BRIEF and BRIEF-SR Scores as Predictors of Academic Success in a Middle School: Initial Findings

William Ellery Samuels

College of Staten Island/City University of New York

Sheldon Blackman

Lavelle Preparatory Charter School

Nelly Tournaki

College of Staten Island/City University of New York

Christopher Ziliniski

Lavelle Preparatory Charter School

Paper Presented at the 2014 Annual Meeting of the Eastern Psychological Association

Boston, Massachusetts
Executive functioning (EF) is a strong predictor of many aspects of children’s and adolescents’ developmental success, including academic performance. The goals of the project were twofold: (1) examining the relationship between EF as measured by BRIEF and BRIEF-SR global scores and academic performance and (2) assessing these relationships throughout the middle score years. BRIEF scores were given by teachers’ assessment of students and student self-reports for 404 adolescents from disadvantaged backgrounds who attended grades 6 – 9 at an urban, chartered middle school; those same students provided BRIEF-SR scores. The findings indicate that (1) both teacher-reported BRIEF and self-reported BRIEF-SR scores were strongly associated with both science/math and ELA/social studies/Spanish grades, as well as the school’s “wellness” courses which address academic preparedness and life skills; and that (2) BRIEF and BRIEF-SR scores changed little over the measured grades, nor did the relationship between these scores and grades itself change. Therefore, BRIEF and BRIEF-SR were found to be a relatively stable trait that strongly affected middle school students’ academic performance. We argue, therefore, that these scores obtained during early middle grades can be used to predict academic performance in subsequent grades.
Executive Functioning as a Predictor of Academic Success in a Middle School: Initial Findings

Executive functioning (EF) is defined as a “collection of processes that are responsible for guiding, directing, and managing cognitive, emotional, and behavioral functions, particularly during active, novel problem solving” (Gioia, Isquith, Guy, & Kenworthy, 2000, p. 1). More specifically, in schools, students are expected to perform complex cognitive activities and also exhibit overt goal-directed behaviors; when successful, we refer to them as exhibiting high EF (Anderson, 2002; Isquith, Gioia, & Espy, 2004; Lezak, 1983; Welsh, & Pennington, 1988). Although there is general consensus that EF represents a unifying latent 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, attentional set shifting (Huizinga, Dolan, & van der Molan, 2006; Miyake et al., 2000; Welsh, 1991; van der Sluis, de Jong, & van der Leij, 2007) as well as complex planning and meta-cognitive tasks (Baughman & Cooper, 2007; Miyake et al., 2000). In education we are looking for the identification of early, modifiable predictors of achievement (e.g., Ginsburg, 1997) 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). Such identification is particularly relevant in improving the success of educational strategies aimed at assisting children who struggle academically (Rack, Snowling, & Olson, 1992; Savage, Pillay, & Melidona, 2007). We assume that such a predictor is EF.

EF and IQ

Related to effortful-control (Liew, 2012) 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 et al., 2010). More specifically, studies have suggested that EF shows a distinct association with concurrent mathematics performance, in addition to measures assessing children’s general cognitive ability (Bull, & Scerif, 2001; Espy, et al., 2004; St. Clair-Thompson, & Gathercole, 2006). Moreover, each measure of EF (including the Stroop task, Wisconsin Naming Task, and counting span) predicted unique variance in mathematical ability beyond that ac-
Executive Functioning & Academic Success

counted for by IQ or reading achievement. Further, a recent longitudinal study predicted early mathematics achievement of six year old children based on their EF when they were four years old (Clark, Pritchard, & Woodward, 2010). 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 sub-test.

The Development of EF over Time

As explained in a review by Best and Miller (2010), EF is observable and measurable early in human development and continues to improve into adolescence. More specifically, 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, by switching between task demands while updating the tasks’ goals and specifics. During elementary school years, further progression in EF performance is typically found, mirrored by the emergence of a speed-accuracy trade-off that corresponds to children’s growing awareness of a discrepancy between tasks demands, on the one hand, and their own performance, on the other. More specifically, 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 toddler’s self restraint and found that individual differences predict their EF 14 years later.

