Examining the Validity of Behavioral Self-Regulation Tools in Predicting Preschoolers' Academic Achievement

Sara A. Schmitt

Purdue University

Megan E. Pratt

Megan M. McClelland

Oregon State University

Schmitt, S. A., Pratt, M. E., & McClelland, M. M. (2014). Examining the validity of behavioral self-regulation tools in predicting preschoolers' academic achievement. *Early Education and Development, 25*(5), 641-660. doi: 10.1080/10409289.2014.850397

Author Note

Sara A. Schmitt, Department of Human Development and Family Studies, Purdue University; Megan E. Pratt, College of Public Health and Human Sciences, Oregon State University; Megan M. McClelland, College of Public Health and Human Sciences, Oregon State University.

This research was supported by the U.S. Department of Education Institute for Education Sciences (grant #R305A100566).

Correspondence concerning this article should be addressed to Sara A. Schmitt, 1202 W. State St., Room 235, Purdue University, West Lafayette, IN 47904. E-mail: saraschmitt@purdue.edu.

Abstract

The current study investigated the predictive utility among teacher-rated, observed, and directly assessed behavioral self-regulation skills to academic achievement in preschoolers. Specifically, this study compared how a teacher report, the Child Behavior Rating Scale, an observer report, the Observed Child Engagement Scale, and a direct assessment, the Head-Toes-Knees-Shoulders task, relate to early math and literacy skills. The sample consisted of 247 children from 31 preschool classrooms. Trained research assistants observed a subsample of 104 children. Results indicated significant, positive relationships for teacher-rated and directly assessed behavioral self-regulation for early math and literacy skills. Teacher ratings were the strongest predictors of literacy, and the direct assessment emerged as the strongest predictor of math. Observed behavioral self-regulation was not significantly related to either academic domain. Discussion focuses on domain specificity of behavioral self-regulation assessments and the importance of utilizing multiple measurement tools when assessing behavioral self-regulation and relations to early achievement.

Keywords: behavioral self-regulation, academic achievement, measurement

Examining the Validity of Behavioral Self-Regulation Tools in Predicting Preschoolers' Academic Achievement

A child's success in school depends on a constellation of factors including early academic skills, socio-emotional skills, and self-regulation (Blair, 2002). In an attempt to improve academic achievement of young children in the U.S., early math and literacy skills have been targets of curriculum development in early education programs. Recent research, however, argues that although these early academic skills are important for school success, aspects of self-regulation, specifically *behavioral self-regulation,* provide a strong foundation for school achievement in conjunction with early academic skills and other environmental factors (McClelland et al., 2007). Behavioral self-regulation consists of three aspects of executive function: working memory, attentional or cognitive flexibility, and inhibitory control. In practice, children's ability to regulate their behavior in classroom contexts includes maintaining and switching attention, remembering and following instructions, engaging in and completing classroom tasks, complying with classroom rules, and resisting inappropriate responses or activities (McClelland, Cameron Ponitz, Messersmith, & Tominey, 2010). Although considerable research supports the short- and long-term relations between behavioral self-regulation and academic outcomes (Andersson & Bergman, 2011; Blair & Razza, 2007; McClelland, Acock, Piccinin, Rhea, & Stallings, 2013; McClelland et al., 2007; Moffitt et al., 2011; Wanless, McClelland, Acock, Chen, & Chen, 2011), measurement concerns remain, and there is a lack of consensus on which assessments are the most practical and effective in predicting early academics (Blair, Zelazo, & Greenberg, 2005; Carlson, 2005). Traditional methods for measuring behavioral self-regulation in classroom contexts include teacher ratings, which have been very useful (McClelland, Acock, & Morrison, 2006), but may also be

susceptible to rater bias (Waterman, McDermott, Fantuzzo, & Gadsden, 2012). Recent advancements in the measurement of behavioral self-regulation include observational tools and direct assessments, but there has been limited research comparing the predictability of these contemporary measures to teacher reports. As evidence of the value of behavioral self-regulation for academic success accumulates, it is increasingly important that parents, teachers, and researchers employ the most effective measurements when assessing and identifying children at risk for poor behavioral self-regulation. The current study extends current literature by comparing the predictive utility of three global measures of behavioral self-regulation in preschool: a teacher report called the Child Behavior Rating Scale (CBRS; Bronson, Tivnan, & Seppanen, 1995), an extramural observer report, which is a modified version of the Observed Child Engagement Scale (OCES; Rimm-Kaufman, 2005), and a direct assessment, the Head-Toes-Knees-Shoulders task (HTKS; McClelland et al., 2007).

Importance of Behavioral Self-regulation for Academic Success

A large body of work suggests that behavioral self-regulation is important for school success in early childhood and beyond (Andersson & Bergman, 2011; Blair & Razza, 2007; Cameron Ponitz, McClelland, Matthews, & Morrison, 2009; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; McClelland et al., 2007; Moffitt et al., 2011). Indeed, the development of strong working memory, attention, and inhibitory control skills is related to higher math and literacy outcomes during preschool (Blair & Razza, 2007; McClelland et al., 2007) and kindergarten (Cameron Ponitz, McClelland, et al., 2009; Howse et al., 2003). Moreover, research supports the notion that these skills lay the foundation for successful academic trajectories (McClelland et al., 2006; McClelland et al., 2013). For example, in a recent study, children's attention at age 4 was related to college completion by age 25, controlling for math

and literacy skills and other demographic variables (McClelland et al., 2013). Behavioral self-regulation has also emerged as a personal asset for achievement (McClelland & Wanless, 2012) and a compensatory factor for children experiencing demographic risk, such as homelessness and poverty (Obradovic, 2010; Sektnan, McClelland, Acock, & Morrison, 2010). Taken together, as research continues to demonstrate the importance of behavioral self-regulation for achievement trajectories, it is critical to evaluate the predictive validity of various tools measuring these skills.

Teacher Report Measures of Behavioral Self-Regulation

Teacher ratings have been commonly used to assess children's behavioral self-regulation in classroom contexts, and a number of benefits are associated with these measures. For example, teacher ratings can be completed quickly and have shown moderate to strong levels of internal reliability (Bronson et al., 1995; Konold, Jamison, Stanton-Chapman, & Rimm-Kaufman, 2010; Sektnan et al., 2010). In addition, teachers typically have ample opportunities to observe children in various classroom contexts and activities where behavioral self-regulation is often required (e.g., during frustrating or demanding tasks) and, therefore, are able to provide information about these achievement-relevant behaviors (Sattler, 1988). Finally, teacher ratings demonstrate stability over time such that children's relative ranking on teacher reports at one time point is related to their relative ranking at a later time point (McClelland & Morrison, 2003). In other words, teacher-rated behavioral self-regulation at one time point predicted behavioral self-regulation at a subsequent time point.

