-1649.43.6.1428

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamy, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*, 465 – 486.

Everitt, B. S. & Skrondal, A. (2010). *The Cambridge Dictionary of Statistics* (4th Ed.) Cambridge: Cambridge University Press.

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*, 1410 – 1430.

Garner, J. K. (2009). Conceptualizing the relations between executive functions and self-regulated learning. *Journal of Psychology*, *143*, 405 – 426.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review and integrative framework. *Psychological Bulletin*, *134*, 31 – 60.

Geiser, S. & Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *CSHE Research & Occasional Paper Series*, CSHE.6.07., 1 – 35.

Ginsburg, H. P. (1997). Mathematics learning disabilities: A view from developmental psychology. *Journal of Learning Disabilities*, *30*, 20 – 31.

Gioia, G. A., Andrews Epsy, K., & Isquith, P. K. (2003). *BRIEF-P. Behavior rating inventory of executive function—Preschool version. Professional manual.* Lutz, FL: Psychological Assessment Resources, Inc.

Gioia, G. A. & Isquith, P. K. (2004).Ecological assessment of executive function in traumatic brain injury. *Developmental Neuropsychology*, *25*, 135 – 158.

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

Gioia, G. A., Isquith, P. K., Kenworthy, L., & Baron, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, *8*, 121 – 137.

Hartman, E., Houwen, S., Scherder, E. & Visscher, C. (2010). On the relationship between motor performance and executive functioning in children with intellectual disabilities. *Journal of Intellectual Disability Research: JIDR*, *54*(5), 468 – 477.

Haveman, R. & Wolfe, B. (1995). The determinants of children’s attainments: A review of methods and findings. *Journal of Economic Literature*, *33*(3), 1829 – 1878.

Hintermair, M. (2013). Executive functions and behavioral problems in deaf and hard-of-hearing students at general and special schools. *Journal of Deaf Studies & Deaf Education*, *18*(3), 344 – 359.

Hostinar, C. E., Stellern, S. A., Schaefer, C., Carlson, S. M., & Gunnar, M. R. (2012). *Associations between early life adversity and executive function in children adopted internationally from orphanages.* PNAS Proceedings of the National Academy of Sciences of the United States of America, 109, 17 208 – 17 212.

Hughes, C. & Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. *Journal of Experimental Child Psychology*, *108*, 663 – 676.

Huizinga, M., Dolan, C., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017 – 2036.

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*, 193 – 215.

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 – 132.

Isquith, P. K., Gioia, G. A., & Espy, K. (2004). Executive function in preschool children: Examination through everyday behavior. *Developmental Neuropsychology*, *26*, 403 – 422.

Jansen, B. R. J., De Lange, E. & Van der Molen, M. J. (2013). Math practice and its influence on math skills and executive functions in adolescents with mild to borderline intellectual disability. *Research in Developmental Disabilities: A Multidisciplinary Journal*, *34*(5), 1815 – 1824.

Kaufman, C. (2010). *Executive Functioning in the Classroom: Practical Strategies for Improving Performance and Enhancing Skills for All Students*. Baltimore, MD: Paul H. Brookes Publishing Co.

Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Frontal lobes: Cognitive neuropsychological aspects. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioural Neurology and Neuropsychology* (pp. 409 – 418). New York, NY: McGraw Hill.

Kotsopoulos, D. & Lee, J. (2012).A naturalistic study of executive function and mathematical problem-solving. *The Journal of Mathematical Behavior*, *31*, 196 – 208.

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*, 677 – 692.

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology*, *32*, 455 – 462.

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*, 373 – 387. doi: 10.1037/a0013843.

Lehto, J. (1996). Are executive function tests dependent upon working memory capacity? *Quarterly Journal of Experimental Psychology*, *49A*, 29 – 50

Lehto, J. E., Juujarvi, 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

Liew, J. (2012). Effortful control, executive functions, and education: Bringing self-regulatory and social-emotional competencies to the table. *Child Development Perspectives*, *6*, 105 – 111.