In the present study we examined EF of middle school students; there is evidence that EF perhaps increases more gradually during adolescence (Best, & Miller, 2010; Davidson et al., 2006; Huizinga & van der Molen, 2007). For example, investigating a large sample of 2036 children aged five 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). In conclusion, EF is considered a key indicator of risk or resilience for school outcomes.

The goals of the study were therefore to (1) assess the growth of EF during adolescence among at-risk children, (2) investigate the extent and strength of the relationship between EF and various academic topics, (3) determine whether students’ self-reports or teachers’ ratings of students were more predictive of academic performance, and (4) make recommendations for school-based professionals about the use of EF-related scores in planning academic interventions.

Methods

Sample

Participants included all male and female students, aged 12 to 15 years, attending a charter school in grades 6 through 9 and the 21 teachers who taught each grades’ wellness classes; the number of students participating varied by year, increasing as those in the initial sixth grade graduated and additional students were added as grade levels were added; Table 1 summarizes the number of participants each year. The students tended to live in disadvantaged neighborhoods, with 87% of the students receiving free (70%) or reduced-price (17%) school lunches. Approximately 42% of the students are African American, 17% are European American, 8% are Asian American, and 32% identify with a Latin culture.

Setting

The study was conducted at a single, urban charter school in the northeastern US. The school’s charter dictates that at least 40% of the student body have IEPs; the diagnoses can be physical, psychological, and/or behavioral. The mission of the school is to provide “a rigorous college preparatory education that equips and empowers . . . all students, including those living with emotional challenges.” Classrooms 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.”

The school opened in the 2009-2010 academic year with a single sixth grade class. The school added a grade each year as this cohort progressed; grades 6 – 12 will ultimately exist. When data were analyzed here, grades included 6 through 9.

Materials

BRIEF and BRIEF-SR

There are two ways through which EF is assessed: performance-based tests and rating scales. Historically, the assessment of EF in clinical and research settings relied on the former but a number of authors have argued that such measures have limited ecological validity—that is, there is little relationship between a neuropsychological test and the participant’s behavior in the real-world settings (Isquith, Roth, & Gioia, 2013; Sbordone, 1996). On the other hand, several authors have argued that rating scales, by their very nature, capture observations of those every day, real-world behaviors (Gioia & Isquith, 2004; Silver, 2000).

In the present study, we chose to use one of the rating scales, namely the Behavior Rating Inventory of Executive Function (BRIEF, Gioia, Andrews, Espy, & Isquith, 2003). The BRIEF is a well-researched and established instrument published by Psychological Assessment Resources, Inc. (PAR). The right to use the BRIEF and BRIEF-SR was obtained from PAR. The BRIEF is comprised of 86 items through which a long-time observer assesses a subject’s executive functioning. Items ask such things as how often the students “needs help from an adult to stay on task,” “becomes overwhelmed by large assignments,” “makes careless errors,” “does not take initiative,” and “interrupts others.” The BRIEF was completed by teachers and paraprofessionals who know the students well.

The BRIEF-SR, also published by and obtained from PAR, is the companion instrument to the BRIEF that measures self-reported responses to items pertaining to executive functioning. The
80 items cover topics similar to those on the BRIEF: “I don’t plan ahead for school assignments,” “I make careless errors,” “does not take initiative,” and “I interrupt others.”

Although both BRIEF instruments produce several 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

Academic performance was measured through students’ annual GPAs for the following courses: English/language arts (ELA), mathematics, science, social studies, Spanish, and “wellness.” This last course is taken every year by all students and addresses various social, emotional, and academic preparedness topics to help students better address challenges that may hinder their academic success. Students’ birth dates and genders were also obtained from the schools.

Procedure

With approval from both the College of Staten Island’s and the Lavelle Preparatory School’s IRBs, both instruments were administered by the school administration within two weeks of the end of every academic year during students’ wellness classes. Students completed the BRIEF-SR on themselves; teachers completed the BRIEF for each student in their wellness class. Absent students completed the BRIEF-SR upon their return; teachers returned all completed BRIEF forms within one week of initial administration during the wellness course. With permission, the BRIEF/BRIEF-SR and academic data were linked, de-identified, and then furnished to the authors for analysis.
# Table 1

