Developmental Psychology
2009, Vol. 45, No. 3, 605–619

0012-1649/09/$12.00 DOI: 10.1037/a0015365

A Structured Observation of Behavioral Self-Regulation and Its Contribution to Kindergarten Outcomes

Claire Cameron Ponitz
University of Virginia

Megan M. McClelland
Oregon State University

J. S. Matthews and Frederick J. Morrison
University of Michigan

The authors examined a new assessment of behavioral regulation and contributions to achievement and teacher-rated classroom functioning in a sample ($N = 343$) of kindergarteners from 2 geographical sites in the United States. Behavioral regulation was measured with the Head-Toes-Knees-Shoulders (HTKS) task, a structured observation requiring children to perform the opposite of a dominant response to 4 different oral commands. Results revealed considerable variability in HTKS scores. Evidence for construct validity was found in positive correlations with parent ratings of attentional focusing and inhibitory control and teacher ratings of classroom behavioral regulation. Hierarchical linear modeling indicated that higher levels of behavioral regulation in the fall predicted stronger levels of achievement in the spring and better teacher-rated classroom self-regulation (all ps $< .01$) but not interpersonal skills. Evidence for domain specificity emerged, in which gains in behavioral regulation predicted gains in mathematics but not in language and literacy over the kindergarten year ($p < .01$) after site, child gender, and other background variables were controlled. Discussion focuses on the importance of behavioral regulation for successful adjustment to the demands of kindergarten.

Keywords: academic achievement, behavioral self-regulation, structured observation, kindergarten, interpersonal skills

Making a successful transition to kindergarten is a critical milestone for young children. This move entails adjusting from the relatively unstructured activities of preschool to a formal school environment that demands more self-control (Pianta & Rimm-Kaufman, 2006). Researchers have suggested that mastering behavioral aspects of self-regulation helps children adjust to school, benefit from learning experiences, and succeed in social interactions (Blair, 2002; Dobbs, Doctoroff, Fisher, & Arnold, 2006; Fowler & Cross, 1986; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; McClelland, Cameron, Connor, et al., 2007; Miech, Essex, & Goldsmith, 2001; National Institute of Child Health and Human Development Early Childcare Research Network [NICHD ECCRN], 2003; Shonkoff & Phillips, 2000; Vitaro, Brendgen, Larose, & Tremblay, 2005). How well children navigate this change shapes their future educational trajectories (Aram, 2005; Entwisle & Alexander, 1998; McClelland, Acock, & Morrison, 2006). Further, poorly regulated children are at greater risk for low achievement, emotional and conduct problems, and school dropout in middle childhood and adolescence (Duncan et al., 2007; Eisenberg et al., 2000; Shaw, Gilliom, Ingoldsby, & Nagin, 2003; Vitaro et al., 2005). This research provides a strong incentive to develop reliable and predictive measures of behavioral regulation viable over the early childhood period.

Defining Behavioral Regulation

Behavioral regulation requires cognitively based operations that fall under the broad self-regulation construct (Baumeister & Vohs, 2004; Blair, 2002; McClelland, Cameron, Wanless, & Murray, 2007). We define behavioral regulation as the manifestation of executive function skills in overt, observable responses in the form of children's gross motor actions, which are also important for success in classrooms. Thus defined, behavioral regulation involves multiple components of executive function including *attentional focusing, working memory*, and *inhibitory control*. The focusing aspect of attention includes selecting and attending to

Claire Cameron Ponitz, Center for the Advanced Study of Teaching & Learning (CASTL), University of Virginia; Megan M. McClelland, Department of Human Development and Family Sciences, Oregon State University; J. S. Matthews, Combined Program in Education and Psychology, University of Michigan; Frederick J. Morrison, Department of Psychology, University of Michigan.

We thank the members of the University of Michigan Pathways to Literacy Project and Oregon State University Kindergarten Readiness Study and the participating families, teachers, and school district administration and personnel at both sites who made this study possible. We are also grateful to Alan Acock and CASTL researchers for guidance with data analyses. The Pathways to Literacy Project was supported by the National Institute of Child and Human Development and the National Science Foundation under Grants R01 HD27176 and 0111754, respectively. Additional funding was provided by the U.S. Department of Education Institute for Education Sciences under Grant R305H04013. The Kindergarten Readiness Study was supported by Oregon State University.