A number of teacher-rated measures assess aspects of behavioral self-regulation in the classroom environment and have demonstrated predictive validity with academic outcomes in preschool and elementary school (Adler & Lange, 1997; Bronson et al., 1995; Clark et al., 2010; Cooper & Farran, 1991; Gioia, Isquith, Guy, & Kenworthy, 2000; Gresham & Elliot, 1990;

Howse et. al, 2003; Matthews, Cameron Ponitz, & Morrison, 2009; McClelland, Morrison, & Holmes, 2000; McClelland et al., 2006; Sektnan et al., 2010; Wanless, McClelland, Acock, Chen, et al., 2011). Many of the commonly used teacher-rated assessments conceptualize behavioral self-regulation within a more broadly defined learning-related skills framework (Adler & Lange, 1997; Cooper & Farran, 1991; Gresham & Elliot, 1990). For example, the work-related skills subscale of the Cooper-Farran Behavioral Rating Scales (CFBRS; Cooper & Farran, 1991) includes both indicators of behavioral self-regulation and aspects of social competence. Also, the Instrumental Competence Scale for Children (COMPSCALE; Adler & Lange, 1997), a teacher rating that includes both children's behavioral self-regulation *and* motivation in the classroom. Although both the CFBRS and COMPSCALE have been useful in predicting children's achievement (Howes et al., 2003; McClelland et al., 2000; McClelland et al., 2006), these teacher-rated assessments are more broadly focused on behavioral self-regulation as one part of a larger learning-related skills concept and do not explicitly focus on the predictive utility of working memory, attention, and inhibitory control.

There are few teacher-rated measures that are more narrowly focused on behavioral self-regulation as conceptualized in this study. One example is the Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000). The BRIEF is a 63-item standardized checklist that assesses an array of distinct sub-domains of executive function skills, including working memory, attention, and inhibitory control. Although the BRIEF has been validated for early elementary-age children, the majority of the work has examined clinical samples, such as children with ADHD and developmental disorders (Gioia, Isquith, Kenworthy, Barton, 2002; Mahone et al., 2002). Fewer studies have examined the validity of this scale in normative

populations (e.g., Clark, Pritchard, & Woodward, 2010). Additionally, the BRIEF is relatively long (63-items), which may make it less practical to employ in a school setting.

Thus, considering both conceptual and practical issues, the current study utilized the CBRS (Bronson et al., 1995), a short 10-item teacher report that explicitly mesaures the behavioral aspects of self-regulation important for success in classroom contexts. The CBRS has been validated and used to measure behavioral self-regulation in several previous studies, and has shown predictive utility for achievement outcomes in culturally and economically diverse samples (Matthews et al., 2009; Suchodoletz et al., 2013; Wanless, McClelland, Acock, Chen, et al., 2011).

Although employing teacher ratings of behavioral self-regulation like the CBRS carries many benefits, these ratings could also be susceptible to rater bias. In particular, teacher ratings could be influenced by the amount of contact the teacher has with children's parents (Serpell & Mashburn, 2012), the order in which they complete ratings (Brandon, Kehle, Jenson, & Clark, 1990), and/or the gender of the child (i.e., teachers tend to rate boys lower than girls on behavioral measures; Beaman, Wheldall, & Kemp, 2006). Moreover, evidence suggests potential bias in teacher ratings of behavior when teacher and child ethnicities are different (e.g., White teachers rate Black children's behavior more negatively than white children; Downey & Pribesh, 2004), although these findings are mixed (de Ramirez & Shapiro, 2005; Hosterman, DuPaul, & Jitendra, 2008). Teachers rely on reference groups when rating children's behavior as a way to gauge what is developmentally appropriate. Differences in teachers' classroom experiences and education may create highly variable reference groups, which could result in additional biases in teacher ratings (Phillips & Lonigan, 2010). Finally, when comparing teacher-administered academic assessments to outside/independent raters, teachers have been

found to vary significantly more in their appraisals (i.e., 28-31% of children's score variation was attributed to teachers, not children), indicating that extramural assessors may provide more accurate depictions of children's individual differences in behavior (Waterman et al., 2012).

Although subjectivity is always a potential downfall when utilizing any adult report measure, some research has found that trained classroom observers may provide less subjective reports than teachers on children's behavior due to specialized training, exposure to various reference groups from which to compare children, and a lack of emotional investment with study samples (Phillips & Lonigan, 2010; Whitebread et al., 2009).

Observational Measures of Behavioral Self-Regulation

Global classroom observational assessments of children's behavior are gaining popularity in the field and are often used in conjunction with individualized teacher reports and direct assessments. To date, few observational measures have been developed to assess behavioral self-regulation although broader observational tools are available that capture some of the self-regulation domains. For example, the Individualized Classroom Assessment Scoring System (inCLASS; Downer, Booren, Lima, Luckner, & Pianta, 2010) is an observational assessment that measures children's classroom competence across three domains: teacher interactions, peer interactions, and task orientation. Although the inCLASS is a promising global assessment of children's behavior in classroom contexts, it was not specifically designed to measure children's behavioral self-regulation.

Another observational measure that taps aspects of children's behavioral self-regulation is the Academic Engaged Time (AET; Walker & Severson, 1990) coding system. This measure compares the amount of time that children are engaged appropriately with academic tasks with the overall observational time. Behaviors that constitute academic engaged time include

listening and paying attention to the teacher, asking for assistance, and following directions. Although this measure captures many of the aspects of behavioral self-regulation in the classroom, it was developed to identify clinical behavioral problems as part of the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1990), not to assess typical classroom behavioral self-regulation. Moreover, this measure has primarily been used with children in kindergarten and above (Walker, Golly, McClane, & Kimmich, 2005; Walker et al., 2009), and the behaviors recorded during AET observations may not be appropriate in younger, preschool-aged samples.

To our knowledge, only one published study has utilized an observational tool specifically designed to measure self-regulation in classroom settings. Whitebread and colleagues (2009) developed the Cambridgeshire Independent Learning in the Foundation Stage (C.Ind.Le) Coding Scheme, an observational framework used to assess self-regulation and metacognition for children ages 3-5 years. Although this measure proved useful in measuring self-regulation and metacognitive skills, it relied on observers coding lengthy video recordings of children's behavior during classroom activities, which can be an arduous process. Moreover, this measure was designed to assess children's self-regulation more broadly by including affective and emotional components (i.e., emotional/motivational monitoring and control) in addition to the more cognitive components. Finally, the cognitive components that are coded using the C.Ind.Le Coding Scheme (e.g., monitoring, planning, evaluation) do not explicitly map onto the components of behavioral self-regulation (working memory, attentional flexibility, and inhibitory control) of interest in the present study.

To determine the predictive utility of extramural observer ratings in the current study, we were interested in utilizing a practical, observational tool that focused on the same components

of behavioral self-regulation (working memory, attentional flexibility, and inhibitory control) as the CBRS and HTKS. Thus, we employed the OCES (Rimm-Kauffman, 2005), a relatively short (8-minute), live-coded, observer rating measuring behavioral self-regulation in classroom contexts.

OCES Measure and Previous Studies

The OCES was adapted from the National Institute of Child Health and Human Development Early Child Care Research Network (2005) Classroom Observation Scale. The OCES has been used to evaluate children's engagement in learning during classroom activities (Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Cameron Ponitz, Rimm-Kaufman, Grimm, & Curby, 2009; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009), and has shown utility in predicting reading proficiency in young children (Cameron Ponitz, Rimm-Kaufman, et al., 2009). This observational rating scale consists of eight categories that capture aspects of children's global classroom engagement. In the current study, we used four of the categories in the OCES (self-reliance, engagement, attention, and compliance) to measure behavioral self-regulation, which reflects the work of researchers who have defined behavioral self-regulation as the ability to inhibit inappropriate actions, attend to stimuli, and follow directions (Baumeister & Vohs, 2004; McClelland et al., 2007).

