Early Gender Differences in Self-Regulation and Academic Achievement

Article *in* Journal of Educational Psychology · August 2009

DOI: 10.1037/a0014240

3 authors:

Jamaal.S. Matthews
Montclair State University
10 PUBLICATIONS **400** CITATIONS
SEE PROFILE

Claire Elizabeth Cameron
University at Buffalo, The State University o…
48 PUBLICATIONS **1,973** CITATIONS
SEE PROFILE

Frederick J Morrison
University of Michigan
113 PUBLICATIONS **5,736** CITATIONS
SEE PROFILE

Journal of Educational Psychology
2009, Vol. 101, No. 3, 689–704

0022-0663/09/$12.00 DOI: 10.1037/a0014240

Early Gender Differences in Self-Regulation and Academic Achievement

J. S. Matthews
University of Michigan

Claire Cameron Ponitz
University of Virginia

Frederick J. Morrison
University of Michigan

This study examined gender differences in self-regulation in the fall and spring of kindergarten and their connection to gender differences in 5 areas of early achievement: applied problems (math), general knowledge, letter–word identification, expressive vocabulary, and sound awareness. Behavioral self-regulation was measured using both an objective direct measure ($N = 268$; Head-Toes-Knees-Shoulders task) and, for a subsample of children, a teacher report of classroom self-regulatory behavior ($n = 156$; Child Behavior Rating Scale). Results showed that girls outperformed boys in both assessments. Although gender differences in self-regulation were clear, no significant gender differences were found on the 5 academic achievement outcomes, as measured by the Woodcock–Johnson III Tests of Achievement. Self-regulation consistently predicted math and sound awareness, although links were stronger with the direct measure as compared with teacher reports. Implications for understanding the role of gender and self-regulation in early and later academic achievement and the role of self-regulation in particular areas of achievement are discussed.

Keywords: gender differences, behavioral self-regulation, academic achievement, kindergarten, structured observational assessment

Mounting evidence points to a growing gender gap in classroom functioning and academic achievement. Historically, boys have been largely advantaged in the classroom and most academic settings (Weaver-Hightower, 2003). In the mid 1990s, research regarding gender equity in education highlighted the difficulty and educational neglect many girls experienced in American classrooms. Overlooked by teachers, recipients of biased pedagogy and testing, and dealing with poor academic self-concepts, girls were seen as educationally stifled in an environment more suited to male intellectual development (American Association of University Women, 1992; Eisenhart & Finkel, 1998; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Sadker & Sadker, 1994; Weaver-Hightower, 2003).

The tide of male preeminence in the classroom and academic achievement has begun to turn in recent years. With the exception of some advanced mathematics and science courses, girls have narrowed, closed, or overcome many academic gaps that previously favored boys, whereas other long-standing gaps that favored girls have widened. Research today on gender and education in kindergarten through 12th grade school settings reveals that girls tend to build stronger relationships with teachers, attain higher grades, achieve higher levels of education, and progress better scholastically overall than boys (Birch & Gary, 1998; Duckworth & Seligman, 2006; Silverman, 2003). These findings are evident even after controlling for contributions of students' backgrounds, including maternal education, an important indicator of socioeconomic status and a proxy for children's early home learning experiences (Morrison & Cooney, 2002; Sameroff, Bartko, Baldwin, Baldwin, & Seifer, 1998). In addition to academic achievement, girls now also predominate in academic and social groups, such as debate teams, honor societies, student government, and a variety of other groups, with the exception of sports (Sommers, 2000).

In contrast, boys are more likely candidates for expulsion, suspension, and dropping out. Nationally, 42% of boys have been suspended from school at least once by age 17, compared with 24% of girls (Office of Juvenile Justice and Delinquency Prevention, 2006). Further, boys are four times more likely than girls to be referred for remedial and special education services (Brook & Boaz, 2005; Flynn & Rahbar, 1994). The National Center for Education Statistics has shown that nearly twice as many boys than

J. S. Matthews, The Combined Program in Education and Psychology, University of Michigan; Claire Cameron Ponitz, Social Development Lab, University of Virginia; Frederick J. Morrison, Department of Psychology, University of Michigan.

This study was funded by the National Institute of Child and Human Development and the National Science Foundation Grants R01 HD27176 and 0111754, respectively, awarded to Frederick J. Morrison. This research was also funded by a training fellowship from the Institute of Education Science, U.S. Department of Education Award R305 B040049, to Claire Cameron Ponitz. The opinions expressed are ours and do not represent the views of the funding agencies.

We thank the Pathways to Literacy Project members and staff for their important contributions and extend our sincere appreciation to the schools, teachers, and families that made this study possible. Thanks are also due to Megan McClelland and Andrew Mashburn for their helpful comments in preparing the article.

Correspondence concerning this article should be addressed to J. S. Matthews, Department of Psychology, University of Michigan, 2229 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043. E-mail: jamaalm@umich.edu

girls between the ages of 5 and 12 are held back at least one grade (Freeman, 2004).

Despite these statistics, conflicting findings regarding gender differences in achievement have hampered our understanding of the issue. The recent "boy crisis" has been contested as an exaggeration of what some call minor gender differences (Mead, 2006). In her examination of National Assessment of Educational Progress data in recent years, Mead (2006) commented that gender differences in achievement and the boy crisis are a matter of perspective on equivocal findings in national data. She found that girls sometimes exhibit an advantage in some academic achievement domains; however, the advantage tends to be small in most normative populations (i.e., middle-class, Caucasian children). Further, she argued that boys are not really falling behind; it is simply the case that girls are doing better than ever (Mead, 2006).

In addition, other researchers have shown few or no gender differences in the early school years in many subject matter tests but have found that gender contributes to achievement disparity as children move into their middle school years, with differences fully manifesting in high school and beyond (Entwisle, Alexander, & Olson, 1997; Hyde, Fennema, & Lamon, 1990; Willingham & Cole, 1997). On the contrary, others have shown consistent elementary-level gender differences favoring girls, especially in literacy achievement (Coley, 2001; Gambell & Hunter, 1999; Lummis & Stevenson, 1990; Ready, LoGerfo, Burkham, & Lee, 2005). Finally, in a middle school sample, Duckworth and Seligman (2006) found a significant gender gap favoring girls in classroom grades in English, math, and social studies but a nonexistent or small gap favoring boys on standardized achievement and IQ tests. This conflicting evidence raises questions about the true state of gender differences in achievement that necessitates greater scrutiny and directs us to examine possible factors that may help illuminate confusion in the research literature. In the current study, we explore gender differences in academic achievement and also consider children's behavioral skills in school settings as a potential predictor of these differences.

In Search of an Explanation

Debate has surrounded potential contributing factors to achievement disparities between boys and girls. Some distal explanations include a change in societal expectations for females as professionals, classroom structures that suit the talents of girls, a female advantage in self-discipline and interpersonal skills, inherent bias in teacher expectations and curricula, and childhood socialization factors at home (Cooper & Speece, 1990; Duckworth & Seligman, 2006; Eccles, 1994; Gambell & Hunter, 1999; Silverman, 2003; Sommers, 2000). Although a number of these factors undoubtedly play a contributing role, increasing attention is being paid to self-regulation as a proximal factor that consistently predicts educational experiences and outcomes in early childhood and ultimately leads to differences in achievement. Self-regulatory skills, which help children direct and control their attention and behavior, are crucial for successful school performance and adaptation (Blair, 2002).

The literature on self-regulation is diverse in its conceptualization of the term. Our working definition of self-regulation in early childhood includes behavioral self-regulation, which depends on cognitive skills, including working memory, attention control and switching, and inhibitory control (Barkley, 1997; Bronson, 2000; McClelland, Cameron, Wanless, & Murray, 2007). Behavioral regulation requires children to integrate these multiple component skills and form behavioral responses, such as remembering a classroom rule to raise their hand before participating. Behavioral regulation is distinct from emotion regulation, or the regulation of emotional responses to stimuli; emotion regulation is also important for social and educational outcomes but is not a focus here (Eisenberg, Spinrad, & Smith, 2004). Behavioral regulation includes children's ability to remember directives, as well as monitor, inhibit, and direct their attention and behavior (Gathercole & Pickering, 2000; Rueda, Posner, & Rothbart, 2005).

Successfully regulating one's behavior is associated with executive function, a primarily cognitive construct (Shonkoff & Phillips, 2000; Zelazo, Müller, Frye, & Marcovitch, 2003). Another term includes cognitive regulation, defined as "the regulation of attention and selective strategy use in the execution of cognitive tasks" (Blair, 2002, p. 112). Our focus on children's overt behavior, including in their gross motor responses and actions in the classroom, leads to our label of behavioral self-regulation (Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003). In this study, we highlight task-related behaviors; however, we acknowledge that successful classroom behavior also includes social skills, such as controlling aggression and getting along with peers (Ladd, Birch, & Buhs, 1999; McClelland et al., 2007). However, growing evidence suggests that behavioral self-regulation and its underlying cognitive skills, including attention, is a stronger predictor of multiple areas of achievement than are social skills (Duncan et al., 2007; Howse et al., 2003). One argument is that strong behavioral regulation early in the school trajectory sets the stage for academic success by predicting increased school engagement and motivation and children's adoption of positive learning strategies (Fredricks, Blumenfeld, & Paris, 2004; Zimmerman & Schunk, 2001).

Thus, early attentional and behavioral regulatory skills have been argued to help construct a foundation for successful classroom functioning, making room for opportunities necessary for optimal academic performance throughout all levels of school (Howse et al., 2003; Kuhl & Kraska, 1989). Consistent with this claim, strong behavioral self-regulation in the fall of kindergarten predicts higher year-end achievement in multiple areas (Howse et al., 2003; McClelland, Morrison, & Holmes, 2000). In contrast, children with difficulty regulating their attention have been shown to exhibit low achievement (Howse, Lange, Farran, & Boyles, 2003).