Correspondence concerning this article should be addressed to Claire Cameron Ponitz, Center for Advanced Study of Teaching & Learning (CASTL), 350 Old Ivy Way, Suite #300, Charlottesville, VA 22903. E-mail: ccponitz@virginia.edu

relevant information, such as listening to the teacher and focusing on a task (Barkley, 1997; Rothbart & Posner, 2005; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; Zelazo, Müller, & Goswami, 2002). Working memory entails cognitively maintaining and manipulating information; for example, remembering and then carrying out teacher instructions (A.-M. Adams, Bourke, & Willis, 1999). Inhibitory control helps children prevent or modify incorrect responses, such as remembering to raise their hand in class (M. J. Carlson & Corcoran, 2001; Diamond, Kirkham, & Amso, 2002; Dowsett & Livesey, 2000; Rennie, Bull, & Diamond, 2004).

Measuring Behavioral Regulation

Several direct assessments of behavioral regulation and its underlying cognitive skills have recently been developed (S. M. Carlson, 2005; Lengua, Honorado, & Bush, 2007; Ponitz et al., 2008; Schwebel, 2004). Out of 13 tasks reviewed by Rothbart, Posner, and Kieras (2006), 11 of them were performed successfully by children prior to 5 years of age. S. M. Carlson (2005) documented a similar pattern of ceiling effects on regulatory tasks for 5-year-olds. The few tasks that require abilities beyond those of 5-year-olds demand substantial time or specialized materials to administer; these include the attention network task (Rueda et al., 2004), Tower of London and Tower of Hanoi (Bull, Espy, & Senn, 2004); and advanced dimensional change card sort (Müller, Dick, Gela, Overton, & Zelazo, 2006). Moreover, psychometric data from diverse samples are not available for many of these assessments (Isquith, Crawford, Espy, & Gioia, 2005; Ponitz et al., 2008). See also Rothbart et al. and S. M. Carlson for further detail.

In the present investigation, we examined an original assessment of behavioral regulation, and aimed to extend previous work in this area. In these prior studies, we reported on the Head-to-Toes task (McClelland, Cameron, Connor, et al., 2007; Ponitz et al., 2008), adapted from an assessment by McCabe, Rebello-Britto, Hernandez, and Brooks-Gunn (2004) that required children to remember and use two rules to respond to behavioral commands (e.g., touch their head when told to touch their toes, and vice versa). The Head-to-Toes task showed convergent validity with teacher ratings of classroom behavioral regulation and variability in 4-year-olds but approached ceiling levels in participants older than 5 years (Ponitz et al., 2008). On the basis of these findings, we developed a new version, the Head-Toes-Knees-Shoulders (HTKS) task, with the goals of increasing task complexity and variability in kindergarten-age children, while maintaining the psychometric and pragmatic advantages of the Head-to-Toes task. Whereas the Head-to-Toes task included two rules, the HTKS ultimately requires children to remember four rules: "touch your head" is first paired with "touch your toes"; then, two commands are added, where "touch your knees" is paired with "touch your shoulders." To succeed, children must master and apply three cognitive skills to gross motor movements: (a) focusing on instructions and commands, (b) using working memory to remember and execute new rules while processing commands, and (c) inhibiting the automatic response while responding correctly.

The role each cognitive component plays in regulating overt behavior is still debated (Bronson, 2000; Müller, Zelazo, Hood, Leone, & Rohrer, 2004; Ponitz et al., 2008; Zelazo, Carter, Reznick, & Frye, 1997). Nonetheless, recent advances indicate that in the first several years of development, biological, cognitive, and contextual factors synergistically contribute to concrete skills such as whether a child can control his or her behavior in a classroom setting (Bulotsky-Shearer, Fantuzzo, & McDermott, 2008; Calkins, 2007). Moreover, research documents that early in life, individuals vary in their intensity of reaction to stimuli as well as in regulation of this reactivity (Fox & Calkins, 2003). Temperament research suggests the attentional focusing and inhibitory control aspects of effortful control are particularly relevant for the development of behavioral regulation (Eisenberg, Smith, Sadovsky, & Spinrad, 2004; Kagan, 1989; Rothbart, Ahadi, Hersey, & Fisher, 2001; Rothbart & Jones, 1998; Rothbart et al., 2006; Wachs, Gurkas, & Kontos, 2004). Temperament is often assessed on the basis of parent ratings and thus captures typical responses across multiple situations and stimuli (Rothbart et al., 2001). In the present study, our first goal was to explore the psychometric properties of the HTKS measure of behavioral regulation by examining variability in a sample of 5- to 6-year-olds and relations with parent temperament ratings and teacher reports of classroom behavioral regulation.