In classroom contexts, a child with strong observed behavioral self-regulation is likely to: (a) show active *engagement* and focus during learning tasks (Shonkoff & Phillips, 2000); (b) maintain and switch *attention* during classroom activities, while ignoring distractions (Rothbart & Posner, 2005; Rueda, Posner, & Rothbart, 2004); (c) demonstrate *self-reliance* by working well on his or her own in both structured and unstructured classroom settings (Schunk & Zimmerman, 1997); and (d) readily show *compliance* with teacher requests (Ramini, Brownell,

& Campbell, 2010). A previous study demonstrated that behavioral self-regulation, as measured by direct assessments similar to the HTKS, at the beginning of the kindergarten year was significantly related to levels of observed classroom engagement as measured by the OCES later in the year (Brock et al., 2009), suggesting that the OCES may tap the same behavioral self-regulation components as the teacher-rated CBRS and directly-assessed HTKS. Little work, however, exists that has explicitly compared the utility of observer report to teacher ratings or direct assessments. Similar to teacher ratings, observer reports may still be susceptible to rater bias, so there is a need to evaluate whether teacher-rated, observed, and direct assessments vary in their predictive relations to academic outcomes to determine the most effective tools for assessing behavioral self-regulation in preschool children.

Direct Assessments of Behavioral Self-Regulation

The development of direct assessments has increased dramatically in the last decade given the noted importance of behavioral self-regulation for young children's academic success (Andersson & Bergman, 2011; Blair & Razza, 2007; McClelland et al., 2007). Although direct assessments have been viewed as improvements in the measurement of behavioral self-regulation, many were designed for clinical or laboratory populations and require substantive materials and time to administer (McClelland & Cameron Ponitz, 2012). Moreover, several tasks measure only one aspect of behavioral self-regulation such as inhibitory control, which may not be useful when trying to assess an integrative construct that includes attentional flexibility and working memory *in addition* to inhibitory control (McClelland & Cameron Ponitz, 2012). To address these limitations, the HTKS (Cameron Ponitz et al., 2008; Cameron Ponitz, McClelland, et al., 2009; McClelland et al., 2007) task was developed to assess the integration of self-regulation domains most relevant to classroom success. This task has

demonstrated strong psychometric properties, including predictive validity for achievement outcomes (Matthews et al., 2009; McClelland et al., 2007), and some evidence suggests the HTKS may be a stronger predictor of early academics than teacher report on the CBRS (Matthews et al., 2009; Wanless, McClelland, Acock, Chen, et al., 2011). However, additional research that explicitly focuses on comparing measures is needed to validate this finding.

The Present Study

In the present study, we examined the relative efficacy of two adult reports (teacher and observer ratings) and a direct assessment of behavioral self-regulation in predicting achievement outcomes in preschool. We expected that all three assessments would be significantly related to academic outcomes. In addition, given previous reports indicating that the HTKS is more strongly related to achievement compared to teacher reports (Matthews et al., 2009; Wanless, McClelland, Acock, Chen, et al., 2011), we hypothesized that the HTKS would explain the most unique variance in math and literacy scores. Finally, we expected that observer ratings of behavioral self-regulation would better predict achievement compared to teacher ratings due to the relative neutrality of extramural observations, which tend to involve less emotional investment, a larger comparison group, and more intensive training procedures compared to teacher reports (Phillips & Lonigan, 2010).

In all statistical analyses, we included important covariates related to academic outcomes: child age, gender, and parent education level (as a proxy for socioeconomic status; Davis-Kean, 2005; Halpern, 2000; Sirin, 2005). We used parent education level based on research showing strong predictive relations to children's language and academic outcomes (e.g., Magnuson, Sexton, Davis-Kean, & Huston, 2009; McClelland et al., 2007).

Method

Participants

Participants were recruited from 16 local preschools participating in a larger, longitudinal study focused on children's behavioral self-regulation. Participants included 247 children (118 female) enrolled in 31 classrooms. Fifty-one percent of children were enrolled in Head Start classrooms. Children ranged in age from 42- to 70-months-old and attended NAEYC accredited preschools located in a small city in the Pacific Northwest. Children were 60% White; 20% Latino; 1% African American; 4% Asian; and 15% other ethnic groups; 14% of children were Spanish-speaking. Parent education level was approximately two years of college. Written consent was obtained from 42 teachers (including head and assistant teachers) and parents prior to participation, and all participants received a $20 gift card.

A subsample (n = 104) of children was observed by trained research assistants. T-tests and logistic regressions revealed that children in the observed subsample were not statistically different from the full sample based on demographic information or child measures such as early math and literacy scores or behavioral self-regulation.

Procedure

Data collection took place in the spring of the academic year. Parents completed demographic questionnaires, teachers completed the CBRS to assess children's behavioral self-regulation in classroom contexts, and trained research assistants conducted single, 8-minute observations of individual children in their classrooms during free play activities. The HTKS and academic assessments were administered in a quiet area of the classroom. A fluent, Spanish-speaking research assistant assessed children identified as Spanish-speaking by their teachers.

Measures

Parent demographic questionnaire. Parents completed a demographic questionnaire in English or Spanish, which provided information including child's age, gender, ethnicity, native language, and the level of parent education.

Teacher reports of behavioral self-regulation. Teachers completed the CBRS for individual children in their classes, which measures behavioral self-regulation in the classroom (Cameron Ponitz, McClelland, et al., 2009; Matthews et al., 2009: McClelland et al., 2007). In previous studies, factor analyses revealed a behavioral self-regulation factor consisting of 10 items on the CBRS, tapping the integration of working memory, attentional flexibility, and inhibitory control (Cameron Ponitz, McClelland, et al., 2009; Wanless, McClelland, Acock, Chen, et al., 2011). In the current study, teachers rated these 10 items on a 5-point Likert scale, with 1 indicating children *never* displayed certain behaviors and 5 indicating that children *always* displayed certain behaviors. Items included statements such as "Observes rules and follows directions without requiring repeated reminders" and "Completes tasks successfully." Item scores were totaled to create a sum score for each child. Consistent with past research (α >.94; Bronson, et al., 1995; Cameron Ponitz, McClelland, et al., 2009; Wanless, McClelland, Acock, Chen, et al., 2011), this measure demonstrated high internal reliability in the current sample (α >.96). In addition, the CBRS has demonstrated concurrent and longitudinal validity in predicting academic outcomes in past work (Lim, Rodger, & Brown, 2010; Matthews et al., 2009; Suchodoletz et al., 2013; Wanless, McClelland, Acock, Chen, et al., 2011).