Research details the prevalence of poor self-regulatory skills and their eventual impact on achievement. Rimm-Kaufman, Pianta, and Cox (2000) reported 46% of teachers to rate nearly half of children entering kindergarten as lacking in basic classroom competencies to achieve in the school environment, such as following directions and taking turns. Kindergarten teachers in another investigation identified one out of six children as having poor learning-related skills, including behavioral self-regulation (McClelland et al., 2000). In a follow-up study, kindergarten teacher ratings of behavioral self-regulation predicted trajectories of literacy and math skills through second grade and levels of literacy and math in sixth grade (McClelland, Acock, & Morrison, 2006). Moreover, the gap in achievement for children with weak versus strong skills widened between kindergarten and second grade. Other researchers have also found early learning and regulatory

skills to positively predict academic achievement 2 to 4 years later, heightening the importance of identifying these skills, predictors, and links to achievement in kindergarten (Alexander, Entwisle, & Dauber, 1993). In the present study, behavioral self-regulation is examined as a possible source of gender discrepancies in academic achievement.

Self-Regulation: Gender Differences and Measurement

Recently, investigators have reported gender differences in self-regulation in the early school years (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006; Kochanska, Coy, & Murray, 2001; McClelland et al., 2000; Ponitz et al., 2008; Ready et al., 2005; Silverman, 2003). Despite strong evidence connecting self-regulation with achievement, little of this research has incorporated gender as a central focus or expanded to look beyond literacy and language outcomes (Ready et al., 2005). Differing methodologies may, in part, contribute to conflicting results on these issues. Teacher and parent reports are traditional methods of measuring self-regulation and self-control in the early school years (Blair, 2003; Bronson, Tivnan, & Seppanen, 1995; Duncan et al., 2007; Howse, Lange, et al., 2003; Rothbart, Ahadi, Hersey, & Fisher, 2001). However, these measures may be susceptible to observer bias (Rothbart et al., 2001). For example, controlling for other measures of performance, teachers have been shown to rate boys lower than girls on achievement and behavior (Beaman, Wheldall, & Kemp, 2006; Ferguson, Lloyd, & Horwood, 1991). In addition, Duckworth and Seligman (2006) found discrepancies in the size and significance of gender differences in self-control (i.e., delay of gratification) across different measures in their study (i.e., an objective measure vs. parent and teacher reports). Thus, multiple measures of behavior, both observer and direct, may best illuminate how gender differences in behavioral regulation relate to achievement. In addition, direct measures provide a more proximal assessment of behavior, in a similar context as achievement measures are administered.

Recently, a number of objective measures of self-regulation have been developed (Diamond, Kirkham, & Amso, 2002; Dowsett & Livesey, 2000; Hughes, Dunn, & White, 1998; Kochanska et al., 2001; Korkman, Kemp, & Kirk, 2001; McCabe, Rebello-Britto, Hernandez, & Brooks-Gunn, 2004). To varying degrees, self-regulation tasks tap working memory, attention, and inhibitory control, which work separately and in conjunction to support self-directed classroom behavior (Blair, 2002; McClelland et al., 2007; Ponitz et al., 2008). Working memory refers to the cognitive process involved in maintaining and manipulating information and predicts achievement in 7-year-olds (Gathercole & Pickering, 2000). Attention refers to the ability to focus on information and ignore distraction; kindergarten levels have been linked with high school graduation rates (Vitaro, Brendgen, Larose, & Tremblay, 2005). Inhibitory control enables children to modify their automatic responses by responding more adaptively, such as inhibiting impulses and initiating a new, unnatural response. Prekindergarteners with strong inhibitory control have higher kindergarten achievement in both math and reading (Blair & Razza, 2007).

Evidence suggests that scores from the recently developed Head-to-Toes task (McClelland et al., 2007; Ponitz et al., 2008) are efficient, reliable, and valid for measuring overall behavioral self-regulation that predicts academic achievement in preschool. Further, the measure has also shown small but statistically significant gender differences in behavioral self-regulation in children aged 3 to 6 years (Ponitz et al., 2008). However, there is a ceiling effect, making this task most appropriate for 4- and 5-year-olds. In this study, we use an extended, more complex version developed for kindergarteners. Currently, few overt measures of behavioral regulation extend to this age group, and data from samples larger than 50 children are limited (Carlson, 2005; Rothbart, Posner, & Kieras, 2006). Similar to *Simon Says* and other games common in early childhood settings, the task requires children to integrate three cognitive skills (working memory, attention, and inhibitory control) and apply them to their overt behavior. In this structured observational task, children must attend to the researcher, remember two and then four rules, and respond with a conflicting behavioral response (i.e., they must touch their head when asked to touch their toes). The addition of two additional rules, to require children to remember a total of four different rules, is a unique modification to currently available behavioral regulation measures and makes the Head-Toes-Knees-Shoulders task (HTKS) appropriate for children older than 5 years (Diamond et al., 2002).

The Present Study

In the kindergarten year, a considerable degree of emphasis is placed on children acquiring the behavioral skills necessary to facilitate the learning process (Entwisle & Alexander, 1998; Piotrkowski, Botsko, & Matthews, 2000). The development of skills within this year can have implications throughout elementary and into middle school achievement, beyond the contributions of family background factors such as maternal education (McClelland et al., 2006). In general, research also suggests gender differences in self-regulation emerging in early childhood but gender differences in achievement emerging later in the school trajectory. At least one set of findings has connected regulatory differences to gender differences in literacy (Ready et al., 2005). Ready et al.'s (2005) study, however, considered only literacy skills and relied on teacher reports of self-regulation. It is not yet clear how gender differences on an observational measure of self-regulation correspond with teacher assessments of behavior or whether self-regulation and gender predict multiple aspects of early achievement, including mathematics and general academic knowledge. The present study extends previous work in these ways.

This study seeks to (a) clarify whether gender differences in self-regulation appear in kindergarten in two types of assessments, (b) assess the size and effect of these disparities, and (c) examine whether gender and self-regulation predict achievement at the end of the kindergarten year in several important academic domains. We examined three research questions.

First, how does a direct measure of behavioral self-regulation compare with teacher ratings of child behavior? We expected both assessments to yield variability and to correspond with one another, with children scoring higher on the direct assessment receiving better ratings from their teachers regarding classroom behavioral regulation.

Second, what is the extent of gender differences in self-regulation on teacher ratings in the spring of kindergarten and on a direct objective measure in the fall and spring of kindergarten? It was hypothesized that girls would obtain higher scores than boys on both assessments.

Third, controlling for background variables, do gender and self-regulation at kindergarten entry predict achievement gains over the kindergarten year? Further, are gender differences in achievement comparable with those for self-regulation? Girls and children with stronger self-regulation at kindergarten entry were expected to demonstrate greater gains compared with boys and children with weaker self-regulation. Effect sizes for gender differences in achievement were not expected to be as large as those for gender differences in self-regulation. All analyses controlled for maternal education, age, and fall achievement.

Method

Participants

Children, families, and teachers were part of a 5-year longitudinal study investigating the nature and sources of cognitive and social development during the transition from preschool to elementary school. First-time kindergarteners ($N = 268$) in southeast Michigan participated after entering the study through their preschool programs. As 3- and 4-year-olds, children entering preschool, housed within one public school district, were recruited through fall orientations and backpack mailing at the district's six participating schools during the first 2 years of the larger study. Recruiting was stopped after the target sample size had been achieved. Recruitment efforts enrolled approximately 38% of the districts' entering preschoolers in the longitudinal study, with a final sample size for the present investigation of 268 children in their kindergarten year. Current study participants were enrolled in 87 classrooms in 12 schools; parents of child participants and teachers provided written informed consent prior to participation. Parents and teachers received a $20 gift card each year they participated.

For this investigation, children's mean age was 5.45 years ($SD = 0.33$). The sample had comparable numbers of girls ($n = 139$) and boys ($n = 129$). Of all the children who participated, we obtained ethnicity data for 84%. Of these, 83.1% were White/Caucasian, 7.5% were Asian or Asian American, 6.1% were African American, and 1.3% were Hispanic. The final 2% were biracial children representing a variety of ethnicities. The mean parent education level for reporting families was 16 years ($SD = 1.8$), and nearly 40% of parents held a master's degree.

Measures

Child Behavior Rating Scale (CBRS; Bronson et al., 1995)

The specific focus of our study was to assess children's self-regulatory (goal-oriented) behaviors for academic outcomes. The CBRS (Bronson et al., 1995) is based on the Bronson Social Task and Skill Profile (Bronson, 1991), an observational instrument designed to assess children's classroom goal-oriented behaviors and strategies used to regulate behavior in academic and social situations; the measure also contains items on social relations with peers and adults, though this was not a present focus (Bronson, 1994; Schunk, 2001). A principal-components analysis was conducted to identify the items within this measure that assess the academic self-regulatory behaviors of interest and create one factor for further analysis. Sample items include, "Completes learning tasks in an organized way," "Observes rules and follows directions without requiring repeated reminders," and "Sees own errors on task and corrects them." Items were rated on a 5-point Likert scale from 1 (*never*) to 5 (*always*).