Linking Behavioral Regulation with Early Academic and Classroom Success

Extensive research has connected the cognitive skills of behavioral regulation with success in multiple domains. For example, stronger attentional focusing skills predict increased achievement, and high school graduation rates (Blair & Razza, 2007; Howse, Lange, Farran, & Boyles, 2003; Vitaro et al., 2005), as well as adaptive interpersonal outcomes such as peer competence and control of positive and negative emotion (Raver, Blackburn, Bancroft, & Torp, 1999; Trentacosta & Izard, 2007). Working memory has been tied to early reading, mathematics, and cognitive skills in elementary students (Gathercole & Pickering, 2000; Kail, 2003). Working memory deficits are less clearly associated with behavioral and social difficulty, bolstering the argument that the cognitive components involved in behavioral regulation are distinct (A.-M. Adams et al., (1999); Bull & Scerif, 2001; Pickering & Gathercole, 2004). Finally, lower levels of inhibitory control are associated with lower achievement in multiple samples (J. W. Adams & Snowling, 2001; Blair & Razza, 2007; Lawrence et al., 2002; St Clair-Thompson & Gathercole, 2006). The social problems exhibited by disruptive preschoolers can be attributed, in part, to their poor inhibitory control (Hughes, White, Sharpen, & Dunn, 2000), linked to weak rule internalization in toddlerhood (Kochanska, Murray, Jacques, & Koenig, 1996).

Overall Behavioral Regulation

Many relevant developmental tasks, including adapting one's behavior to classroom demands, require multiple cognitive skills working together (Blair, 2002). This integration warrants the construct, *behavioral regulation*, which we used in the present inquiry. Converging evidence makes clear that overall behavioral regulation predicts how well children adapt to school and their subsequent achievement, including achievement trajectories through second grade and achievement levels at sixth grade (Bronson, Tivnan, & Seppanen, 1995; Eisenberg et al., 2000; Howse, Calkins, et al., 2003; McClelland et al., 2006; McClelland, Cameron, Connor, et al., 2007; McClelland, Cameron, Wanless, et al., 2007; McClelland, Morrison, & Holmes, 2000). However, most studies of kindergarten behavioral regulation and academic and social outcomes have not gathered information directly from children with structured observational assessments.

Behavioral Regulation and Interpersonal Skills

Research has connected stronger effortful control, important for behavioral regulation, with better interpersonal skills requiring emotion regulation, such as being able to control negative outbursts of anger and aggression and modulate positive reactions like excitement (Denham et al., 2003; Diener & Kim, 2004; Eiden, Edwards, & Leonard, 2007; Eisenberg et al., 2004; Murray, Sektnan, & McClelland, 2006). This and other evidence suggests behavioral regulation might be significantly predictive of interpersonal skills as well as of achievement (Howse, Calkins, et al., 2003; Trentacosta & Izard, 2007). Nonetheless, prior investigations of these constructs have often relied solely on questionnaire reports of children's regulation (Diener & Kim, 2004; Howse, Calkins, et al., 2003; Mashburn, Hamre, Downer, & Pianta, 2006; Trentacosta & Izard, 2007). In our prior work with the Head-to-Toes task, we looked only at its associations with achievement and demographic variables (McClelland, Cameron, Connor, et al., 2007; Ponitz et al., 2008). To clarify these issues, we had as our second aim an examination of the links between the HTKS and multiple markers of kindergarten functioning, including achievement, classroom behavioral regulation, and interpersonal skills requiring emotion regulation. We expected that because of the cognitive processing demands of the HTKS, it would most strongly predict achievement and classroom behavioral regulation, relative to interpersonal skills.