**Number of Students by Gender, Grade Level, and Academic Year**

| Grade | 2009-2010 | | 2010-2011 | | 2011-2012 | | 2012-2013 | |
|-------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|       | Males | Females | Males | Females | Males | Females | Males | Females |
| 6     |  |  |  |  |  |  |  |  |
| 7     | 40 | 25 | 37 | 35 | 49 | 34 | 54 | 55 |
| 8     | 44 | 29 | 43 | 49 | 58 | 37 | 37 | 41 |
| 9     | 29 | 19 | 29 | 41 |
| Total | 40 | 25 | 81 | 64 | 121 | 102 | 149 | 145 |
Results

Analytic Plan

Multilevel models of change using full maximum likelihood estimations were conducted to predict the effects of executive functioning, gender, and age on annual course grades. A different model was tested for teacher responses (BRIEF scores) and student self-reports (BRIEF-SR scores) for each of the six courses: ELA, math, science, social studies, Spanish, and wellness. Analyses were conducted on standardized scores for all variables except gender where males = 0 and females = 1. Age was measured as standardized scores of students’ Julian age on the day the BRIEF or BRIEF-SR was completed. Since twelve models were run (6 courses x 2 instruments), the model α-levels were set at .01 to reduce the chance of experiment-wise Type I (false positive) errors.

As Figure 1 shows, a full four years of data were only available for 18 students. An advantage of multilevel change models over, e.g., multiple regression general linear models is that the former handles well data where not all participants provide data on all levels (Singer & Willett, 2004), as is the case here.

Detailed Findings

Table 2 presents the means and standard deviations for girls’ and boys’ BRIEF scores, BRIEF-SR scores, ages, and annual GPAs in ELA, math, science, social studies, Spanish, and wellness courses. Scores are given for each grade; note that a grade here does not represent a cohort, but summaries of all students’ scores when they were in that grade. Table 2 suggests that grades remain consistent across the grades, but that boys’ EF scores may differ from girls’ EF scores; the analyses found these differences to be significant for half of the models.

Table 3 presents the β- and, for gender, b-weights for the predictors of annual course grades for each of the models. Being based on standardized scores, these weights also represent partial
correlation coefficients between the given term and annual course GPAs (e.g., between a BRIEF/BRIEF-SR score and GPA while controlling for age and gender).

The results presented in Table 3 indicate that EF, as measured by BRIEF and BRIEF-SR scores, significantly predicted all course grades, with one exception: BRIEF-SR scores’ relationship with the wellness grades. These strong relationships between BRIEF/BRIEF-SR scores and grades remained constant over this time frame: The BRIEF/BRIE-SR x age interaction terms were not significant ($p = .01$).

Gender significantly predicted course GPA in 6 of the 12 models. It was significantly associated with GPAs for all models employing teacher data except for math; it was also significant for student data predicting wellness grades.

Students’ age at the time of BRIEF/BRIEF-SR administration was significant twice, when predicting math grades with student data and wellness grades with teacher data.

The covariance residual term tests whether the initial status on the predictor terms is related to later academic performance. In all cases it was. This means that knowing initial scores on the predictors can rather successfully predict students’ academic performance in subsequent grades. Given that the time factor (age) was rarely significant, this is not unexpected.
Table 2