Test–retest reliability was not possible in this study because teacher ratings were only sampled at one time point; however, previous work has shown CBRS scores to have a test–retest reliability of .67 and an internal consistency (Cronbach's alpha) of .96 (Abt Associates, 1988). Other research on the CBRS shows its scores to have strong construct as well as concurrent validity through its strong factor structure, which accounted for a large proportion of the variance, and its correlation ($r = .43$) with the Bronson Social Task and Skill Profile, an observational measure of goal-oriented in classroom settings (e.g., planning, organizing, mastery, interacting cooperatively; Goodwin & Driscoll, 1980; Seppanen, Godin, Metzger, Bronson, & Cichon, 1993). Furthermore, The CBRS has been used for Head Start and Giant Step evaluations as well as a number of other individual preschool programs (Abt Associates, 1988; Bronson et al., 1995; Meleen, Love, & Nauta, 1988). This measure asks teachers to rate children on practical behaviors, such as following instructions and completing and persisting on classroom tasks. These behaviors require regulating responses based on cognitive skills, including remembering instructions (connected to working memory), focusing on the task at hand (attention), and completing one task before moving onto another (inhibitory control).

Direct Assessment of Self-Regulation

The direct assessment of self-regulation, the HTKS task, is a more complex version of the Head-to-Toes task reported by Ponitz et al. (2008) and based on the Head-and-Feet task (McCabe et al., 2004). Study results indicate scores from the Head-to-Toes task demonstrate strong interrater reliability and construct and predictive validity in early childhood (Ponitz et al., 2008; McClelland et al., 2007). For example, Head-to-Toes task performance correlated positively ($r = .29$ to $.42$) with CBRS teacher ratings, and as we would expect with a task capturing a developmental construct, in a growth curve analysis, chronological age explained the most variance in children's scores relative to background and demographic variables. Consistency among overall scores obtained by different examiners was 100%. In addition, in another study, children who performed better on the task had higher preschool levels of spring achievement in emergent literacy, vocabulary, and mathematics, and those who made greater gains on the task made greater gains as well, with effect sizes ranging from $d = 0.09$ to 0.15 (McClelland et al., 2007). However, this is the first study of the more complex, four-rule assessment, the HTKS.

In the original Head-to-Toes task, children first follow one of two commands naturally and then are instructed to respond with a conflicting, nonautomatic action. For example, if the administrator says, "Touch your head," then the correct response would be to touch one's toes. Similarly, if the administrator says "Touch your toes," then the correct response is to touch one's head. The HTKS increases task complexity by adding two additional commands, including "Touch your shoulders" and "Touch your knees." After the task is explained and children are given chances to practice, they are given the first part of the task (first 10 items), with two possible commands. Then the two additional commands are added,

with a total of four commands requiring a deliberate, unnatural response (second 10 items). This task requires children to use multiple cognitive skills—to remember the rules and attend to the commands—and apply them to their overt behavior by inhibiting the tendency to follow the command as given and instead give the opposite, conflicting response. These demands may be similar to those in classrooms, when children need to follow multiple instructions and finish one project before starting another or remember to raise their hand before participating. Assessing the correspondence between this structured observational task and teacher ratings of behaviors in the classroom context was one goal of the present study.

On each item, children were scored 2 points for responding correctly, 1 point for self-correcting (initial movement to the incorrect response but ending with the correct respond), and 0 points for responding incorrectly. There were 20 items, with a scoring range from 0–40. Children were tested once in the fall and once in the spring.

Measures of Achievement

Subtests of the Woodcock–Johnson III Tests of Cognitive Abilities (WJ-III; Woodcock, McGrew, & Mather, 2001) were used to measure applied math skills, general academic knowledge, literacy, vocabulary, and sound awareness. *W* scores were used in analyses for all achievement tests based on a centered *W* score of 500, which is the average achievement level for a 10-year-old (Mather & Woodcock, 2001). *W* scores have properties similar to those of the Rasch ability scale, including equal-interval measurement characteristics, which make them suited to analyses of achievement gains over time. All subtests include items increasing in difficulty and were administered until children reached a ceiling level (e.g., six items incorrect in a row). Construct and concurrent validity is sufficient for all 22 of the WJ-III achievement measure scores (McGrew & Woodcock, 2001; Woodcock et al., 2001). For this study, we used the following five measures.

Applied problems (math). The Applied Problems subtest includes orally administered word problems accompanied by pictures; as the test progresses in difficulty, children are provided with a paper and pencil. To solve the problems, the child must listen to the item, recognize the procedure to be followed, and perform the appropriate calculations before responding. Applied Problems scores have a median internal reliability of .92 among 5- and 6-year-olds (Mather & Woodcock, 2001).

General knowledge. The subtest of Academic Knowledge includes questions presented orally, with accompanying images, and asks children to demonstrate factual knowledge of science-related information (e.g., "In what ways is winter weather different from summer?"), social studies (e.g., "Who is the person that delivers the mail?"), and humanities (e.g., "Do you know this story?" about a picture of Little Red Riding Hood). Scores have an internal reliability of .88 for 5- and 6-year-olds (Mather & Woodcock, 2001).

Vocabulary. The Picture Vocabulary subtest measures oral language development and lexical knowledge. It is an expressive language task at the single-word level. This subtest requires participants to name objects displayed on a page. The median internal reliability is .77 for 5- and 6-year-olds (Mather & Woodcock, 2001).

Literacy skills. The Letter–Word Identification subtest requires children to name letters and read words aloud from a list. Children were asked to read actual words with fluent pronunciation, which increased in difficulty as children progressed. In previous research (Mather & Woodcock, 2001), scores from this measure have shown strong internal consistency reliability of .98 for 4- and 6-year-olds.

Sound awareness. The Sound Awareness subtest tests the rhyming of real words, deletion and substitution at the syllabic and subsyllabic levels, and reversal of syllables to form new words. It is largely a measure of phonological awareness. Sound Awareness scores on this measure have an internal reliability of .81 for 5- and 6-year-olds (Mather & Woodcock, 2001).

Procedure

Data were collected from three sources: parents, teachers, and children themselves. First, each summer following the first 2 years of the study, all families were contacted to schedule a home visit, where they received questionnaires and stamped envelopes to return to our research lab. One questionnaire included the background questionnaire seeking sociodemographic information on the children. Families who did not schedule a home visit were mailed the questionnaires. A total of 167 families (62%) responded. Second, in the spring, teachers completed the 32-item CBRS (Bronson, Goodson, Layzer, & Love, 1990; Bronson & Love, 1987) for each participating child in their kindergarten classroom. From the full sample, 156 surveys were returned by teachers (58%; 80 girls and 76 boys). Thus, 112 children in the full sample of 268 were unrated by their teachers. Third, in the fall and spring of the kindergarten year, child participants were administered the assessments of self-regulation and academic achievement in two individual sessions lasting 30–40 min. Children were given stickers as they progressed through the battery.

Analytic Strategy

In addition to children who were missing parent background questionnaire and CBRS teacher rating data, at any given time point, achievement and HTKS data were missing for 1–5% of children. Table 1 shows the data available by gender for each variable in the study. Listwise deletion would have cut our sample size to only 86 children, reducing our statistical power and producing biased parameter estimates (Acock, 2005). Our approach to missing data had two goals. First, we wanted to maximize data for the analyses predicting achievement. Second, we wanted to ensure our estimates of missing data were accurate and that we did not extend beyond the available data in our analyses and interpretations.

With regard to the first goal, we used a single imputation procedure. Missing values for the entire sample of 268 children were imputed using the expectation maximization algorithm in SPSS Version 15.0. Expectation maximization is a maximum likelihood single imputation procedure using all available variables in the data set on participants to iteratively calculate values for missing data (Acock, 2005). With imputation, it is important to include all available variables that may be theoretically or statistically linked to the values being imputed to improve the precision of the estimates of any missing values (Graham, Cumsille, &

Table 1
Descriptive Statistics for Raw Data by Gender and Imputed Data for Entire Sample

	Raw data				Imputed data	
	M (*SD*)		*N*			
Variable	Girls	Boys	Girls	Boys	*M* (*SD*)	*N*
T1 child age	5.41 (0.33)	5.49 (0.34)	136	128	5.45 (0.33)	268
T1 maternal education	16.12 (1.71)	16.33 (1.99)	86	81	16.10 (1.56)	268
T1 HTKS[a]	29.41 (8.34)	26.10 (10.08)	135	127	27.82 (9.26)	268
T2 HTKS[b]	33.81 (5.44)	30.99 (9.09)	134	122	32.31 (7.44)	268
T2 CBRS self-regulation[c]	4.10 (0.67)	3.78 (0.74)	80	76	3.98 (0.72)	156
T1 Applied Problems	432.92 (12.73)	434.59 (15.75)	138	129	433.68 (14.26)	268
T2 Applied Problems	445.09 (13.12)	447.35 (16.38)	137	125	455.98 (14.78)	268
T1 Academic Knowledge	466.67 (11.27)	464.49 (13.98)	138	129	465.57 (12.67)	268
T2 Academic Knowledge	475.40 (11.65)	472.82 (13.02)	137	125	474.05 (12.34)	268
T1 Letter–Word ID	374.38 (23.98)	375.05 (32.58)	138	129	374.64 (28.37)	268
T2 Letter–Word ID	405.21 (26.58)	404.82 (32.99)	137	125	404.69 (29.57)	268
T1 Picture Vocabulary	477.51 (9.86)	479.69 (12.93)	138	129	478.53 (11.47)	268
T2 Picture Vocabulary	484.39 (8.77)	484.84 (11.52)	137	125	484.51 (10.13)	268
T1 Sound Awareness	470.68 (11.86)	468.27 (16.50)	138	128	469.31 (14.48)	268
T2 Sound Awareness	481.15 (11.41)	479.49 (133.33)	137	125	480.15 (12.37)	268

Note. T1 = time 1 (fall); T2 = time 2 (spring); HTKS = Head-Toes-Knees-Shoulders task; CBRS = Child Behavior Rating Scale; ID = identification. [a] For gender differences (raw data), $t = 2.90$, $p < .01$. [b] For gender differences (raw data), $t = 2.90$, $p < .01$. [c] For gender differences (raw data), $t = 2.60$, $p < .05$.