Developmental Implications of Behavioral Regulation in Kindergarten

Success in kindergarten has short- and long-term developmental consequences. In one study, the cognitive and linguistic skills with which children entered kindergarten contributed uniquely to their positive social functioning (teacher and peer relationships; teacher-reported classroom participation) and academic achievement during the year (Ladd, Birch, & Buhs, 1999). In a longitudinal inquiry, teacher ratings of kindergarteners' classroom functioning, including behavioral regulation, predicted achievement trajectories through second grade and achievement levels through sixth grade (McClelland et al., 2006). Another recent study revealed kindergarteners with poor behavioral regulation (attention and hyperactivity problems) were more likely to drop out of high school than their well-regulated peers (Vitaro et al., 2005).

This evidence frames children's regulatory repertoires in kindergarten as laying a foundation for multiple areas of development. Recent research reveals the importance of early behavioral regulation for multiple areas of achievement, including reading and mathematics (Aram, 2005; Blair & Razza, 2007; McClelland, Cameron, Connor, et al., 2007; Ready, LoGerfo, Burkam, & Lee, 2005). Therefore, the third aim of the present study was to examine the relative contribution of overall behavioral regulation to literacy, vocabulary, and mathematics gains over the kindergarten year.

Recent evidence suggests particularly strong links between aspects of behavioral regulation and mathematics achievement, relative to associations between behavioral regulation and language and literacy skills (Blair & Razza, 2007; Duncan et al., 2007). However, in our work with the Head-to-Toes task in prekindergarteners (many of whom participated in the current study as kindergarteners), we found that behavioral regulation predicted mathematics, vocabulary, and literacy similarly (McClelland, Cameron, Connor, et al., 2007). During early childhood, dynamic and interrelated changes occur in regulatory skill development as well as learning in multiple academic areas (Snow, 2007). This highlights a need for further exploration of possible domain-specific relations between behavioral regulation and achievement domains emerging in the early years (McClelland, Cameron, Connor, et al., 2007). These relations may mean that behavioral regulation relates broadly to achievement earlier in the school trajectory but becomes more distinctly predictive of specific skills as children develop.

The Current Study

We examined the psychometric properties of a behavioral regulation measure, links to parent- and teacher-ratings, and relations to achievement and social behavior. In hierarchical analyses, we controlled for several theoretically important child-level variables associated with skill development: age, gender, parent education, and ethnicity, denoted by Spanish-speaking status (Christian, Bachman, & Morrison, 2001; McLoyd, 1998; Ready et al., 2005). Our study design was correlational and included data collected in two sites in the United States (Oregon and Michigan); thus, we also examined site as a classroom-level predictor of each outcome.

The following three research questions were posed. First, what is the extent of variability and gains in kindergarteners' scores on a new task measuring behavioral regulation (the HTKS), and how does HTKS task performance correlate with parent and teacher ratings? Significant variability and gains over the kindergarten year were expected in behavioral regulation, and children with better performance on the task were expected to receive higher ratings of attentional focusing and inhibitory control (from parents) and higher ratings of classroom behavioral regulation (from teachers).

Second, does behavioral regulation at kindergarten entry show predictive validity for end-of-kindergarten mathematics, literacy, and vocabulary achievement and teacher ratings of classroom functioning? Controlling for background variables and geographical site, we expected that children demonstrating better behavioral regulation at fall of kindergarten would earn higher scores on end-of-year achievement tests and teacher ratings of classroom behavioral regulation, and interpersonal skills, relative to those with initially lower behavioral regulation. Behavioral regulation performance was expected to be less strongly predictive of interpersonal skills ratings than achievement and teacher-rated classroom behavioral regulation.

Third, does initial behavioral regulation predict kindergarten gains in achievement? We expected that children whose behavioral regulation was strong upon kindergarten entry would show gains in mathematics, literacy, and vocabulary, after initial competence levels and background variables were controlled. This expectation was derived from our prior research showing performance on the simpler (two-rule) Head-to-Toes task predicted gains in these three domains in prekindergarten (McClelland, Cameron, Connor, et al., 2007).

Method

Participants

Children were recruited from two sites: a predominantly middle- to upper-middle-socioeconomic status (SES) urban fringe area with a range of economic and ethnic diversity in Michigan and a mixed-SES rural site in Oregon. Parents of all children and kindergarten teachers provided written informed consent prior to participation.