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.

Lohmeier, J. H. & Raad, J. (2012). The predictive value of selection criteria in an urban magnet school. *Urban Education*, *47*(3), 612-624. doi: 10.1177/0042085911434569.

Lorsbach, T. C., Wilson, S., & Reimer, J. F. (1996). Memory for relevant and irrelevant information: Evidence for deficient inhibitory processes in language/learning disabled children. *Contemporary Educational Psychology*, *21*, 447 – 466.

Luciana, M., Conklin, H. M., Hooper, C., J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, *76*, 697 – 712.

Lyons, K. E. & Zelazo, P. D. (2011). Monitoring, metacognition, and executive function: Elucidating the role of self-reflection in the development of self-regulation. In B. B. Janette (Ed.), *Advances in Child Development and Behavior* (Vol. 40, pp. 379 – 412). Oxford: Elsevier.

Lyons, G. R., Fletcher, J. M., & Barnes, M. C. (2003). Learning disabilities. In E. J. Mash & R. A. Barkley (Eds.), *Child Psychopathology* (pp. 520 – 586). New York, NY: Guilford Press.

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., et al. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*, 643 – 662.

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., Sapienza, J. K., Long, J. D., & Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. *Educational Researcher*, *41*, 375 – 384.

McLean, J. F. & Hitch, J. (1999).Working memory impairments in children with specific arithmetical learning difficulties. *Journal of Experimental Child Psychology*, *74*, 240 – 260.

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*, 49 – 100.

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*, 158 – 173. doi:10.1016/j.jecp.2011.01.008.

Pennington, B. F. (1998). Dimensions of executive funcitons in normal and abnormal development. In N. A. Krasnegor, G. R. Lyon, & P. S. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior.* (pp. 327-348). Baltimore: Brookes.

Pinheiro, J., Bates, D., Deb Roy, S. Sarkar, D., and the R Development Core Team (2013). *nlme: Linear and Nonlinear Mixed Effects Models.* R package version 3.1 – 113.

R Core Team (2013).*R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Retrieved from http://www.R-project.org/.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The noword reading deficit in Developmental dyslexia: A review. *Reading Research Quarterly*, *27*, 29 – 53.

Revelle, W. (2014) *psych: Procedures for Personality and Psychological Research.* Evanston, Illinois Northwestern University. Retrieved from http://cran.r-project.org/web/packages/psych/

Roebers, C. M., Cimeli, P. Rothlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary schoolchildren: An explorative study of their interrelations and their role for school achievement. *Metacognition Learning*, *7*, 151 – 173.

Romine, C.B. & Reynolds, C.R. (2005). A model of the development of frontal lobe function: Findings from a meta-analysis. *Applied Neuropsychology*, *12*, 190 – 201.

Rourke, B. P. & Conway, J. A. (1997). Disabilities of arithmetic and mathematical reasoning. *Journal of Learning Disabilities*, *30*, 34 – 46.

Rosenthal, M., Wallace, G. L., Lawson, R., Wills, M. C., Dixon, E., Yerys, B. E., & Kenworthy, L. (2013). Impairments in real-world executive function increase from childhood to adolescence in autism spectrum disorders. *Neuropsychology*, *27*, 13 – 18.

Russell, J., Jarrold, C., & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry*, *37*, 673 – 686.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*(6), 110 – 114, doi: 10.2307/3002019.

Savage, R., Pillay, V., & Melidona, S. (2007). Deconstructing rapid automatized naming: Component processes and the prediction of reading difficulties. *Learning and Individual Differences*, *17*, 129 – 146.

Sbordone, R. J. (1996). Ecological validity: Some critical issues for the neuropsychologist. In R. J. Sbordone & C. J. Long (Eds.), *Ecological Validity of Neuropsychological Testing* (pp. 15 – 42). Boca Raton, FL: St. Lucie Press.