Elek-Fisk, 2003). Thus, we included teacher report data, all achievement variables, HTKS scores, and children's demographic and background information to impute missing values for maternal education, HTKS scores, and achievement data.

The second goal of ensuring the accuracy of our estimates of missing data influenced how we decided to handle missing teacher-report CBRS data. Our concern was that teachers who did not complete CBRS ratings of children also did not complete other measures of children's competencies; therefore, there would be limited available data associated with CBRS ratings, which may compromise the accuracy of the imputed values. In other words, we did not have good predictor variables that would enable us to adequately impute missing CBRS data for these children. Therefore, for CBRS analyses, we used the imputed data set, selecting a subsample of only those 156 children whose teachers originally reported CBRS data. For all other analyses, we used the entire imputed data set, with sample size of 268.

We conducted an analysis of attrition bias to test whether children missing CBRS data differed in some way from children with these data. However, checks conducted prior to single imputation revealed that the 156 children with CBRS data did not differ from the 112 lacking CBRS data by age, $t(266) = -0.75$, $p > .05$; gender, $t(266) = 0.20$, $p > .05$; minority status, $t(266) = 0.26$, $p > .05$; or fall HTKS score, $t(266) = 0.27$, $p > .05$. Children with CBRS reports had mothers with slightly lower education levels (15.9 years) compared with children without CBRS data (16.6 years), $t(266) = -2.35$, $p < .05$. In a logistic regression, missing CBRS data was also not related to the number of child study participants in the teacher's class ($B = -0.04$, $p = .40$). These analyses indicated that missing CBRS data was not systematically related to other child variables and was likely due to unmeasured teacher variables not a focus of our study.

We also conducted all analyses with both data sets (raw data with listwise deletion and imputed); patterns were highly similar, although in the results using listwise deletion, because of the limited sample size, few associations were statistically significant. Table 1 displays descriptive information from the imputed data. Sample size ($n = 156$ for analyses using the CBRS and $N = 268$ for all other analyses) is clearly marked in reported results.

Results

This study investigates associations between teacher ratings of child classroom behavior and performance on a direct assessment of self-regulation (HTKS). We also examine the presence, size, and influence of gender differences in self-regulation in kindergarten as measured by these two behavioral assessments. Finally, we evaluate the predictive value of gender and self-regulation for academic achievement.

Associations Between Teacher Ratings of Self-Regulation and HTKS Performance

We first report findings using the CBRS teacher ratings, which included the subsample of 156 children nested in 41 classrooms. Exploratory factor analysis was used to examine the number of factors obtained with the CBRS in this sample and whether particular items representing classroom behavioral regulation comprised a single factor, which could then be compared with the HTKS. Principal-components analysis with varimax rotation was used to derive the maximum difference between multiple factors for ease and clarity in interpretation. The factors were mildly correlated. Therefore, a principal-components analysis using oblimin rotation produced the same factors and same configuration of items on each factor with highly similar factor structure coefficients as our principal-components analysis using orthogonal rotation (Field, 2000).

Analyses revealed five distinct factors (see Table 2). This five-factor structure has been validated in other work in kindergarten students in Oregon (Ponitz, McClelland, Matthews, & Morrison, in press). The first factor had 10 items and described classroom behavioral regulation. All 10 items loaded highly, most with .8 or .7 coefficients on the factor. This factor accounted for 41.9% of the variance. This factor alone was used for additional analyses. The second factor consisted of seven items pertaining to interpersonal skills and accounted for 10.3% of the variance. The third factor consisted of seven items regarding nonacademic social play and explained 8.5% of the variance. The fourth factor had four items pertaining to engagement and explained 3.9% of the variance. Finally, the fifth factor had four items and described social problem solving, accounting for 3.6% of the variance.

We next assessed scale reliability, gender differences, and associations with children's HTKS performance for the first factor only (Classroom Self-Regulation). Cronbach's alpha for the 10 items on this factor was .95. The mean score of these 10 items was calculated and used as the child's teacher-rated self-regulation score in further analyses.

Teacher reports of student regulation in the classroom were higher for girls than for boys, $t(154) = 2.70$, $p < .01$. Gender had a small but significant effect on teacher ratings, with girls rated significantly higher than boys ($r^2 = .04$, $p < .05$). There was no statistically significant difference in the variance of boys' and girls' teacher ratings, Levene's $F(1, 155) = 0.65$, $p = .49$.

Children who received higher ratings on the CBRS (collected in the spring) also had higher HTKS scores in the fall ($r = .29$, $p < .01$) and spring ($r = .25$, $p < .01$; see Table 3). For girls, the fall HTKS and CBRS rating correlation was significant and positive ($r = .36$, $p < .01$), but the spring HTKS and CBRS rating correlation was not statistically significant ($r = .08$, $p = .48$). In contrast, for boys, fall HTKS and CBRS ratings were not correlated ($r = .16$, $p = .16$), but in the spring, boys with stronger HTKS scores received higher CBRS ratings ($r = .28$, $p < .05$).

Gender Differences in HTKS Performance

The second research aim was to examine gender differences in kindergarten HTKS performance, including mean scores, improve-

Table 2
Factor Analysis of Child Behavior Rating Scale Items (N = 156)

Factor and item	1	2	3	4	5
Factor 1: Classroom Self-Regulation					
Item 24: Responds to instructions and begins appropriate task	.868				
Item 23: Concentrates when working, not easily distracted	.860				
Item 21: Completes tasks successfully	.832				
Item 20: Completes learning tasks in an organized way	.797				
Item 29: Returns to unfinished tasks after interruption	.783				
Item 27: Finds and organizes materials	.747				
Item 28: Sees own errors on task and corrects them	.724				
Item 25: Takes time to do his/her best work	.723				
Item 22: Attempts new and challenging tasks	.635				
Item 15: Observes rules and follows directions without reminders	.580				
Factor 2: Interpersonal Skills					
Item 6: Expresses hostility—Physically		−.814			
Item 5: Expresses hostility—Verbally		−.773			
Item 7: Cooperates with playmates		.764			
Item 8: Takes turns without being told to do so		.714			
Item 13: Complies with adult directives—With little or no resistance		.680			
Item 3: Willing to share		.641			
Item 16: Does not fuss when doesn't get teacher's attention		.577			
Factor 3: Social Play–Interaction					
Item 9: Offers suggestions for play			.782		
Item 4: Plays with other children			.753		
Item 1: Joins in play with others			.738		
Item 11: Engages in pretend play			.733		
Item 10: Suggestions for play are accepted by other children			.732		
Item 14: Initiates social interaction with adults			.512		
Item 2: Comforts peers			.457		
Factor 4: Engagement					
Item 32: Shows enthusiasm for activities				.754	
Item 31: Conveys confidence about being able to succeed				.733	
Item 30: Interested in trying new activities, games, etc.				.730	
Item 19: Shows satisfaction when completes a project				.582	
Factor 5: Social Problem Solving					
Item 18: Tries to solve a problem before asking for help					.771
Item 12: Resolves potential social conflicts					.653
Item 17: Can deal with normal criticism or teasing					.649
Item 26: Feels s/he can cope well with classroom situations					.481

Note. The total variance explained is 68.3%. Factor 1 explained 41.9% of the variance, Factor 2 explained 10.3%, Factor 3 explained 8.5%, Factor 4 explained 3.9%, and Factor 5 explained 3.6%.

Table 3
Correlations and Descriptives

Variable	1	2	3	4	5	6	7	8	9	10
1. Gender (male = 1, female = 0)	—	−.17**	−.19**	−.24**	.13*	−.01	.06	.01	−.12	−.08
2. HTKS (fall)		—	.52**	.34**	.12	.27**	.44**	.26**	.33**	.40**
3. HTKS (spring)			—	.25**	.20**	.29**	.46**	.21**	.30**	.35**
4. Teacher ratings (n = 156)				—	.19**	.34**	.42**	.15*	.13*	.24**
5. Age					—	.09	.21**	.21**	.18**	.08
6. Letter–Word Identification (spring)						—	.61**	.39**	.43**	.68**
7. Applied Problems (spring)							—	.37**	.44**	.63**
8. Picture Vocabulary (spring)								—	.69**	.50**
9. Academic Knowledge (spring)									—	.57**
10. Sound Awareness (spring)										—

Note. N = 268, except where noted. Teacher ratings are reports of classroom self-regulation only (Factor 1). HTKS = Head-Toes-Knees-Shoulders task.
* $p < .05$. ** $p < .01$.

ment from fall to spring, and overall variability. These analyses used the full sample of 268.

Mean Scores

Independent sample t tests revealed that girls outperformed boys on the HTKS in both the fall and spring. Repeated-measures analyses indicated that both boys, $F(1, 120) = 28.23$, $p < .01$, and girls, $F(1, 128) = 38.26$, $p < .01$, improved significantly from fall to spring on the task, and girls had higher scores than boys in the fall and spring. There was no interaction between time and gender, $t(267) = -0.34$, $p = .74$. Boys' scores at the end of kindergarten did not differ statistically from girls' scores at the beginning of kindergarten, $t(266) = -1.43$, $p = .15$. Gender explained a small portion of the score variance in fall ($r^2 = .03$, $p < .05$) and spring ($r^2 = .04$, $p < .05$).