Recruitment and Attrition at Each Site

Michigan site. Participants in Michigan were part of a 5-year longitudinal study on early academic and social development. All 3- and 4-year-olds 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 study. Recruiting was stopped after the target sample size had been achieved. Recruitment efforts resulted in approximately 38% of the district's entering preschoolers being enrolled in the longitudinal study, with a final sample size for the present investigation of 281 children retained through their kindergarten year. Current study participants were enrolled from 89 classrooms in 12 schools.

Of 412 children recruited, 100 left the longitudinal study before entering kindergarten; the other 30 participants were not yet in kindergarten. Comparisons of background variables showed the attrited group did not differ from the participants by gender or parent education levels ($p > .05$). However, 81% of retained families reported parent education, compared with 59% of those who left the study. In addition, the ethnic make-up of the two groups differed significantly: Compared with the attrited group, the participant group had more participants who were Asian, $\chi^2(1, n = 407) = 10.31$, $p < .01$, and more participants who were White, $\chi^2(1, n = 407) = 55.44$, $p < .01$. One additional child, who was younger than 4 years old at the beginning of kindergarten, was excluded.

Oregon site. Children and parents in Oregon were part of a 2-year investigation of factors related to self-regulation in preschool and kindergarten. Three National Association for the Education of Young Children–accredited preschools and three Head Start preschools located in Oregon were invited to participate in the study, and all agreed. Recruitment letters were sent home with approximately 165 children entering prekindergarten, and 95 families (or 58%) agreed to participate. Two children were excluded from the study because neither English nor Spanish was their first language, making the final sample size 93 preschoolers. In Year 2 (the year of the present inquiry), 62 children remained in the study and entered 23 kindergarten classrooms in 16 elementary schools.

In comparisons of the 31 children leaving the Oregon study with the 62 retained, there were no significant differences in gender, parent education level, age, or Spanish-speaking status ($p > .05$). There was a significant difference in ethnicity, such that there were significantly more White participants who stayed in the study, $\chi^2(1, n = 93) = 4.84$, $p < .05$ and significantly more Asian participants who left the study, $\chi^2(1, n = 93) = 4.96$, $p < .05$.

Combined Sample

Site comparison. Comparisons by site revealed several differences. Compared with Michigan kindergarteners, children in Oregon were 2.3 months older in the fall, $t(113.2) = 4.97$, $p < .01$, and their parents had lower levels of education by 1.5 years, $t(54.9) = -2.78$, $p < .01$. Participants in Michigan had on average 10 fewer months of child care and preschool experience relative to Oregon participants, $t(63.1) = 3.53$, $p < .01$. There were more members of ethnic minority groups in Oregon, $t(78.3) = 2.38$, $p < .01$. The Michigan sample was 77% White, 8% Asian, 7% African American, 6% Arabic, and 3% Latino. The Oregon sample was 56% White, 24% Latino (of whom 13 of 15 were primarily Spanish speakers), 13% Asian, and 7% other.

Descriptives. The full sample included data for the 343 children (52% girls, 48% boys) from Michigan and Oregon in the fall and spring of kindergarten. The mean age at Time 1 (fall) was 5.48 years ($SD = 0.33$ years). On average, children from both sites had spent 14.18 months in child care and preschool prior to data collection at Time 1 ($SD = 16.36$ months). The mean parent education level was almost a college degree ($M = 15.68$ years, $SD = 2.30$) across the two sites; 18% of children came from non-Asian ethnic minority groups. Two small subgroups were Spanish speaking ($n = 13$) or had parents with a high school degree or less ($n = 14$).

Procedure

Data were collected from three sources: parents, teachers, and children. In addition to completing background questionnaires, parents rated children's temperament (attentional focusing and inhibitory control) in the spring and summer before kindergarten. Teachers reported on children's classroom behavioral regulation and interpersonal skills in the spring of kindergarten. As part of two batteries, the behavioral regulation assessment was administered to children in the fall and spring of kindergarten along with three achievement measures— mathematics, literacy (letter–word reading), and vocabulary. Each battery was administered by research assistants in individual sessions at the school (30–40 min in Michigan; 10–15 min in Oregon). In Oregon, teachers identified students who had Spanish as their first language; these children were given the achievement tests and the behavioral regulation task in Spanish. Assessments were conducted in a quiet room or hallway outside participants' classrooms.