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

<table>
<thead>
<tr>
<th></th>
<th>Grade 6</th>
<th></th>
<th>Grade 7</th>
<th></th>
<th>Grade 8</th>
<th></th>
<th>Grade 9</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>BRIEF (Teacher Ratings of Students’ EF)</td>
<td>Mean</td>
<td>126.2</td>
<td>106.8</td>
<td>133.2</td>
<td>114.9</td>
<td>164.8</td>
<td>146.6</td>
<td>126.1</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>33.8</td>
<td>29.2</td>
<td>33.3</td>
<td>32.1</td>
<td>31.1</td>
<td>46.5</td>
<td>37.7</td>
</tr>
<tr>
<td>BRIEF-SR (Student Self-Reported EF)</td>
<td>Mean</td>
<td>130.5</td>
<td>122.3</td>
<td>132.2</td>
<td>133.1</td>
<td>127.6</td>
<td>131.9</td>
<td>128.9</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>26.3</td>
<td>21.7</td>
<td>23.6</td>
<td>26.0</td>
<td>27.5</td>
<td>18.3</td>
<td>32.4</td>
</tr>
<tr>
<td>Student Age</td>
<td>Mean</td>
<td>12.2</td>
<td>12.2</td>
<td>13.1</td>
<td>13.1</td>
<td>14.0</td>
<td>14.0</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>ELA GPA</td>
<td>Mean</td>
<td>4.8</td>
<td>5.7</td>
<td>4.8</td>
<td>5.7</td>
<td>2.9</td>
<td>3.4</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>2.9</td>
<td>2.6</td>
<td>2.4</td>
<td>2.3</td>
<td>2.4</td>
<td>2.2</td>
<td>3.2</td>
</tr>
<tr>
<td>Math GPA</td>
<td>Mean</td>
<td>4.2</td>
<td>4.4</td>
<td>4.1</td>
<td>4.7</td>
<td>3.3</td>
<td>2.6</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>2.9</td>
<td>2.9</td>
<td>2.6</td>
<td>2.7</td>
<td>2.9</td>
<td>2.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Science GPA</td>
<td>Mean</td>
<td>3.8</td>
<td>4.3</td>
<td>3.7</td>
<td>4.5</td>
<td>5.1</td>
<td>5.5</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>2.9</td>
<td>2.8</td>
<td>2.7</td>
<td>2.9</td>
<td>2.7</td>
<td>1.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Social Studies GPA</td>
<td>Mean</td>
<td>3.8</td>
<td>4.3</td>
<td>4.1</td>
<td>4.3</td>
<td>4.2</td>
<td>2.9</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>2.8</td>
<td>2.9</td>
<td>2.1</td>
<td>2.4</td>
<td>2.8</td>
<td>2.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Spanish GPA</td>
<td>Mean</td>
<td>4.7</td>
<td>5.7</td>
<td>4.4</td>
<td>4.5</td>
<td>4.8</td>
<td>4.3</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>3.1</td>
<td>2.9</td>
<td>2.5</td>
<td>2.5</td>
<td>2.8</td>
<td>2.7</td>
<td>3.5</td>
</tr>
<tr>
<td>Wellness GPA</td>
<td>Mean</td>
<td>6.2</td>
<td>7.2</td>
<td>7.0</td>
<td>7.7</td>
<td>5.4</td>
<td>6.7</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>2.4</td>
<td>2.0</td>
<td>2.1</td>
<td>1.8</td>
<td>2.2</td>
<td>1.7</td>
<td>2.3</td>
</tr>
</tbody>
</table>
Table 3

Summary of Multi-Level Change Models Predicting Annual Course Grades

<table>
<thead>
<tr>
<th>Table</th>
<th>ELA</th>
<th>Math</th>
<th>Science</th>
<th>Social Studies</th>
<th>Spanish</th>
<th>Wellness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher $N = \sim 120$</td>
<td>Student $N = \sim 200$</td>
<td>Teacher $N = 124$</td>
<td>Student $N = 202$</td>
<td>Teacher $N = \sim 120$</td>
<td>Student $N = 202$</td>
</tr>
<tr>
<td>BRIEF/BRIEF-SR Score</td>
<td>$\beta$</td>
<td>0.60***</td>
<td>0.18**</td>
<td>0.63***</td>
<td>0.20***</td>
<td>0.66***</td>
</tr>
<tr>
<td>BRIEF/BRIEF-SR Score x Age (Change Over Time in Strength of BRIEF Effect)</td>
<td>$\beta$</td>
<td>-0.11</td>
<td>0.06</td>
<td>-0.07</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Gender (Male = 0, Female = 1)</td>
<td>$b$</td>
<td>0.07**</td>
<td>0.13</td>
<td>-0.02</td>
<td>0.02</td>
<td>-0.21**</td>
</tr>
<tr>
<td>Age (Rate of Change)</td>
<td>$\beta$</td>
<td>0.03</td>
<td>-0.17*</td>
<td>-0.01</td>
<td>-0.21***</td>
<td>0.01</td>
</tr>
<tr>
<td>Covariance Residual (Stability over time of prediction of criterion by predictors)</td>
<td>$\beta$</td>
<td>0.38***</td>
<td>0.79***</td>
<td>0.34***</td>
<td>0.76***</td>
<td>0.60***</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$; *** $p < .001$
Overall Findings

Both BRIEF and BRIEF-SR GEC scores significantly predicted course grades in all cases except when using BRIEF-SR scores to predict wellness GPAs. This relationship did not change over the three years measured here.