Variability in HTKS Scores

In general, according to guidelines by Kline (2005), skewness (the extent to which scores were skewed on one or both sides of the mean) and kurtosis (the extent to which scores were near or far from the mean) did not indicate severe nonnormality. However, there were observable differences in variability of boys' and girls' HTKS scores. Using Levene's test, gender differences in variability were statistically significant in the fall, $F(1, 266) = 5.80$, $p < .05$, and in the spring, $F(1, 266) = 14.50$, $p < .001$. Boys had more variable scores in the fall (SD = 10.3) and spring (SD = 9.3), whereas the variability of girls' scores was initially smaller (SD = 8.3) and diminished further by spring (SD = 5.4). Figures 1 and 2 display histograms of scores by gender in the fall and spring, respectively. Inspection of these distributions revealed that the overall HTKS distribution in the fall showed greater variability

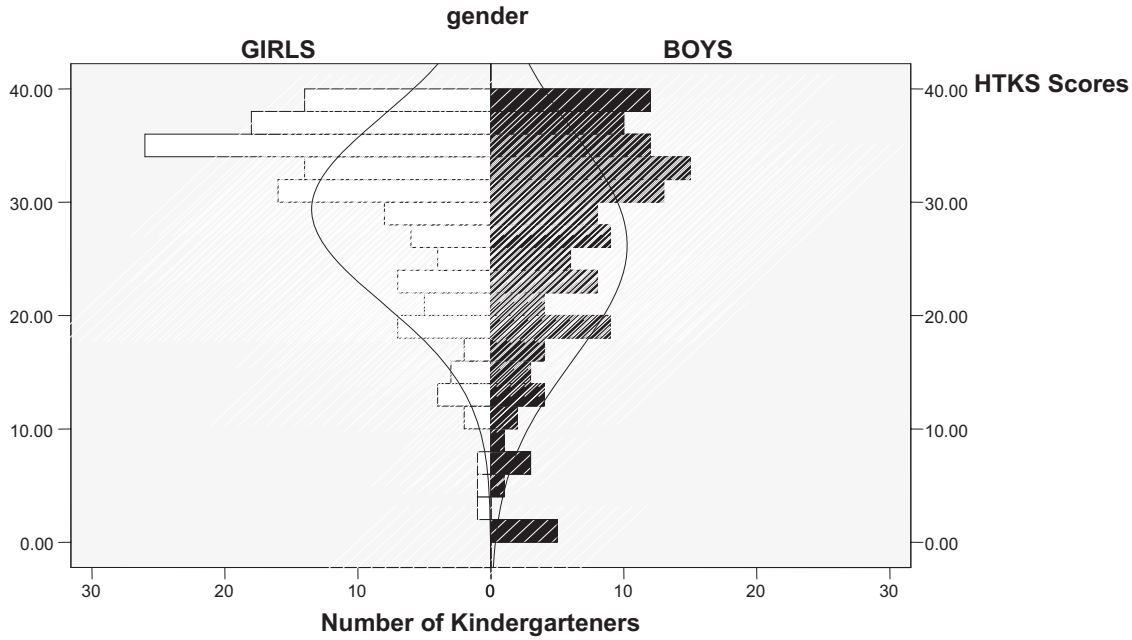

Figure 1. Distribution of fall Head-Toes-Knees-Shoulders task (HTKS) scores by gender (N = 268).

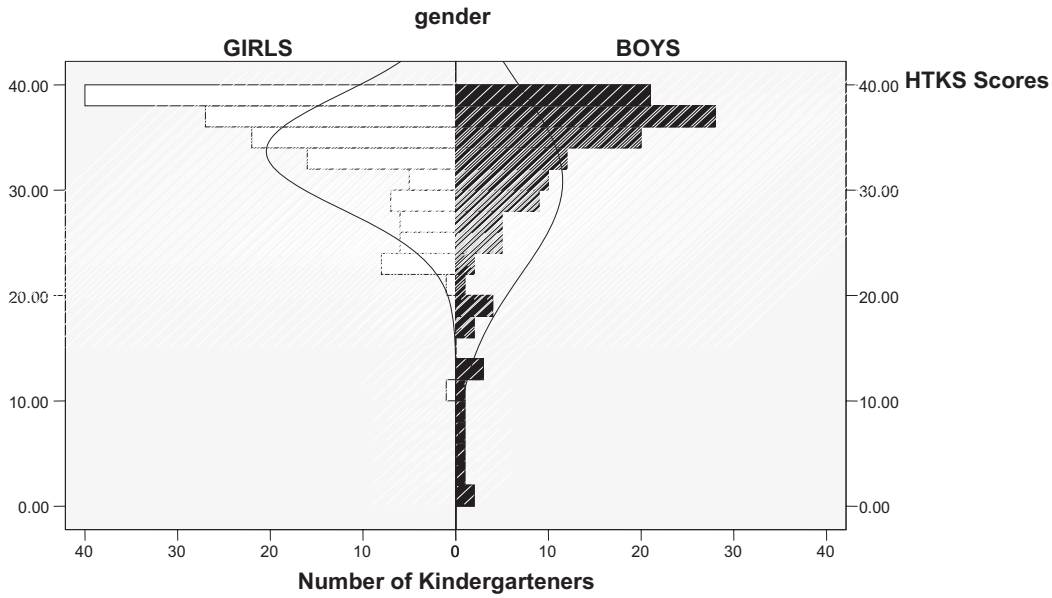

Figure 2. Distribution of spring Head-Toes-Knees-Shoulders task (HTKS) scores by gender ($N = 268$).

than in the spring, with some girls and boys represented at the bottom portions of the distribution curve. As expected, over the course of the year, students improved significantly in their scores, as detailed by our previous analyses and the spring distribution of HTKS. At both time points, girls' scores were more tightly clustered (skewed) toward the top of the distribution, whereas boys' were more evenly distributed.

Despite significant HTKS gains over time on average, a group of students continued to linger at the bottom end of the distribution in the spring. Twenty-two students scored 1 standard deviation below the mean in both the fall (18.1; $M = 11.2$) and the spring (24.5; $M = 16.4$). Sixteen of these 22 students (73%) were boys. Using another metric, 12 boys (86%) scored under 20 points (out of 40) in both the fall and spring, compared with 2 girls. We reviewed background data on children in this group and confirmed that none of the children suffered from documented cognitive or physical disability.

HTKS Performance and Kindergarten Achievement Gains

The third research aim was to explore whether HTKS scores and gender predicted gains in academic achievement over the kindergarten year. We used hierarchical linear modeling (HLM) to assess individual and classroom differences in academic skills as well as variability (Raudenbush & Bryk, 2002; Raudenbush, Bryk, Cheong, & Congdon, 2004). Our decision to use HLM was based on the nesting of 268 child participants within 87 classrooms (ranging from 1 to 8 child study participants per class). Although the average child n per classroom was small (about 3 children), the assumption of independence of observations was violated for kindergarteners sharing classroom membership. Unlike typical regression using ordinary least squares, HLM estimation algorithms (full information maximum likelihood) calculate standard errors accurately by accounting for nesting. This is a valid approach even when the Level 1 units (children) within each Level 2 unit (classroom) are relatively few (Raudenbush & Bryk, 2002).

Examination of the intraclass coefficient (ICC) confirmed the presence of differences in achievement outcomes attributable to classroom membership. We calculated the ICC for gain scores in five academic outcomes—applied problems (math), general knowledge, word reading, vocabulary, and sound awareness. We used gain scores (spring W score − fall W score) instead of residualized change (spring achievement, controlling for fall) because our goal was to identify predictors associated with new learning during kindergarten (Rogosa, Brandt, & Zimowski, 1982). Classroom differences were statistically significant at the $p < .05$ level for four outcomes: math (the ICC was 0.18, indicating 18% of the variance was at the classroom level), general knowledge (12%), vocabulary (11%), and sound awareness (12%). Less than 3% of the variance existed between classrooms in literacy, which was not statistically significant. We proceeded with HLM models because for four of the five outcomes, the ICCs indicated significant variance at the classroom level. Analyses using linear regression were conducted and yielded equivalent results. We built five HLM models for gains, controlling for age, maternal education, and fall achievement (see final model in the Appendix). We also wanted to assess whether gender and behavioral self-regulation at kindergarten entry, measured with the HTKS, predicted children's spring achievement beyond the contributions of these other variables. Continuous variables were grand mean centered, and gender was uncentered, with male coded as 1 and female coded as 0. Thus, the reference group for interpreting results includes girls, with average values on all predictor variables.

Gains in achievement were then modeled, along with predictors associated with gains above or below this average. We calculated effect sizes for achievement associated with a 1 standard deviation change in each predictor. To estimate the educational impact, we

also calculated achievement in months associated with predictor effects, based on the average monthly gain in that outcome (mean time between testing = 5.64 months). Comprehensive results from the HLM findings are presented in Table 4 for each outcome. For all models, the intercept (average gain) was significantly different from zero (see *t*s for $\gamma_{00,}$ which represent the grand mean gain for all children; see the Appendix for an explanation of each element in the model equations). Because we modeled gains, higher fall scores were associated with fewer gains, which can be expected because children initially scoring low in the fall had more room to improve (Rogosa et al., 1982). Finally, on the basis of descriptive results suggesting more boys lagged behind in HTKS performance, we tested the interaction between gender and fall HTKS scores to see whether boys with weak initial self-regulation had lower achievement relative to other children. This interaction was significant for only one outcome (vocabulary) and was thus omitted from the final models. Next, we provide a thorough practical interpretation for each predictor of gains in mathematics and briefly summarize findings for the other outcomes.