Observations of behavioral self-regulation. During an 8-minute observation in the classroom during free play, trained research assistants rated children using the OCES (Rimm-Kaufmann, 2005). The full scale consists of eight observable categories. A rubric was used during classroom observations that contained child behavioral codes within each category to

guide observers' ratings. Exploratory factor analyses were conducted on the eight observable OCES categories, and three factors were identified. One of the factors was pertinent to the present study. This factor consisted of four OCES categories (engagement, self-reliance, attention, and compliance). The child behavioral codes within these four categories reflected the attention and inhibitory control aspects of behavioral self-regulation. For example, child behavioral codes tapping attention were "Resists intrusion to attention" and "Shows little or no distractibility" and tapping inhibitory control was "Manages self, materials, and tasks in less structured settings and challenging environments." The two factors that were not utilized in the current analyses were defined as: problem behavior (consisting of negative affect and disruptive behaviors) and positive classroom behavior (consisting of positive affect and cooperation with peers). These factors did not tap the aspects of behavioral self-regulation of interest in the current study, and thus, were not used in analyses.

Past research has identified a factor from the OCES representing global child engagement, a construct similar to behavioral self-regulation, consisting of five categories of the full OCES scale (i.e., engagement, self-reliance, attention, compliance, and disruptive behavior; Brock et al., 2009; Cameron Ponitz, Rimm-Kaufman, et al., 2009; Rimm-Kaufman et al., 2009). We excluded the disruptive behavior category for statistical and conceptual reasons. First, the disruptive behavior category did not load strongly onto the behavioral self-regulation factor, it loaded onto a second factor. Moreover, it was only significantly correlated with two of the four categories (self-reliance and compliance) that loaded onto the behavioral self-regulation factor. In addition, the child behavioral codes associated with the disruptive behavior category were not conceptually related to the components of behavioral self-regulation of interest in the present study. For example, the codes "Earns two or more reprimands" or "Behaviors annoy other

children" do not reflect the integration of working memory, attentional flexibility, or inhibitory control.

For each of the four OCES categories used, a score ranging from 1 to 7 was recorded following each observation. A rating of 1 indicated low levels of the behavior and a 7 indicated high levels of the behavior. OCES scores on the four categories were averaged to create a composite behavioral self-regulation score for each child. These four items were all moderate to highly correlated with each other *(rs* > .41, *ps* < .001) and showed strong internal reliability (α >.87). Prior to data collection, OCES raters were trained on video segments and live observations of children until inter-rater reliability of >.80 was reached for each category. Coder drift was tested bi-weekly and high rater agreement was maintained throughout the data collection period (*ICC* = .92).

Direct measure of behavioral self-regulation. The HTKS task is a behavioral assessment that taps the integration of the three executive function components of behavioral self-regulation: attentional flexibility, working memory, and inhibitory control (Cameron Ponitz, McClelland, et al., 2009; McClelland et al., 2007). The HTKS consists of three test phases. During the first phase, children are asked to respond naturally to two commands (e.g., "Touch your head" and "Touch your toes"). Then children are asked to do the opposite of the original instruction. In subsequent parts of the task, additional rules are added to increase cognitive complexity. The HTKS consists of 30 items, with a range in scores of 0 to 60. Children were given a score of 0 for an incorrect response; 1 for a self-corrected response; and 2 for a correct response. Previous research indicates high inter-rater reliability (κ > .90) and validity of this measure in assessing children's behavioral self-regulation with culturally diverse samples (McClelland et al., 2007; Suchodoletz et al., 2013; Wanless, McClelland, Acock, Chen, et al.,

2011; Wanless, McClelland, Tominey, & Acock, 2011). In addition, this measure has been positively and significantly correlated with the CBRS ($rs > .28$, $ps < .01$; Cameron Ponitz, McClelland, et al., 2009; McClelland et al., 2007).

Academic outcomes. Academic achievement was assessed using two subtests of either the Woodcock-Johnson III (WJ-III) or the Batería III Woodcock-Munoz: Letter-Word Identification and Applied Problems. Previous research has demonstrated high reliability and validity ($\alpha > .80$) for both of the subtests (Schrank et al., 2005; Woodcock & Mather, 2000). *W*-scores were used in the analyses, which are standardized scores based on the average performance of a typical child at a given age (based on normative data; Jaffe, 2009).

Early math skills. The Applied Problems subtest of the WJ-III or the Batería III Woodcock-Muñoz was employed to measure early mathematical problem-solving skills, including counting, reading numbers, and basic addition and subtraction.

Emergent literacy skills. The Letter-Word Identification subtest of the WJ-III or the Batería III Woodcock-Muñoz was employed to measure emergent literacy skills, including naming letters and reading words out loud.

Results

The primary goal of the present study was to determine the relative efficacy of two adult reports and a direct assessment of behavioral self-regulation in classroom contexts as they relate to children's achievement in preschool.

Analytic Strategy

Data analyses were conducted using Stata 12.0 (StataCorp., 2011). Due to the nested structure of the data (children nested in classrooms), multilevel modeling was utilized to answer all research questions. Unconditional models were first estimated, which provided ICCs for WJ:

Applied Problems (*ICC* = .20) and for WJ: Letter-Word Identification (*ICC* = .10). The effect of behavioral self-regulation assessments on math and literacy outcomes in preschool was modeled as:

$Y_{ij} = \beta_{0j} + \beta_{1ij}$ (child age) + β_{2ij} (parent education)+ β_{3ij} (gender) + β_{4ij} (CBRS) + β_{5ij} (OCES) + β_{6ij} (HTKS) + r_{ij}

Y_{ij} represents the average academic outcome in preschool for child *i* in classroom *j*, accounting for child age, parent education, gender, and error all at the child level. Standardized effect sizes were calculated using the following equation:

$$\beta_{y.x} = (b_{y.x} \times s_x)/ s_y$$

Missing data. A subsample of 104 children in the study was observed in the classroom, meaning that OCES data were missing for 60% of children. In addition, data were missing on CBRS (4% missing), WJ: Applied Problems (10% missing), and WJ: Letter-Word (7% missing). To handle missing data, multiple imputation was employed for hypothesis testing using Stata Version 12.0 (StataCorp., 2011) to reduce potential bias that could result from using listwise deletion (Acock, 2012). All data were assumed to be missing at random (MAR), which requires that all variables associated with missingness are included in models and that all other patterns of missingness are random (Schafer & Graham, 2002). There is no definitive test of the MAR assumption, however, tests were conducted to determine whether auxiliary variables not included in original models were related to missingness. Logistic regressions were run using dummy variables that were created for all variables with > 5% missingness (0 = present; 1 = missing). Variables already in the model and demographic variables theoretically related to missingness were included as predictors in the logistic regressions. None of the auxiliary variables included

in the models predicted missingness, suggesting that it is reasonable to assume that missing data were missing at random.

Due to the large percentage of missing data on the OCES measure, all analyses were also run including only children who were observed (n = 104). Similar patterns of results that emerged using the full sample were found in analyses with only the subsample. Thus, analyses utilizing the full sample and imputed data are reported.

Descriptive statistics. Descriptive statistics can be found in Table 1, and correlations for all predictor, outcome and control variables can be found in Table 2. The CBRS and OCES were significantly related to each other (r = .27, p = .006), and the CBRS and HTKS were significantly correlated (r = .35, p < .001), but the HTKS was not significantly related to the OCES. The CBRS and HTKS were correlated with early math (r = .40 and r = .64, respectively) and literacy (r = .41 and r = .43, respectively). The OCES was significantly related to math (r = .21, p = .036). Parent education and child age were positively and significantly related to math (r = .48 and r = .26, respectively) and early literacy (r = .42 and r = .17, respectively). Gender was not significantly correlated with any variables with the exception of the CBRS (r = -.24, p < .001) where teachers rated boys lower than girls on classroom behavioral self-regulation in preschool.