Parent and Teacher Questionnaire Measures

Background questionnaire. Parents completed a background questionnaire in English or Spanish (Oregon only) seeking information about child age, gender, prior child care/preschool experience, ethnicity, and parent education level. In Oregon, two native Spanish-speakers translated the questionnaire into Spanish and backtranslated it into English.

Parent ratings of temperament. Parents at both sites rated children on the Short Form of the Child Behavior Questionnaire (CBQ), a widely used instrument that captures temperament in children over the ages of 3–7 years, with ratings showing strong reliability and validity (Putnam & Rothbart, 2006; Rothbart et al., 2001). In the present study, we used the attentional focusing (six items, e.g., "When drawing or coloring in a book, shows strong concentration") and inhibitory control scales (six items, e.g., "Can wait before entering into new activities if asked") from the CBQ Short Form. Ratings are made from 1 to 7, with 1 representing *weak* and 7 representing *strong attentional focusing or inhibitory control*. Scale reliability analyses at both sites resulted in interitem reliability alphas ranging .67 to .82. In Michigan, 162 families (58%) returned temperament data, and in Oregon, 52 families (84%) completed temperament ratings.

Teacher ratings of classroom functioning. We obtained teacher ratings of classroom functioning using items on a 5-point Likert-type scale culled from the Child Behavior Rating Scale (CBRS; Bronson et al., 1995); ratings of 1 indicated that a child *never* exhibited the desired behavior, and ratings of 5 indicated that a child *always* exhibited the desired behavior . Previous evidence indicates strong reliability and validity for the CBRS (Bronson et al., 1995; Matthews, Ponitz, & Morrison, in press; McClelland, Cameron, Connor, et al., 2007; McClelland & Morrison, 2003). In previous research, the CBRS was found to be positively correlated ($r = .43$) with the Bronson Social Task and Skill Profile (BSTSP), an observational measure of children's classroom goal-oriented behaviors and strategies used to regulate behavior in academic and social situations (Bronson, 1994; Goodwin & Driscoll, 1980; Seppanen, Godin, Metzger, Bronson, & Cichon, 1993). In other research, CBRS items loaded primarily on a behavioral regulation factor (.60–.80), accounting for roughly 42% of the total variance. An interpersonal factor was the second strongest factor, also with high item loadings (.60–.80) accounting for 10% of the total variance (Matthews et al., in press).

For the present study, we conducted exploratory factor analyses at each site (using principal axis factoring and promax rotation), and the two largest factors tapped classroom behavioral regulation and interpersonal skills. Patterns were highly similar at each site. The first factor included 10 items assessing classroom behavioral regulation, such as "observes rules and follows directions without reminders," "completes learning tasks involving multiple steps in an organized way," "concentrates when working," and "returns to unfinished tasks after interruption." The second factor included 7 interpersonal skill items, for example, "cooperates with playmates," "expresses hostility physically (reverse scored)," "is willing to share," and "takes turns." For analyses, we created scale composites using the average of the item ratings; scores of 1 indicated *low teacher-rated functioning* and scores of 5 indicated *high teacher-rated functioning*. Internal reliability coefficients ranged .89–.95 for the two scales. In Michigan, teacher-report data were available for 157 children (56%). In Oregon, 56 children (90%) had teacher-report data.

Child Achievement Measures

To assess achievement in mathematics, literacy, and expressive vocabulary, we used three subtests from the Woodcock–Johnson Psychoeducational Battery III Tests of Achievement (Woodcock & Mather, 2000) or the *Batería Woodcock–Muñoz—Revisada: Pruebas de Aprovechamiento* [Woodcock–Muñoz Review Battery: Tests of Achievement] (Woodcock & Muñoz-Sandoval, 1996; Oregon site only). For all three measures, we used two different forms in the fall and spring. Basal and ceiling guidelines for each subtest indicated that testing should stop after the participant answered six or more items incorrectly. We used W scores, a conversion of raw scores similar to the Rasch ability scales (Rasch, 1960), in analyses based on a centered W score of 500, representing the average achievement level for a 10-year-old child (Foorman, Francis, Fletcher, & Lynn, 1996; Mather & Woodcock, 2001; Merrell & Tymms, 2007). W scores are well suited for investigations of achievement within and across years.