Students’ gender and age were occasionally also predictive of course GPAs. Even when either factor was significantly predictive, the β-weights were not large. The only reliable exception to this pattern may in the prediction of wellness GPAs. Gender was significant in both the BRIEF and BRIEF-SR models; age was significant in the BRIEF model. The β-weights for gender and age when they significantly predicted wellness grades were in the .30s, rather high values that suggest these factors account for about 10% (0.30² = 0.09, or ~10%) of the overall variance in predictions of course grades.

BRIEF/BRIEF-SR scores did not significantly change over this time. Similarly, the relationships between the predictors—primarily BRIEF/BRIEF-SR scores—and course grades was quite stable over time.

Discussion

We found that executive functioning (EF) operationalized as either BRIEF and BRIEF-SR GEC scores was strongly associated with academic performance, both when reported by the student himself/herself (BRIEF-SR) or by the student’s wellness teacher (BRIEF). Neither the BRIEF nor the BRIEF-SR scores reliably changed over the middle school years. It is therefore not surprising that relationships between executive functioning and grades were themselves rather constant, and that initial EF status well predicted academic performance in subsequent years.

These results held true for both science/math grades and ELA/social studies/Spanish grades. Therefore, higher levels of EF allow middle school students to perform well in a range of content areas.
The relationships were somewhat different for wellness grades. Gender reliably predicted wellness grades; girls usually received higher course grades than boys. When using teacher (BRIEF) ratings, age mattered: Wellness grades tended to be higher among older students. This pattern may reflect general development trajectories in which girls show faster cognitive and behavioral development than boys. Although well established in the literature, this conclusion must remain tentative in the present study, since it was the wellness teachers who rated the EF of their students, so we cannot rule out biases in either the teachers’ ratings or the grades they gave their students. This bias is much less likely among the models predicting grades in courses where those teachers did not complete the BRIEF, or when students were self-reporting EF.

The model predicting wellness grades with student (BRIEF-SR) data do show some similarity to those predicting wellness grades from teacher (BRIEF) data: Gender still predicted wellness grades. Interestingly, this was the only occasion when executive functioning was not related to course grades. If this pattern is maintained in future waves, this may represent an interesting divergence in executive functioning reported by students and their teachers. The wellness curriculum is unique to this school, but nonetheless covers topics like de-escalation of interpersonal conflict and time management that we expected would be well associated with executive functioning. Little more can be said on the topic without more focused analysis of what exactly wellness grades measure.

Despite the somewhat different pattern of results found when predicting wellness grades, EF proved itself to be an important predictor of course grades among a diverse population of middle school students. We did not find sufficient evidence for either a change in EF or in the effect of EF on academics. These findings support of Best, Miller, and Naglieri (2011) who found that the growth of EF among the general population tends to slow down during mid- to late-adolescence.

The sampled student body contained a large portion of students with a variety of physical, emotional, behavioral, and cognitive needs. Many of those with acute needs have been diagnosed and prescribed individual education plans (IEPs). The developmental trajectory of EF is expected to be affected by factors included in at least some of these IEPs; investigating
sub-groups of the students based on their varied diagnoses may afford a more sophisticated understanding of this interaction that presents differing trajectories overlooked in the overall population.

Conclusion

Whatever additional details further study reveals, we can conclude now that BRIEF and BRIEF-SR scores were strongly associated with a broad range of academic performance here, and that this relationship was rather constant in a population of urban, middle schools students. It is stable enough that an initial diagnosis in early grades can well predict academic performance in subsequent middle school grades. For example, it appears that the BRIEF/BRIEF-SR could be included in a battery of intake instruments to help school personnel plan a few years of academic support for middle school students.

Limitations

In addition to the somewhat limited range of time, all students hailed from the same school. Admittance to the school is by lottery among the many applicants, so they at least initially well represent a population of diverse students from disadvantaged backgrounds. However, their developmental trajectories after enrollment are not independent nor fully generalizable.
References