How Do Teacher-Rated, Observed, and Direct Assessments of Behavioral Self-Regulation in Preschool Vary in Their Ability to Predict Academic Outcomes? Results indicated significant and positive relations between behavioral self-regulation, as measured by the CBRS and the HTKS, and early math (WJ-III: Applied Problems) and literacy (WJ-III: Letter-Word Identification; see Table 3). Children who were rated higher on the CBRS, scored significantly higher on early math (β = .21, p < .001) and early literacy (β = .32, p < .001). Similarly,

children with higher scores on the HTKS had significantly higher scores on early math (β = .43, $p < .001$) and early literacy (β = .18, $p = .004$). Hypotheses were partially supported in that the HTKS accounted for more unique variance than either adult report on math. However, the CBRS accounted for more unique variance than the OCES and the HTKS on literacy. Contrary to hypotheses, observer ratings (OCES) were not significantly associated with math or literacy.

Of the covariates, parent education was significantly related to math (β = .20, $p = .001$) and literacy (β = .28, $p < .001$) and age was significantly related to early math (β = .12, $p = .014$).

Discussion

In the present study, we explored how three assessments of behavioral self-regulation (teacher-rated, observed, and directly assessed) compared in their predictability of math and literacy in preschool. Results indicated that teacher ratings and the direct assessment were significantly related to early math and emergent literacy, but observer reports were not significantly associated to either academic domain. Domain specific patterns emerged in our comparisons of the predictive utility of assessments, such that teacher ratings were the strongest predictors of literacy, and the direct assessment was the strongest predictor of math. This study contributes to current literature by exploring the efficacy of three different assessments of behavioral self-regulation in preschool classroom contexts as they relate to achievement.

Efficacy of Teacher-Rated, Observed, and Directly Assessed Behavioral Self-Regulation in Predicting Early Achievement

Consistent with previous literature, we found that teacher reports on the CBRS were significantly and positively associated with math and literacy achievement (Matthews et al., 2009; Suchodoletz et al., 2013; Wanless, McClelland, Acock, Chen, et al., 2011). Also in line

with other work, children who scored higher on the HTKS were more likely to perform better on tests of early math and early literacy (Cameron Ponitz, McClelland, et al., 2009; Matthews et al., 2009; McClelland et al., 2007; Wanless, McClelland, Acock, Chen, et al., 2011). As expected, the HTKS emerged as a stronger predictor of math relative to the CBRS. Previous research has demonstated that the HTKS is a robust predictor of math (Cameron Ponitz, McClelland, et al., 2009; McClelland et al., 2007; Wanless, McClelland, Acock, Chen, et al., 2011), and in a recent study, Matthews et al. (2009) found that links between behavioral self-regulation and math were stronger for the HTKS than for the teacher-rated CBRS. The HTKS task may be a better predictor of math than the CBRS because of the cognitive demands it places on children (Cameron Ponitz, McClelland, et al., 2009; Matthews et al., 2009). For example, the three components of behavioral self-regulation (working memory, attentional flexibilty, and inhibtory control) that the HTKS tap are also required to do well on early math assessments, such as the WJ Applied Problems subtest. As items in this subtest increase in difficulty, children are required to draw upon behavioral self-regulation: children must pay attention to the assessor's instructions, remember mathematical rules, and inhibit the inclination to quickly respond with incorrect solutions. Therefore, the cogntive skills that are required to do well on direct assessments of behavioral self-regulation, such as the HTKS, also lay the foundation for success in math achievement (Bull, Espy, & Wiebe, 2008; Clark et al., 2010; Espy et al., 2004; Willoughby, Blair, Wirth, & Greenberg, 2012).

Athough the HTKS emerged as a stronger predictor of math (and was related to children's literacy), the CBRS was the strongest predictor of literacy. It was unexpected that the HTKS was not the most predictive of literacy given other work demonstrating robust links between the HTKS and emergent literacy (Matthews et al., 2009; McClelland et al., 2007;

Wanless, McClelland, Acock, Chen, et al., 2011). The current results were consistent with findings from two recent studies, however, where the CBRS outperformed the HTKS in predicting early literacy (Duncan & McClelland, 2013; Suchodoletz et al., 2013). These results, along with those in the current study, could reflect the tendency for preschool classrooms to be rich in literacy activities, with teachers spending more time explicitly engaged in langauge and literacy instruction than in math instruction (Layzer, Goodsen, & Moss, 1993). This time spent in literacy instruction could allow teachers to better evaluate the regulatory behaviors important for literacy development than a direct assessment that was administered on just one occasion. Furthermore, in the current study, data collection took place in the spring of the academic year, which gave teachers the advantage of observing children's behaviors over a variety of contexts and activities over the school year, which may be particularly important when measuring behavioral self-regulation as it relates to literacy outcomes.

The current findings suggest that the utility of measures of behavioral self-regulation may be domain specific. In other words, certain measures may be particularly useful for predicting math, and other measures may be better for predicting literacy. Domain specificity has emerged in previous studies exploring the predictive relations between measures of behavioral self-regulation and academic outcomes (Cameron Ponitz, McClelland, et al., 2009; Suchodoletz et al., 2013). For example, Cameron Ponitz and colleagues (2009) found that behavioral self-regulation, as measured by the HTKS, predicted gains in math over the kindergarten year, but not gains in literacy. Moreover, when comparing the utility of the HTKS and CBRS in a sample that included German and Icelandic chidlren, Suchodoletz et al. (2013) found that the HTKS was a more robust predictor of math, and the CBRS was a stronger predictor of different aspects of literacy for the Icelandic children.

This domain specificity could be, in part, explained by contextual variations in assessments. For example, although the CBRS and the HTKS may measure the same conceptual construct (Cameron Ponitz et al., 2008; McClelland et al., 2007), the HTKS is individualy administered and provides a snapshot of children's behavioral self-regulation during one cognitively demanding task. During this task, the integration of working memory, attentional flexibilty, and inhibitory control are scored as children play a challenging game. Children likely call upon these same skills during a challenging math activity. For example, during a color and size sorting and ordering activity, children must first remember the instructions to sort by color and then order by size. They must also sustain attention and simultaneously consider two different properties of an object (color and size). Finally, they must resist the tendency to attend to other properties of the object or give up on the task. In contrast to the HTKS, the teacher-rated CBRS provides a perception of classroom behavior across a variety of activities and contexts over the course of an academic year, which may allow teachers opportunities to observe the skills that are important for literacy development. For example, while observing children engaged in literacy activities, teachers may be able to report on the automaticity of letter recognition (utilizing working memory skills) that may not be so easily measured by a direct assessment of behavioral self-regulation. Because contextual differences may play a role in the utility of measures in predicting academic domains, utilizing direct assessment in conjunction with teacher ratings may be the most effective for measuring children's behavioral self-regulation comprehensively as it relates to achievement.