Applied Problems (Math)

The average increase in mathematics achievement over the year was 11.67 points (see Table 4). Compared with children with initially stronger fall math achievement and mothers with more years of education, those with lower fall math scores ($d = -0.40$) and whose mothers had lower levels of education ($d = -0.19$) tended to demonstrate greater gains over the school year, controlling for the other variables. For example, kindergarteners whose fall math scores were 1 standard deviation (12.74 points) below the fall average of 433.68 points were expected to gain 16.77 instead of 11.67 points (compute: the grand mean math gain – the coefficient for fall score multiplied by its standard deviation, or $\gamma_{00} - [\gamma_{10} \times SD]$; $11.67 - [-0.40 \times 12.74] = 16.77$). This translated to 1.81 additional months of learning (based on 11.67/5.64 months = 2.07 points per month). In addition, children with fall HTKS scores 1 standard deviation higher than average, or those scoring 37 instead of 28 points, made greater math gains, with an effect size roughly 50% larger than that associated with maternal education ($d = 0.29$). Compared with the initial model, the final model for math explained 28% of the total variance.

General Knowledge

The average gain in general academic knowledge was 9.22 points. Children scoring 1 standard deviation higher initially tended to gain 4.05 fewer points than average ($d = -0.43$). One additional predictor was significant; girls made greater gains than did boys in general knowledge at a level of marginal significance, $t(266) = -1.66$, $p = .10$, $d = 0.08$. The final model for general knowledge explained 18% of the variance.

Letter–Word Reading

The average gain in word reading was 30.59 points. Only one predictor was significant, such that children with weaker fall word reading scores tended to grow about 3.40 points more than average ($d = -0.23$). This model explained 4% of the total variance in letter–word reading.

Vocabulary

For gains in expressive vocabulary, fall score ($d = -0.54$) was associated with improving 4.70 points for children 1 standard deviation below the fall mean. There were no other significant predictors of vocabulary. The final model (shown in Table 4) explained 29% of the overall variance.

Sound Awareness

Children improved significantly on sound awareness (on average, 10.82 points), and children with fall scores 1 standard deviation below the mean improved 6.08 points beyond this ($d = -0.60$). In addition, relative to those scoring at the fall HTKS mean, kindergarteners with HTKS scores 1 standard deviation above the mean gained 1.67 points more ($d = 0.16$), which corresponded with 0.9 months of learning. The final model for sound awareness explained 32% of the overall variance.

Table 4
Hierarchical Linear Modeling Results for Five Models of Kindergarten Achievement Skills (N = 268)

Variable	Applied Problems	Academic Knowledge	Letter–Word Identification	Picture Vocabulary	Sound Awareness
Fixed effects					
Intercept, γ_{00}	11.67 (13.02**)	9.22 (11.92**)	30.59 (20.54**)	6.45 (9.41**)	10.82 (14.52**)
Fall age, γ_{10}	3.88 (1.78)	1.51 (0.97)	−0.62 (–0.24)	1.81 (1.51)	−1.20 (–0.71)
Gender, γ_{20}	1.36 (1.23)	−1.57 (–1.66)	−1.33 (–0.69)	−1.11 (–1.20)	−0.08 (–0.09)
Maternal education, γ_{30}	−1.59 (–3.51**)	−0.32 (–0.93)	0.57 (0.85)	−0.09 (–0.36)	0.00 (0.01)
Fall achievement, γ_{40}	−0.40 (–7.28**)	−0.32 (–5.46**)	−0.12 (–4.10**)	−0.41 (–8.93**)	−0.42 (–8.44**)
Fall HTKS, γ_{50}	0.40 (5.23**)	0.08 (1.32)	0.07 (0.78)	0.08 (1.47)	0.18 (2.97**)
Random effects					
Intercept, u_{0j}	16.22 (130.29**)	9.67 (128.48**)	12.55 (104.87†)	4.70 (124.19**)	3.79 (103.30†)
Level 1 effects, r_{ij}	98.93	63.59	205.02	49.45	66.75

Note. For fixed effects, values are coefficients (and *t* ratios); for random effects, values are variance (chi-square values). For gender, female = 0, male = 1. Gender was uncentered in the final model. All other variables were grand mean centered. Degrees of freedom for all models are as follows: child $df(\gamma_{10}–\gamma_{50}) = 262$; classroom $df(\gamma_{00}) = 86$. HTKS = Head-Toes-Knees-Shoulders task.
** $p < .01$. † $p < .10$.

Teacher Ratings of Classroom Behavioral Regulation and Achievement Gains

In a final set of analyses, we assessed the utility of CBRS teacher ratings of classroom self-regulation as a predictor of achievement gains, compared with HTKS scores. Thus, we ran a model for each outcome using the subset of 156 children with CBRS data. We modeled gain score and included the predictors of fall achievement, gender, maternal education, and CBRS teacher ratings collected in the spring of the kindergarten year. After examining these models, we added fall HTKS as a predictor to see which self-regulation measure—the teacher CBRS report or the structured HTKS task—better predicted achievement.

The pattern of results was highly similar to analyses using HTKS as a predictor for all variables. CBRS significantly predicted achievement outcomes in math, $t(155) = 3.20$, $p < .01$, $d = 0.20$, and sound awareness (at a level of marginal significance), $t(267) = 1.66$, $p < .10$, $d = 0.12$. When fall HTKS was added to the model for math, both measures of self-regulation were significant predictors: fall HTKS, $t(267) = 4.01$, $p < .01$, $d = 0.25$, and CBRS, $t(155) = 3.31$, $p < .01$, $d = 0.17$. For sound awareness, HTKS remained a significant predictor, $t(267) = 2.01$, $p < .05$, $d = 0.15$, but CBRS fell to nonsignificance, $t(155) = 1.40$, $p = .16$, $d = 0.10$.

Discussion

The HTKS direct measure of self-regulation obtained similar findings as an established teacher-report measure of child classroom regulation and revealed gender differences in both the fall and spring of kindergarten. Teacher ratings of self-regulation favored girls, and these differences were reflected more starkly in HTKS scores. On this direct measure, boys had more variable scores, including a number of consistently low performers, in comparison with girls who had a tight cluster of high performers and very few low performers. Despite gender differences in self-regulation, there were no significant gender differences on the five academic achievement outcomes at the end of the kindergarten year. However, both HTKS and teacher ratings of self-regulation predicted significant variance in math and phonological awareness gains, and results were stronger for the HTKS compared with the teacher-report measure.

Gender, HTKS Performance, and Teacher Ratings of Self-Regulation

We found clear gender differences in self-regulation at the beginning and end of kindergarten. Boys began the school year at a significant disadvantage in self-regulation in comparison with girls, and although they improved, they did not catch up by spring. Our results indicated that not until the end of the kindergarten year did boys develop the self-regulatory skills with which girls began the kindergarten year. Further, gender differences in self-regulation were manifested in both assessments. The corroboration of teacher ratings and objective measure scores strengthened the validity of our findings, suggesting that teacher evaluations are not inherently biased in kindergarten. This finding contrasts with other literature detailing how teachers may exaggerate gender differences in their reports (Cooper & Farran, 1988; Duckworth & Seligman, 2006; Rothbart et al., 2001; Rubin & Balow, 1976). In the present study, teacher ratings of self-regulation followed a similar pattern to that of the HTKS task; moreover, the differences described were not as sharp (although both were statistically significant). It should be noted that correlations among the HTKS and teacher reports were modest, although fairly consistent with other studies comparing observer reports with structured observational assessments of children's behavioral regulation (Howse et al., 2003; Smith-Donald, Raver, Hayes, & Richardson, 2007). Identifying behavioral and regulatory assessments that are representative of children's classroom functioning, and that strongly predict achievement outcomes, is a current priority for practitioners and researchers working to promote school success.

A closer inspection of gender differences in measures of self-regulation also uncovered some interesting trends, which future work needs to substantiate. It was not simply the case that all boys struggled in building self-regulatory competence in comparison with girls. Many boys scored very well on the HTKS task at both time points and received positive ratings from their teachers. However, the unsettling discrepancy revealed itself in students at the bottom portion of the performance curve. The bottom 10% of boys scored considerably worse than the bottom 10% of girls. This low group of boys also showed fewer gains on the HTKS over the course of the year compared with all other students. Although this trend was distinct visually (see Figure 2) and descriptively, the small number of boys in this group ($n = 16$) precluded further reliable statistical analyses to uncover achievement implications for this group. It is imperative that future research in the early elementary grades incorporate a specific focus on poorly regulating boys and their achievement outcomes in kindergarten and beyond.

Predicting Academic Achievement

Contrary to our expectations, there were no gender differences in achievement gains and neither measure of self-regulation predicted all achievement outcomes. However, findings were consistent in that both self-regulation measures predicted the same two achievement areas (math and phonological awareness). This was somewhat surprising, given the sizeable gender differences in self-regulation. Nonetheless, results are consistent with other work. Freeman (2004) showed that although gender differences in academic achievement exist in early elementary school, performance on assessments in general knowledge, overall reading, and overall mathematics is similar between boys and girls in kindergarten and the first grade. It would be useful to conduct inquiries throughout the early elementary grades to see whether gender differences in achievement emerge later (i.e., in second or third grade) and have origins in gender differences in kindergarten levels of self-regulation. Also, Duckworth and Seligman (2006) demonstrated that the gender gap is often not prevalent in standardized test scores, which were used in this study, but in grades and daily classroom work. Future research should take into account achievement test scores as well as school grades, which require daily persistence and the ability to delay gratification (e.g., doing homework today for a good mark or teacher praise later in the week or month). Nonetheless, performance on achievement tests, including the assessments used in this study, strongly predicts school grades, which helps explain their widespread use.

Our results lead us to consider that the impact of self-regulation on achievement may not emerge until children progress further into their elementary years. Kindergarten introduces children to formal schooling and requires that they learn a new role and adapt to the classroom context. Despite challenges, kindergarten nonetheless poses fewer academic demands than content taught in the later elementary grades (Entwisle & Alexander, 1998). This may help explain the apparent disconnect between self-regulation and standardized achievement at the kindergarten level. Nevertheless, as children progress, schoolwork intensifies and success requires mastery of new organizational and planning skills as well as motivation (Paris & Newman, 1990). Deficiencies in self-regulation present at a younger age may increasingly undermine academic progress and eventually predict achievement outcomes as well (Green & Francis, 1988; Vitaro et al., 2005).