Schuchardt, K., Bockmann, A.-K., Bornemann, G., & Maehler, C. (2013). Working memory functioning in children with learning disorders and specific language impairment. *Topics in Language Disorders*, *33*(4), 298 – 312.

Semrud-Clikeman, M., Fine, J. G. & Bledsoe, J. (2014). Comparison among children with children with autism spectrum disorder, nonverbal learning disorder and typically developing children on measures of executive functioning. *Journal of Autism and Developmental Disorders*, *44*(2), 331 – 342.

Shonkoff, J. P. (2011). Protecting brains, not simply stimulating minds. *Science*, *333*, 982 – 983.

Silver, C. H. (2000). Ecological validity of neuropsychological assessment in childhood traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *15*, 973 – 988.

Singer, J. D. & Willett, J. B. (2004). *Applied Longitudinal Data Analysis*. New York: Oxford University Press.

Somsen, R. J. M. (2007). The development of attention regulation in the Wisconsin Card Sorting Task. *Developmental Science*, *10*, 664 – 680.

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

Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology*, *56*, 87 – 114.

Swanson, H. L. (1999). Reading comprehension and working memory in learning-disabled readers: Is the phonological loop more important than the executive system? *Journal of Experimental Child Psychology*, *72*, 1 – 31.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*, 427 – 449.

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*, 100 – 119. doi:10.1111/j.2044-8279.2011.02035.x.

Waber, D. P., Gerber, E. B., Turcios, V. Y., Wagner, E. R., & Forbes, P. W. (2006). Executive functions and performance on high-stakes testing in children from urban schools. *Developmental Neuropsychology*, *29*, 459 – 477.

Welch, B. L. (1947). The generalization of "student's" problem when several different population variances are involved. *Biometrika*, *34*, 28 – 35

Welsh, M. C. (1991). Rule-guided behavior and self-monitoring on the Tower of Hanoi disk transfer task. *Cognitive Development, 6*, 59 – 76.

Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology*, *85*(2), 357 – 364.

Table 1

Number of Students by Grade Level and Academic Year.

	Academic Year			
Grade Level	2009-2010	2010-2011	2011-2012	2012-2013
6	64	72	83	107
7		73	92	92
8			48	70
9				19
Total Enrollment	64	145	223	288

Note. The shaded arrows reflect that the school grew a grade per year as, e.g., the 6th grade students in AY 2009-2010 comprise most of the 7th grade students in AY 2010-2011.)

Table 2

Means and Standard Deviations of Executive Functioning Scores, Age, and Annual Grade Point Averages.

		Grade 6		Grade 7		Grade 8		Grade 9	
		Male	Female	Male	Female	Male	Female	Male	Female
Teacher BRIEF GEC Score	Mean	126.2	106.8	133.2	114.9	164.8	146.6	126.1	138.1
	S.D.	33.8	29.2	33.3	32.1	31.1	46.5	37.7	29.9
Teacher Assistant BRIEF GEC Score	Mean	117.2	98.0	125.7	107.6	136.6	108.4	127.6	111.1
	S.D.	30.3	26.5	36.2	32.1	35.4	30.7	35.2	35.3
Student Age	Mean	12.2	12.2	13.1	13.1	14.0	14.0	15.0	15.0
	S.D.	0.3	0.3	0.4	0.2	0.4	0.3	0.2	0.1
ELA GPA	Mean	4.8	5.7	4.8	5.7	2.9	3.4	3.5	5.4
	S.D.	2.9	2.6	2.4	2.3	2.4	2.2	3.2	2.7
Math GPA	Mean	4.2	4.4	4.1	4.7	3.3	2.6	3.6	4.6
	S.D.	2.9	2.9	2.6	2.7	2.9	2.1	3.1	2.0
Science GPA	Mean	3.8	4.3	3.7	4.5	5.1	5.5	3.6	4.1
	S.D.	2.9	2.8	2.7	2.9	2.7	1.7	2.5	3.1
Social Studies GPA	Mean	3.8	4.3	4.1	4.3	4.2	2.9	4.0	4.4
	S.D.	2.8	2.9	2.1	2.4	2.8	2.1	2.0	2.3
Spanish GPA	Mean	4.7	5.7	4.4	4.5	4.8	4.3	3.9	4.9
	S.D.	3.1	2.9	2.5	2.5	2.8	2.7	3.5	3.0

Table 3

Descriptive Statistics of BRIEF Scores and Course GPAs by IEP Status.