Contextal variation in measurement may have also played a role in the null results found with the relations between observer reports on the OCES and achievement. Recent work indicates that context is particularly important when measuring behavioral self-regulation in

classroom settings, in that the extent to which a child regulates his or her behavior depends on various classroom situational demands (Bulotsky-Shearer, Fantuzzo, & McDermott, 2008; McClelland & Cameron Ponitz, 2012). In the current study, extramural classroom observations took place during free choice activities. Although complex behaviors and interactions that require behavioral self-regulation do occur during free choice activities (Kontos, Burchinal, Howes, Wisseh, & Galinsky, 2002), task demands typically increase during more structured activities (e.g., individual seat work; Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). Thus, the ability to *observe* children's regulatory behaviors may be more difficult during free choice time, which in turn, could limit the predictive utility of observer reports to achievement.

Furthermore, observations took place over the course of one relatively short period of time, which may not have been enough time to adequately observe behavioral self-regulation. For example, it is possible that children were engaged in one activity that required low levels of behavioral self-regulation (e.g., coloring) for the entire 8-minute observation, making it difficult for observers to assess global, classroom behavioral self-regulation. Although the HTKS is also a snap-shot of behavioral self-regulation administered during one short period of time, the task is designed to directly measure children's ability to remember instructions, pay attention, and inhibit dominant responses (e.g., respond in the opposite manner) during a challenging game. Placing these cognitive demands on children during the HTKS task allows assessors to observe and record behavioral self-regulation in action, whereas an extramural classroom observation may not. Similarly, preschool teachers are often actively engaged with children during academic activities, which may permit them to tap into and rate the self-regulated behaviors that are related to achievement more accuaratley than extramural observers who typically observe children from a distance.

In the current study, observed behavioral self-regulation was not significantly related to achievement, which is contrary to a prior study that reported significant links between the OCES and reading proficiency (Cameron Ponitz, Rimm-Kaufman, et al., 2009). Cameron Ponitz, Rimm-Kaufman, and colleagues (2009) utilized the OCES, however, in distinct ways that could help explain the differences in results. For example, in the current sample, factor analyses identified one factor representing global behavioral self-regulation comprised of four categories of the OCES (engagement, self-reliance, attention, and compliance), whereas Cameron Ponitz, Rimm-Kaufman, and colleagues used an additional category (disruptive behavior) to measure behavioral engagement in the classroom. Although the disruptive category was not statistically or conceptually relevant to our measure of global behavioral self-regulation, omitting this category could have influenced the utility of the measure in predicting literacy skills as was reported by Cameron Ponitz, Rimm-Kaufman, and colleagues.

In sum, in the current study, the efficacy of teacher-rated, observed, and directly assessed behavioral self-regulation varied depending on the academic domain. Our data suggest that the CBRS, a teacher report, was particularly useful for predicting emergent literacy, and the HTKS, a direct assessment, was better for predicting early math skills in preschool. These findings have implications for researchers when designing studies in which the relation between behavioral self-regulation and academic achievement is of interest. Failing to assess children's behavioral self-regulation with multiple measures could bias study results, and increase Type II error (not detecting effects when effects do in fact exist). These results also have practical implications for teachers, parents, and clinicians when assessing and identifying children at risk for poor behavioral self-regulation. Due to the specificity in measures predicting certain academic domains, utilizing multiple measurements could provide a more comprehensive view of

children's behavior in classroom contexts as it relates to academic success (Kerr, Lunkenheimer, & Olson, 2007), which could in turn, help identify and support children experiencing difficulties with behavioral self-regulation.

Limitations and Future Directions

Although this study adds to existing literature by comparing the predictive utility of adult reports and a direct assessment of behavioral self-regulation, some limitations must be noted. First, the current study was cross-sectional which limits the ability to study individual differences and growth in behavioral self-regulation and achievement over time. Future research should compare the utility of behavioral self-regulation assessments longitudinally by following the same group of children during preschool and beyond to better clarify how the utility of measurements of behavioral self-regulation may change overtime and to better establish causal links.

Second, the OCES was limited to single, 8-minute observations of children during one classroom setting (i.e., free play). Observers may not have had opportunities to see and record certain behaviors, which may partially explain the non-significant links between the OCES and academic domains, as well as between the OCES and the HTKS. Future work should continue to examine the utility of the OCES for measuring behavioral self-regulation by observing the same child over multiple classroom visits and in various classroom contexts to ensure that the most valid and comprehensive view of children's behavior is being captured.

Practical Implications and Conclusions

The current study provides insight into the relationships between three different measures of behavioral self-regulation and academic achievement. Our results suggest that teacher-rated and directly assessed behavioral self-regulation remain valid measurements of children's

behavioral self-regulation in relation to achievement. Domain specificity emerged, however, in that the teacher ratings were better predictors of literacy, and the direct assessment was a stronger predictor of math. Having a greater understanding of the varying utility of assessments for preschoolers has implications for researchers when making decisions about what assessments are most practical and effective when studying specific academic domains, and also, supports the value in utilizing multiple measurements when assessing behavioral self-regulation and relations to early achievement. In addition, in lieu of accumulating evidence indicating the critical nature of behavioral self-regulation for academic success, it is imperative that parents, teachers, and researchers employ measurement tools that allow them to best identify children at risk for poor academic outcomes.

References

Acock, A. C. (2012). What to do about missing values. In H. Cooper (Ed.), *APA handbook of research methods in psychology* (pp. 27–50). Washington, DC: American Psychological Association. doi: 10.1037/13621-002

Adler F., & Lange, G. (1997). Children's mastery orientations and school achievement in the elementary grades. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.

Andersson, H., & Bergman, L. R. (2011). The role of task persistence in young adolescence for successful educational and occupational attainment in middle adulthood. *Developmental Psychology, 47*, 950–960. doi: 10.1037/a0023786

Baumeister, R. F., & Vohs, K. D. (2004). *Handbook of self-regulation: Research, theory, and applications*. New York, NY: Guilford.

Beaman, R., Wheldall, K., & Kemp, C. (2006). Differential teacher attention to boys and girls in the classroom. *Educational Review, 58*, 339–366. doi: 10.1080/00131910600748406

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of child functioning at school entry. *American Psychologist, 57*, 111–127. doi: 10.1037/0003-066X.57.2.111

Blair, C., & Razza, R. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*, 647–663. doi:10.1111/j.1467-8624.2007.01019.x

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28*, 561–571. doi: 10.1207/s15326942dn2802_1

Brandon, K. A., Kehle, T. J., Jenson, W. R., & Clark, E. (1990). Regression, practice, and expectation effects on the Revised Connors Teacher Rating Scale. *Journal of Psychoeducational Assessment, 8,* 456–466. doi:10.1177/073428299000800401

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of `hot' and `cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly. 24,* 337–349. doi: 10.1037/a0015861

Bronson, M. B., Tivnan, T., & Seppanen, P. S. (1995). Relations between teacher and classroom activity variables and the classroom behaviors of preschool children in Chapter 1 funded programs. *Journal of Applied Developmental Psychology, 16,* 253–282. doi: 10.1016/0193-3973(95)90035-7

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33,* 205–228. doi: 10.1080/87565640801982312

Bulotsky-Shearer, R. J., Fantuzzo, J. W., & McDermott, P. A. (2008). An investigation of classroom situational dimensions of emotional and behavioral adjustment and cognitive and social outcomes for Head Start children. *Developmental Psychology, 44,* 139–154. 10.1037/0012-1649.44.1.139

Cameron Ponitz, C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! A direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23,* 141–158. doi:10.1016/j.ecresq.2007.01.004

Cameron Ponitz, C., McClelIand, M. M., Matthews, S. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*, 605–619. doi:10.1037/a0015365

Cameron Ponitz, C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review, 38*, 102–120. Retrieved from http://www.nasponline.org/publications/ spr/index.aspx?vol=42&issue=2

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

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

Cooper, D. H., & Farran, D. C. (1991). *The Cooper-Farran Behavioral Rating Scales*. Brandon, VT: Clinical Psychology Publishing Co., Inc.