Self-Regulation and Links to Math and Phonological Awareness

We found that behavioral self-regulation as assessed by the HTKS in the fall of kindergarten and teacher ratings gathered in the spring predicted gains in early math skills as well as sound awareness. The consistency of our findings increases our confidence in them, at least in this sample. There are two plausible explanations for this pattern. First, we must consider the cognitive requirements of our assessments of these academic domains. In both Applied Problems and Sound Awareness subtest items, children must use working memory and attention to keep information in mind while processing a problem and must also inhibit incorrect responses. For example, in Applied Problems, children were asked to look at an array of squares and assess how many there would be if four more were added. In Sound Awareness, they listened to the sounds of a word (sl/ee/p) and were instructed to reverse the sounds to form a new word (p/ee/ls). These complex cognitive skills were also required by the HTKS behavioral self-regulation measure. Other research has demonstrated the importance of working memory and inhibitory control for mathematics skills (Bull & Scherif, 2001; Espy et al., 2004). Another recent study linked inhibitory control to phonological awareness in kindergarten (Blair & Razza, 2007).

A second possibility, which speaks to the finding that teacher ratings predicted math (even when HTKS was included in the model), and predicted sound awareness until HTKS was included, may point to instructional practices and children's opportunities to learn math and phonological skills. It is possible that only kindergarteners with strong HTKS scores on school entry and those with classroom behavioral regulation skills (e.g., listening, following instructions, and completing activities) were able to benefit from learning opportunities regarding math and the underlying sounds of words. Although early elementary classrooms have been shown to be rich in language and literacy activity, the explicit teaching of phonological skills has been less consistently observed in classrooms serving middle-class populations like the one represented here (Connor, Morrison, & Katch, 2004). Moreover, classrooms spend very little time in math; this may mean that to learn math and phonological skills from the relatively few opportunities provided, children need strong behavioral regulation (Pianta, Cox, & Snow, 2007).

It was interesting to find that performance on a one-time, individually administered assessment (fall HTKS) was an equal or better predictor of gains in two areas of achievement compared with ratings of students who teachers had the benefit of knowing all year. Teacher ratings of behavior have often been used and have their own advantages (McClelland et al., 2000; Rimm-Kaufman et al., 2000). For instance, Sattler (1988) noted that teachers have the benefit of witnessing a child's behavior in a variety of settings and school-related activities. They also have the opportunity to notice rare but significant behaviors that single-occasion assessments likely miss. It is possible that teacher ratings would have proven more predictive of classroom-relevant academic outcomes, such as school grades or teacher ratings of unsatisfactory or satisfactory academic performance, which we did not obtain. In addition, the somewhat low response rate of teachers (58%) yielded a substantial number of unrated children ($n = 112$). We attributed the low response rate to teachers' fatigue from participating in a multiyear longitudinal study and found that children with teacher ratings did not differ from those who were not rated. This presents another advantage of the HTKS task for researchers, who can obtain behavioral regulation information from children even if teachers (or parents) do not return questionnaires about students' behavior.

Scientific and Practical Implications

Although a substantial amount of work has studied gender differences in academic achievement and some work on gender differences in self-regulation, the research literature in general remains unclear about the direct significance of gender on self-regulation and academic achievement. This article does not merely display the presence of a gender gap in self-regulation but provides new important details on the nature of this gap and the influence it may have in kindergarten. This includes variability among boys and girls, growth over the school year, and size of the self-regulation gap and its impact on achievement for boys and girls separately. In addition, it highlights the presence of this low male group, which has not been previously revealed in the literature and requires further examination.

Little work has shown this level of detail and investigation on gender differences in self-regulation. Further, no work to our knowledge has been able to display such consistent findings between an observer and direct measures of self-regulation descriptively or for predicting achievement. Finally, this work extends previous work (Ready et al., 2005) by adding other important academic domains of achievement. Evidence has illuminated the importance of both phonological and early mathematics skills, so identifying predictors of gains in these areas is important both scientifically and practically. Research has shown the blending and segmentation of sounds and phonics has the greatest transfer to emergent reading and spelling (Stuart & Rauth, 2006). Sound awareness and letter knowledge have also been described as the best predictors of reading success during the first 2 years of school, and sound awareness is a critical component of the complex literacy process (National Reading Panel, 2000). This study suggests that entering kindergarten with strong self-regulation early in kindergarten make more progress in the phonological skills vital for later reading success.

Further, one comprehensive analysis of six longitudinal data sets showed that early mathematics skills were better predictors of

achievement trajectories than other measures of achievement, as well as socioemotional competence and attention skills (Duncan et al., 2007). Although only one of the six studies included an established direct assessment of regulatory competence (i.e., the Continuous Performance Task, to measure focused attention), for the most part, attention and regulatory skills were measured with teacher and parent reports, whereas achievement was assessed directly. Until recently, the field has been somewhat limited by a paucity of direct measures of regulatory skills appropriate beyond a narrow age range (Isquith, Crawford, Espy, & Gioia, 2005). An investigation using sensitive measures of self-regulation as well as achievement is necessary for identifying whether early math skills are inherently important, as suggested by Duncan et al. (2007), or whether regulatory skills play a significant role in achievement trajectories as well, as suggested by other research (McClelland et al., 2006).

Conclusion

The HTKS direct self-regulation assessment, which has not been previously reported for a kindergarten sample, was consistent with teacher ratings of children's classroom behavioral regulation and predicted gains on two indexes of early achievement as well. Results also revealed a rift in self-regulation between genders but achievement outcomes that remained largely unaffected by gender or self-regulation in kindergarten. In addition, we identified a group of poorly regulating boys who may be at risk for school difficulties. This requires more research. Previous literature has indicated an achievement gap between boys and girls evident later in elementary and middle-school grades. Our findings combined with prior findings encourage us to ask whether the gender gap in self-regulation in kindergarten is the seed of a problem waiting to take root in achievement or other important outcomes (i.e., school grades) in later years. If so, our results suggest helping children develop self-regulation early in their school lives to increase the likelihood of equal opportunity to learn and positive outcomes for all.

References

Abt Associates. (1988, October). *Evaluation of project giant step* (Technical Progress Report No. 4). Cambridge, MA: Author.

Acock, A. C. (2005). Working with missing values. *Journal of Marriage and the Family, 67,* 1012–1028.

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1993). First-grade classroom behavior: Its short- and long-term consequences for school performance. *Child Development, 64,* 801–814.

American Association of University Women. (1992). *How schools shortchange girls: A study of major findings on girls and education.* Washington, DC: American Association of University Women Educational Foundation, The Wellesley College Center for Research on Women.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94.

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

Birch, S. H. L., & Gary, W. (1998). Children's interpersonal behaviors and the teacher–child relationship. *Developmental Psychology, 34,* 934–946.

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

Blair, C. (2003). Behavioral inhibition and behavioral activation in young children: Relations with self-regulation and adaptation to preschool in children attending Head Start. *Developmental Psychobiology, 42,* 301–311.

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

Bronson, M. B. (1991). *Manual for the Bronson Social and Task Skills Profile* (Rev. ed.). Chestnut Hill, MA: Boston College.

Bronson, M. B. (1994). The usefulness of an observational measure of young children's social and mastery behaviors in early childhood classrooms. *Early Childhood Research Quarterly, 9,* 19–43.

Bronson, M. B. (2000). *Self-regulation in early childhood: Nature and nurture.* New York: Guilford.

Bronson, M. B., Goodson, B. D., Layzer, J. I., & Love, J. M. (1990). *Child Behavior Rating Scale*. Cambridge, MA: Abt Associates.

Bronson, M. B., & Love, J. (1987). *Child Behavior Rating Scale.* Boston, MA: Abt Associates.

Bronson, M. B., Tivnan, T., & Seppanen, P. S. (1995). Relations between teacher and classroom activity variables and the classroom behaviors of pre-kindergarten children in Chapter 1 funded programs. *Journal of Applied Developmental Psychology, 16,* 253–282.

Brook, U., & Boaz, M. (2005). Attention deficit and hyperactivity disorder (ADHD) and learning disabilities (LD): Adolescents perspective. *Patient Education and Counseling, 58,* 187–191.

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

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

Coley, R. J. (2001). *Differences in the gender gap: Comparisons across racial/ethnic groups in education and work.* Princeton, NJ: Educational Testing Service.

Connor, C. M., Morrison, F. J., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading, 8,* 305–336.

Cooper, D. H., & Farran, D. C. (1988). Behavioral risk factors in kindergarten. *Early Childhood Research Quarterly, 3,* 1–19.

Cooper, D. H., & Speece, D. L. (1990). Instructional correlates of students' academic responses: Comparisons between at risk and control students. *Early Education and Development, 1,* 279–299.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38,* 352–362.

Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology, 36,* 161–174.

Duckworth, A., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98,* 198–208.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428–1446.

Eccles, J. (1994). Understanding women's educational and occupational choice: Applying the Eccles et al. model of achievement related choices. *Psychology of Women Quarterly, 18,* 585–609.

Eisenberg, N., Spinrad, T. L., & Smith, C. L. (2004). Emotion-related regulation: Its conceptualization, relations to social functioning, and socialization. In P. Philippot & R. S. Feldman (Eds.), *The regulation of emotion* (pp. 277–306). Mahwah, NJ: Erlbaum.

Eisenhart, M. A., & Finkel, E. (1998). *Women's science: Learning and succeeding from the margins.* Chicago: University of Chicago Press.

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin, 132,* 33–72.