			Student Has an IEP	Student Does Not Have an IEP
BRIEF Score	Teacher	Mean	132.3	110.63
		S.D.	34.3	29.78
		n	78	253
	Teacher Assistant	Mean	127.17	107.8
		S.D.	29.82	34.25
		n	44	157
Mean GPA	ELA	Mean	3.5	5.82

	S.D.	2.06	2.2
	n	44	132
Math	Mean	3.13	5.08
	S.D.	2.25	2.45
	n	44	132
Science	Mean	2.81	4.8
	S.D.	2.47	2.51
	n	44	131
Social Studies	Mean	2.89	4.73
	S.D.	2.37	2.46
	n	44	131
Spanish	Mean	3.4	5.41
	S.D.	2.31	2.46
	n	44	131

Table 4

Number and Percent of Students With and Without IEPS by Gender.

	Has an IEP		Does Not Have an IEP		Total	
	n	%	*n*	%	*n*	%
Female	28	8.7	121	37.6	149	46.3
Male	51	15.8	122	37.9	173	53.7
Total	79	24.5	243	75.5	322	100

Table 5

Number and Percent of Students Eligible for Free or Reduced School Lunches by IEP Status.

	Student Has an IEP		Student Does Not Have an IEP		Total	
	n	%	*n*	%	*n*	%
Free Lunch	50	17.4	146	50.7	196	68.1
Reduced-Price Lunch	10	3.5	32	11.1	42	14.6
Full Price Lunch	18	6.3	32	11.1	50	17.4
Total	78	27.1	210	72.9	288	100

Table 6

Initial Multilevel Change Models Including Gender, Lunch Status, Special Education Status, and Time as Predictors of Course GPAs.

		ELA		Math		Science		Social Studies		Spanish	
		Teacher	Teachin g Assist.	Teacher	Teaching Assist.	Teacher	Teaching Assist.	Teacher	Teaching Assist.	Teacher	Teaching Assist.
Gender (Male = 0, Female = 1)	*b*	0.39 *	0.41 *	0.21	0.24 *	0.30 *	0.30 *	0.20	0.21	0.22	0.23 *
	SE	0.08	0.08	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
Lunch Status (Free = −1, Reduced = 0, Paid = 1)	*b*	−0.03	−0.05	−0.04	−0.05	0.08	0.08	−0.01	−0.02	−0.03	−0.04
	SE	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07
Time (Change in GPAs over Time)	β	−0.24 *	−0.21 *	−0.24 *	−0.18 *	−0.02	−0.02	−0.13	−0.12	−0.19	−0.18 *
	SE	0.07	0.06	0.07	0.06	0.07	0.06	0.08	0.06	0.07	0.06
IEP Status (Has IEP = 1, No Diagnosis = 0)	*b*	−0.65 *	−0.63 *	−0.51 *	−0.49 *	−0.55 *	−0.55 *	−0.54 *	−0.53 *	−0.45 *	−0.43 *
	SE	0.10	0.10	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
Covariance Residual	β	0.69 *	0.69 *	0.78 *	0.79 *	0.82 *	0.82 *	0.90 *	0.89 *	0.82 *	0.81 *
	SE	0.06	0.06	0.07	0.07	0.07	0.07	0.08	0.08	0.07	0.07
Time Variance	β	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Deviance Statistic (−2*LL*)		598.20	596.12	628.08	630.15	639.88	639.87	660.32	659.36	638.43	636.13

* $p < .01$

Table 7

Multilevel Change Models with BRIEF Scores and BRIEF Scores x Time Interaction Terms Added.