Davis-Kean, P. (2005). The influence of parent education and family income on child achievement: The indirect role of parent expectations and the home environment. *Journal of Family Psychology, 18*, 294–304. doi: 10.1037/0893-3200.19.2.294

de Ramirez, R. D., & Shapiro, E. S. (2005). Effects of student ethnicity on judgments of ADHD symptoms among Hispanic and White teachers. *School Psychology Quarterly, 20*, 268–287. doi: 10.1521/scpq.2005.20.3.268

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The

Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interaction. *Early Childhood Research Quarterly, 25*, 1–16. doi:10.1016/j.ecresq.2009.08.004

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

Duncan, R. J., McClelland, M. M., & Acock, A. A. (2013). *Measure of executive function: Convergent validity and links to academic achievement.* Manuscript in Preparation.

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology, 26*, 465–486. doi: 10.1207/s15326942dn2601_6

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). TEST REVIEW Behavior Rating Inventory of Executive Function. *Child Neuropsychology, 6*, 235–238. doi: 10.1076/chin.6.3.235.3152

Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of Everyday Executive Function in Acquired and Developmental Disorders. *Child Neuropsychology, 8*, 121–137. doi: 10.1076/chin.8.2.121.8727

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Halpern, D. F. (2000). *Sex differences in cognition.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hosterman, S. J., DuPaul, G. J., & Jitendra, A. K. (2008). Teacher ratings of ADHD symptoms in ethnic minority students: Bias or behavioral difference? *School Psychology Quarterly, 23*, 418–435. doi: 10.1037/a0012668

Howse, R. B., Calkins, S. D., Anastopoulos, A. D., Keane, S. P., & Shelton, T. L. (2003). Regulatory contributors to children's kindergarten achievement. *Early Education & Development, 14*, 101–120. doi: 10.1207/s15566935eed1401_7

Jaffe, L. E. (2009). Development, interpretation, and application of the W score and the relative proficiency index (Woodcock-Johnson III Assessment Service Bulletin No. 11). Rolling Meadows, IL: Riverside Publishing.

Kerr, D. C. R., Lunkenheimer, E. S., & Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and Psychiatry, 48*, 967–975. doi:10.1111/j.1469-7610.2007.01776.x

Konold, T. R., Jamison, K. R., Stanton-Chapman, T. L., & Rimm-Kaufman, S. E. (2010). Relationships among informant based measures of social skills and student achievement: A longitudinal examination of differential effects by sex. *Applied Developmental Science, 14*, 18–34. doi: 10.1080/10888690903510307

Kontos, S., Burchinal, M., Howes, C., Wisseh, S., & Galinsky, E. (2002). An eco-behavioral approach to examining the contextual effects of early childhood classrooms. *Early Childhood Research Quarterly, 17*, 239–258. doi: 10.1016/s0885-2006(02)00147-3

Layzer, J. L., Goodsen, B. D., & Moss, M. (1993). *Life in preschool: Observational study of early chidhood programs for disadvantaged four-year-olds* (Vol. 1). Cambridge, MA: Abt. Associates.

Lim, S. M., Rodger, S., & Brown, T. (2010). Validation of the Child Behavior Rating Scale in Singapore (part 1): Rasch analysis. *Hong Kong Journal of Occupational Therapy, 20,* 52–62. doi: http://dx.doi.org/10.1016/S1569-1861(11)70004-3

Magnuson, K. A., Sexton, H. R., Davis-Kean, P. E., & Huston, A. C. (2009). Increases in maternal education and young children's language skills. *Merrill-Palmer Quarterly: Journal of Developmental Psychology, 55,* 319–350.

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. doi: http://dx.doi.org/10.1016/S0887-6177(01)00168-8

Matthews, J. S., Cameron Ponitz, C. & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology, 101,* 689–704. doi: http://dx.doi.org/10.1016/S1569-1861(11)70004-3

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21,* 471–490. doi: 10.1016/j.ecresq.2006.09.003

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly, 28,* 314-324. doi: http://dx.doi.org/10.1016/j.ecresq.2012.07.008

McClelland, M. M., Cameron Ponitz, C., Messersmith, E., & Tominey, S. (2010). Self-regulation: The integration of cognition and emotion. In W. Overton (Vol. Ed.) & R. Lerner (Ed.), *Handbook of life-span human development.* (Vol. 1: Cognition, biology and methods, pp. 509–553). Hoboken, NJ: Wiley and Sons.

McClelland, M. M., & Cameron Ponitz, C. (2012). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically-valid measures. *Child Development Perspectives, 6,* 136–142. doi: 10.1111/j.1750-8606.2011.00191.x

McClelland, M. M., Connor, C. M., Jewkes, A. M., Cameron, C. E., Farris, C. L., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43,* 947–959. doi: 10.1037/0012-1649.43.4.947

McClelland, M. M., & Morrison, F. J. (2003). The emergence of learning-related social skills in preschool children. *Early Childhood Research Quarterly, 18,* 206–224. doi: 10.1016/s0885-2006(03)00026-7

McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: the role of learning-related social skills. *Early Childhood Research Quarterly, 15,* 307–329. doi: 10.1016/s0885-2006(00)00069-7

McClelland, M. M., & Wanless, S. B. (2012). Growing up with assets and risks: The importance self-regulation for academic achievement. *Research in Human Development, 9,* 278–297. doi: 10.1080/15427609.2012.729907

Moffitt, T. E., Arseneaulty, L., Belsky, D., Dickson, N., Hancox, R. J., …Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academic of Sciences, 108,* 2693–2698. doi: 10.1073/pnas/1010076108

National Institute of Child Health and Human Development Early Child Care Research Network. (2005). A day in third grade: A large-scale studey of classroom quality and teacher and student behavior. *The Elementary School Journal, 105,* 305–323. doi: 10.1086/428746

Obradovic, J. (2010). Effortful control and adaptive functioning of homeless children: Variable and person-focused analyses. *Journal of Applied Developmental Psychology, 31,* 109–117. doi: 10.1016/j.appdev.2009.09.004

Phillips, B. M., & Lonigan, C. J. (2010). Child and informant influences on behavioral ratings of preschool children. *Psychology in the Schools, 47*, 374–390. doi: 10.1002/pits.20476

Ramini, G. B., Brownell, C. A., & Campbell, S. B. (2010). Positive and negative peer interaction in 3- and 4-year olds in relation to regulation and dysregulation. *Journal of Genetic Pscyhology, 171*, 218–250. doi:10.1080/00221320903300353

Rimm-Kaufman, S. E. (2005). *Children's Early Learning Study Implementation Manual.* Unpublished manuscript, Curry School of Education, University of Virginia, Charlottesville, Virginia.