Entwisle, D. R., & Alexander, K. L. (1998). Facilitating the transition to first grade: The nature of transition and research on factors affecting it. *Elementary School Journal, 98,* 351–364.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1997). *Children, schools, and inequality*. Boulder, CO: Westview.

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.

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher, 27,* 6–11.

Ferguson, D. M., Lloyd, M., & Horwood, L. J. (1991). Teacher evaluations of the performance of boys and girls. *New Zealand Journal of Educational Studies, 26,* 155–163.

Field, A. (2000). *Discovering statistics using SPSS for windows.* London: Sage.

Flynn, J. M., & Rahbar, M. H. (1994). Prevalence of reading failure in boys compared with girls. *Psychology in the Schools, 31,* 66–71.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74,* 59–109.

Freeman, C. (2004). *Trends in educational equity of girls and women.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Gambell, T. J., & Hunter, D. M. (1999). Rethinking gender differences in literacy. *Canadian Journal of Education, 24,* 1–16.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at seven years. *British Journal of Educational Psychology, 70,* 177–194.

Goodwin, W. L., & Driscoll, L. A. (1980). *Handbook for measurement and evaluation in early childhood education.* San Francisco: Jossey-Bass.

Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka & W. F. Velicer (Eds.), *Research methods in psychology* (pp. 87–114). New York: Wiley.

Green, L., & Francis, J. (1988). Children's learning skills at the infant and junior stages: A follow-on study. *British Journal of Educational Psychology, 58,* 120–126.

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 and Development, 14,* 101–119.

Howse, R. B., Lange, G., Farran, D. C., & Boyles, C. D. (2003). Motivation and self-regulation as predictors of achievement in economically disadvantaged young children. *Journal of Experimental Education, 71,* 151–174.

Hughes, C., Dunn, J., & White, A. (1998). Trick or treat? Uneven understanding of mind and emotion and executive dysfunction in "hard-to-manage" preschoolers. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 39,* 981–994.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139–155.

Isquith, P. K., Crawford, J. S., Espy, K. A., & Gioia, G. A. (2005). Assessment of executive function in preschool-aged children. *Mental Retardation and Developmental Disabilities Research Reviews, 11,* 209–215.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The development of self-regulation in the first four years of life. *Child Development, 72,* 1091–1111.

Korkman, M., Kemp, S. L., & Kirk, U. (2001). Effects of age on neurocognitive measures of children ages 5 to 12: A cross-sectional study on 800 children from the United States. *Developmental Neuropsychology, 20,* 331–354.

Kuhl, J., & Kraska, K. (1989). Self-regulation and metamotivation: Computational mechanism, development and assessment. In R. Kanfer, P. L. Ackermann, & R. Cudeck (Eds.), *Abilities, motivation and methodology: The Minnesota symposium on learning and individual differences* (pp. 343–374). Hillsdale, NJ: Erlbaum.

Ladd, G. W., Birch, S. H., & Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child Development, 70,* 1373–1400.

Lummis, M., & Stevenson, H. W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26,* 254–263.

Mather, N., & Woodcock, R. W. (2001). *Examiner's manual for the Woodcock–Johnson III Tests of Achievement.* Itasca, IL: Riverside.

McCabe, L. A., Rebello-Britto, P., Hernandez, M., & Brooks-Gunn, J. (2004). Games children play: Observing young children's self-regulation across laboratory, home, and school settings. In R. DelCarmen-Wiggins & A. Carter (Eds.), *Handbook of infant, toddler, and preschool mental health assessment* (pp. 491–521). New York: Oxford University Press.

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.

McClelland, M. M., Cameron, C. E., Wanless, S., & Murray, A. (2007). Executive function, behavioral self-regulation, and social–emotional competence: Links to school readiness. In O. N. Saracho & B. Spodek (Eds.), *Contemporary perspectives in early childhood education: Social learning in early childhood education* (Vol. 7, pp. 113–137). Charlotte, NC: Information Age Publishing.

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.

McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual: Woodcock–Johnson III.* Itasca, IL: Riverside.

Mead, S. (2006). *The truth about boys and girls.* Retrieved May 17, 2007, from http://www.educationsector.org/analysis/analysis_show.htm?doc_id=378705

Meleen, P., Love, J. M., & Nauta, M. (1988). *Final report. Vol. 1: Technical Report. Study of the home-based option in Head Start* (Contract No. 105–86-1602). Washington, DC: Administration for Children, Youth and Families.

Morrison, F. J., & Cooney, R. (2002). Parenting and academic achievement: Multiple paths to early literacy. In J. Borkowski, S. R. Landesman, & M. Bristol-Power (Eds.), *Parenting and the child's world: Influences on academic, intellectual, and social–emotional development* (pp. 141–160). Mahwah, NJ: Erlbaum.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Retrieved November 7, 2008, from http://www.nichd.nih.gov/publications/nrp/smallbook.pdf

Office of Juvenile Justice and Delinquency Prevention. (2006). *Juvenile offenders and victims: 2006 national report.* Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Paris, S. G., & Newman, R. S. (1990). Developmental aspects of self-regulated learning. *Educational Psychologist, 25,* 87–102.

Pianta, R. C., Cox, M. J., & Snow, K. L. (2007). *School readiness and the transition to kindergarten in the era of accountability.* Baltimore, MD: Paul H Brookes.

Piotrkowski, C. S., Botsko, M., & Matthews, E. (2000). Parents' and

teachers' beliefs about children's school readiness in a high-need community. *Early Childhood Research Quarterly, 15,* 537–558.

Ponitz, C. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23,* 141–158.

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (in press). A structured observation of behavioral regulation and its relation to kindergarten outcomes. *Developmental Psychology*.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International, Inc.

Ready, D., LoGerfo, L., Burkham, D. T., & Lee, V. E. (2005). Explaining girl's advantage in kindergarten literacy learning: Do classroom behaviors make a difference? *The Elementary School Journal, 106,* 21–38.

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly, 15,* 147–166.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92,* 726–748.

Rothbart, M. K., Ahadi, S. A., Hersey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development, 72,* 1394–1408.

Rothbart, M. K., Posner, M. I., & Kieras, J. (2006). Temperament, attention, and the development of self-regulation. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 338–357). Malden, MA: Blackwell.

Rubin, R. A., & Balow, B. (1978). Prevalence of teacher identified behavior problems: A longitudinal study. *Exceptional Children, 45,* 102–111.

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28,* 573–594.

Sadker, M., & Sadker, D. (1994). *Failing at fairness: How America's schools cheat girls.* New York: Macmillan.

Sameroff, A. J., Bartko, W. T., Baldwin, A., Baldwin, C., & Seifer, R. (1998). Family and social influences on the development of child competence. In M. Lewis & C. Feiring (Eds.), *Families, risk, and competence* (pp. 161–185). Mahwah, NJ: Erlbaum.

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

Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 125–151). Mahwah, NJ: Erlbaum.

Seppanen, P. S., Godin, K. W., Metzger, J. L., Bronson, M. B., & Cichon, D. J. (1993, April). *Observational study of Chapter 1-funded early childhood programs*. Hampton, NH: RMC Research.

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

Silverman, I. W. (2003). Gender differences in delay of gratification: A meta-analysis. *Sex Roles, 49,* 451–463.

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-Regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly, 22,* 173–187.

Sommers, C. H. (2000). *The war against boys: How misguided feminism is harming our young men.* New York: Simon and Schuster.

Stuart, R., & Rauth, J. (2006, July). *Sound instruction: Phonemic awareness in kindergarten and first grade.* Paper presented at the annual National Reading First Conference, Reno, NE.

Vitaro, F., Brendgen, M., Larose, S., & Tremblay, R. E. (2005). Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of Educational Psychology, 97,* 617–629.

Weaver-Hightower, M. (2003). The "boy turn" in research on gender and education. *Review of Educational Research, 73,* 471–498.

Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment.* Mahwah, NJ: Erlbaum.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III Tests of Cognitive Abilities.* Itasca, IL: Riverside.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68.*

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Erlbaum.

(*Appendix follows*)

Appendix

Hierarchical Linear Model

Level 1

$Y_{ij} = \beta_{0j} + \beta_{1j}$ (age) $+ \beta_{2j}$ (gender) $+ \beta_{3j}$ (maternal education) $+ \beta_{4j}$ (fall achievement) $+ \beta_{5j}$ (fall Head-Toes-Knees-Shoulders task) $+ r_{ij}$.

Level 2

$\beta_{0j} = \gamma_{00} + u_j$.
$\beta_{1j} = \gamma_{10}$.
$\beta_{2j} = \gamma_{20}$.
$\beta_{3j} = \gamma_{30.}$
$\beta_{4j} = \gamma_{40.}$
$\beta_{5j} = \gamma_{50}$.
$\beta_{6j} = \gamma_{50}$.

The outcome Y_{ij}, achievement gain for child i in classroom j, is a function of the intercept or sample average gain (β_{0j}), plus the contributions of age in the fall (β_{1j}), gender (β_{2j}), maternal education (β_{3j}), fall achievement (β_{4j}), and fall Head-Toes-Knees-Shoulders task (β_{5j}) plus error for the individual child (r_{ij}). The intercept is further defined at Level 2 as the mean of the classroom means (γ_{00}) plus error for classroom j (u_j). Thus, intercept differences were allowed to vary across classrooms. The effects of predictors $\beta_{1j} - \beta_{5j}$ were similar at Level 2 and so were fixed, indicated by the omission of an error term for each predictor (e.g., $\beta_{1j} = \gamma_{10}$). To produce meaningful coefficient estimates, we centered continuous variables and left the dummy variable (gender) uncentered.

Received July 19, 2007
Revision received September 8, 2008
Accepted September 24, 2008 ■