		ELA		Math		Science		Social Studies		Spanish	
		Teacher	Teaching Assist.	Teacher	Teaching Assist.	Teacher	Teaching Assist.	Teacher	Teaching Assist.	Teacher	Teaching Assist.
Gender (Male = 0, Female = 1)	*b*	0.10	0.24	−0.06	−0.06	0.06	−0.02	−0.13	−0.10	−0.05	−0.12
	SE	0.08	0.10	0.08	0.10	0.09	0.10	0.09	0.10	0.09	0.10
Lunch Status (Free = −1, Reduced = 0, Paid = 1)	*b*	−0.04	−0.09	0.00	−0.14	0.13	0.07	−0.02	−0.04	−0.03	−0.09
	SE	0.06	0.07	0.06	0.07	0.06	0.07	0.06	0.07	0.07	0.07
Time (Change in GPAs over Time)	β	−0.14	−0.20 *	−0.08	−0.16	0.08	−0.03	−0.03	−0.04	−0.09	−0.10
	SE	0.07	0.06	0.07	0.06	0.07	0.06	0.07	0.06	0.08	0.06
IEP Status (Has IEP = 1, No Diagnosis = 0)	*b*	−0.24	−0.40 *	−0.07	−0.15	−0.16	−0.19	−0.01	−0.15	−0.11	−0.13
	SE	0.11	0.12	0.11	0.13	0.11	0.12	0.11	0.13	0.12	0.12
BRIEF GEC Score	β	−0.47 *	−0.33 *	−0.53 *	−0.47 *	−0.50	−0.59 *	−0.56 *	−0.52 *	−0.41 *	−0.43 *
	SE	0.06	0.08	0.06	0.08	0.06	0.08	0.06	0.08	0.07	0.08
BRIEF GEC Score x Time (Effect of EF on Rates of Change in GPAs)	β	0.11	0.18	0.08	0.17	0.03	0.02	0.15	0.24 *	0.14	0.27 *
	SE	0.06	0.08	0.06	0.09	0.06	0.08	0.06	0.09	0.06	0.08
Covariance Residual	β	0.52 *	0.60 *	0.53 *	0.66 *	0.58 *	0.65 *	0.57 *	0.71 *	0.66 *	0.64 *
	SE	0.07	0.06	0.05	0.07	0.06	0.07	0.06	0.07	0.07	0.07
Time Variance	β	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

SE	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Deviance Statistic ($-2LL$)	452.77	430.85	454.85	449.21	473.85	443.9	470.8	461.36	501.85	443.27
Change in Deviance Statistic from Initial Model (ΔD)	145.43*	165.27*	173.23*	180.94*	166.03*	195.97*	189.52*	198.00*	136.58*	192.86*

* $p < .01$

Figure 1: Summary of changes in significant and highly significant terms across the multilevel change models.

	Teacher										Teacher Assistant									
	Lunch, Gender, & IEP Status					Lunch, Gender, IEP Status, & BRIEF Score					Lunch, Gender, & IEP Status					Lunch, Gender, IEP Status, & BRIEF Score				
	ELA	Math	Science	Social Studies	Spanish	ELA	Math	Science	Social Studies	Spanish	ELA	Math	Science	Social Studies	Spanish	ELA	Math	Science	Social Studies	Spanish
Gender	‡		‡								‡	•	‡		•					
Lunch Status																				
Time	‡	‡									‡	•			•	‡				
IEP Status	‡	‡	‡	‡	‡						‡	‡	‡	‡	‡	‡				
BRIEF Score						‡	‡	‡	‡	‡						‡	‡	‡	‡	‡
BRIEF Score x Time																			•	•

• p < .01; ‡ p < .001