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology, 45*, 958–972. doi: 10.1037/a0015861

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in the kindergarten classroom. *Elementary School Journal, 105*, 377–394. doi: 10.1086/429948

Rothbart, M. K., & Posner, M. I. (2005). Genes and experience in the development of executive attention and effortful control. In L.A. Jensen & R.W. Larson (Eds.), *New Horizons in Developmental Theory and Research* (pp.101–108). San Francisco, CA: Jossey-Bass.

Rueda, M. R., Posner, M. I. & Rothbart, M. K. (2004). Attentional control and self-regulation. In R.F. Baumeister & K.D. Vohs (Eds), *Handbook of Self Regulation: Research, Theory, and Applications* (pp. 283–300). New York, NY: Guilford Press.

Sattler, J. M. (1988). *Assessment of children's intelligence and special abilities* (2nd ed.). San Diego, CA: Allyn & Bacon.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the start of the art. *Psychological Methods, 7,* 147–177. doi: 10.1037/1082-989x.7.2.147

Schrank, F. A., McGrew, K. S., Ruef, M. L., Alvarado, C. G., Muñoz-Sandoval, A. F., & Woodcock, R. W. (2005). *Overview and technical supplement (Bateria III Woodcock-Muñoz Assessment Service Bulletin No. 1).* Itasca, IL: Riverside Publishing.

Schunk, D. H. & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32,* 195–208. doi: 10.1207/s15326985ep3204_1

Serpell, Z. N., & Mashburn, A. J. (2012). Family-school connectedness and children's early social development. *Social Development, 21,* 21–46. doi: 10.1111/j.1467-9507.2011.00623.x

Sektnan, M., McClelland, M. M., Acock, A., & Morrison, F. J. (2010). Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Childhood Research Quarterly, 25,* 464–479. doi: 10.1016/j.ecresq.2010.02.005

Shonkoff, J. P., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

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

StataCorp. (2011). *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.

Suchodoletz, A. v., Gestsdottir, S., Wanless, S. B., McClelland, M. M., Birgisdottir F., Gunzenhauser, C., & Ragnarsdóttir, H. (2013). Behavioral self-regulation and relations to emergent academic skills among children in Germany and Iceland. *Early Childhood Research Quarterly, 28,* 62–73. doi: http://dx.doi.org/10.1016/j.ecresq.2012.05.003

Walker, H. M., Golly, A. M., McClane, J. Z., & Kimmich, M. (2005). The Oregon First Step to Success replication initiative: Statewide results of an evaluation of the program's impact. *Journal of Emotional and Behavioral Disorders, 13*, 163–172. doi: 10.1177/10634266050130030401

Walker, H. M., Seeley, J. R., Small, J., Severson, H. H., Graham, B. A., Feil, E. G., . . . Forness, S. R. (2009). A randomized controlled trial of the First Step to Success early intervention: Demonstration of program efficacy outcomes in a diverse, urban school district. *Journal of Emotional and Behavioral Disorders, 17*, 197–212. doi:10.1177/1063426609341645

Walker, H. M., & Severson, H. (1990). *Systematic Screening for Behavior Disorders (SSBD).* Longmont, CO: Sopris West.

Wanless, S. B., McClelland, M. M., Acock, A. C., Chen, F., & Chen, J. (2011). Behavioral regulation and early academic achievement in Taiwan. *Early Education & Development, 22*, 1–28. doi: 10.1080/10409280903493306

Wanless, S. B., McClelland, M. M., Tominey, S. L., & Acock, A. C. (2011). The influence of demographic risk factors on children's behavioral regulation in prekindergarten and kindergarten. *Early Education & Development, 22*, 461–488. doi: 10.1080/10409289.2011.536132

Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2012). The matter of assessor variance in early childhood education—Or whose score is it anyway? *Early Childhood Research Quarterly, 27*, 46–54. doi: 10.1016/j.ecresq.2011.06.003.

Whitebread, D., Coltman, P., Pasternak, D., Sangster, C., Grau, V., Bingham, S., . . .Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition & Learning, 4*, 63–85. doi: 10.1007/s11409-008-9033-1.

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment, 24*, 226–239. doi: 10.1037/a0025361

Woodcock, R. W., & Mather, N. (2000). Woodcock Johnson Psycho-Educational Battery–III. Itasca, IL: Riverside.

Table 1

Descriptive Statistics: Demographic Variables, Child Self-Regulation, and Achievement Outcomes in Preschool (*N* = 247*)*

Variables	*M*	*SD*	Range
Covariates			
Age (in months)	61.01	4.30	41.52 - 70.38
Parent education (in years)	14.66	3.97	0 - 26
Gender (male = 1)	% Female	% Male	0 - 1
	50	50	
Predictors			
Teacher report (CBRS)	39.18	7.04	18 - 50
Observer report (OCES)	5.89	1.03	2.50 - 7
Direct assessment (HTKS)	23.03	18.61	0 - 59
Outcomes			
Applied problems	419.35	22.81	332 - 481
Letter-word identification	350.95	27.11	276 - 453

Note. CBRS = Child Behavior Rating Scale; OCES = Observational Child Engagement Scale; HTKS = Head-Toes-Knees-Shoulders task.

Table 2

Correlations Among Study Variables with Academic Outcomes (*N* = 247)

Variables	1	2	3	4	5	6	7
1. Applied problems	–						
2. Letter-word identification	.58**	–					
3. Teacher report (CBRS)	.40**	.41**	–				
4. Observer report (OCES)	.21*	.06	.27**	–			
5. Direct assessment (HTKS)	.64**	.43**	.35**	.13	–		
6. Age (in months)	.26**	.17**	.12†	.08	.20**	–	
7. Gender (male = 1)	.02	.02	-.24**	-.01	.03	.07	–
8. Parent education	.48**	.42**	.20**	.12	.49**	.06	.07

Note. CBRS = Child Behavior Rating Scale; OCES = Observational Child Engagement Scale; HTKS = Head-Toes-Knees-Shoulders task.
†$p < .10$, *$p < .05$, **$p < .01$

Table 3

Teacher-Rated, Observed, and Direct Assessments of Behavioral Self-Regulation Predicting Academic Outcomes (*N* = 247)

	Applied Problems			Letter-Word Identification		
	B	SE B	β	B	SE B	β
Teacher report (CBRS)	.67	.19	.21**	1.22	.25	.32**
Observer report (OCES)	.35	1.91	.02	-2.07	2.27	-.08
Direct assessment (HTKS)	.53	.07	.43**	.27	.09	.18*
Age (in months)	.64	.26	.12*	.40	.35	.06
Gender (male = 1)	1.16	2.25	.03	3.22	3.01	.06
Parent education	1.15	.35	.20*	1.94	.42	.28**

Note. CBRS = Child Behavior Rating Scale; OCES = Observational Child Engagement Scale; HTKS = Head-Toes-Knees-Shoulders task.
*$p < .05$, **$p < .01$