The Applied Problems subscale measures mathematics skills and requires children to manipulate pictured quantities, answer questions about time and money, and complete word problems. The Picture Vocabulary subscale primarily assesses expressive vocabulary, and children are asked to name the images of nouns increasing in difficulty. The Letter–Word Identification subscale measures word-reading skills, with children asked to name real letters and then read actual words, which increase in difficulty. All three subtests are commonly used with interitem score reliability norms for North American 5-year-olds speaking English or Spanish at or above .83 for the Applied Problems and Letter-Word Identification subscales and at or above .68 for the Picture Vocabulary subscale (Mather & Woodcock, 2001; Woodcock & Muñoz-Sandoval, 1996).

Structured Observation of Behavioral Regulation

The Head-Toes-Knees-Shoulders (HTKS) task was developed as a complex, extended version of the Head-to-Toes task, which measures behavioral regulation with children's responses to 10 trial commands, with scores showing reliability and validity with preschoolers in recent research (McClelland, Cameron, Connor, et al., 2007; Ponitz et al., 2008). The HTKS included 20 test trials and was designed for early elementary students. After habituating to two oral commands (e.g., "touch your head" and "touch your toes"), children were asked to respond in an unnatural way to two types (on the first 10 trials from the Head-to-Toes task) and then four types (on the second 10 trials) of paired behavioral commands. For example, if the administrator said, "Touch your toes," the correct response would be for the child to touch his or her head; the correct response to a "Touch your knees" command would be for the child to touch his or her shoulders. Correct responses earned 2 points; incorrect responses earned 0 points; 1 point was given if children made any motion to the incorrect response, but self-corrected and ended with the correct action. Scores ranged 0 to 40.

Commands were given in a consistent, nonrandom order. Higher scores indicated higher levels of behavioral regulation. For Spanish-speaking children, the HTKS was translated into Spanish and backtranslated into English by two Spanish speakers, including a professor of Spanish. A dummy variable indicated whether children received the HTKS in Spanish or English and was used as a control variable (see Table 1 and Results section). Two forms of the task were used, counterbalancing the commands used on the first 10 trials (heads/toes vs. knees/shoulders). No significant form differences in overall scores emerged in Oregon or Michigan at either fall or spring ($p > .05$).

Training and interrater reliability. Examiners learned to administer the task by reading the protocol, watching a trained

Table 1
Descriptive Statistics for Michigan Site (N = 281), Oregon Site (N = 62), and Combined Sample (N = 343)

Variable	Raw data by site								Imputed data (combined sample)[a]		
	M		*SD*		Range		*N*				
	MI	OR	MI	OR	MI	OR	MI	OR	*M*	*SD*	Range
Direct assessment											
Fall HTKS	27.50	26.80	9.62	10.57	0–40	1–39	265	62	27.45	9.72	0–41.74
Spring HTKS	32.30	28.00	7.55	11.39	0–40	0–40	255	59	31.31	8.61	−0.0–48.15
Child age (in years)											
Fall	5.44	5.63	0.34	0.26	4.14–6.21	5.08–6.17	272	62	5.48	0.33	4.14–6.21
Spring	5.91	6.13	0.34	0.26	4.55–6.69	5.58–6.67	269	62	5.96	0.33	4.55–6.69
Gender (% male)	0.48	0.47	0.50	0.50	0–1	0–1	281	62	0.48	0.50	0–1
Fall prior child care experience (in months)	12.48	22.45	15.34	18.72	0–48	0–57	245	50	13.96	16.47	−26.42 to 57.87
Parent education level (in years)	15.96	14.42	1.59	3.93	12.5–18	4–21	226	52	15.56	2.41	3.98–21.27
Spanish-version HTKS (% Spanish-speaking)	0	0.21	0	0.41	0–0	0–1	281	62	0.03	0.19	0–1
Parent-rated CBQ											
Attentional focusing	5.05	5.01	0.92	1.02	2.5–7	2–6.67	172	56	5.00	0.96	2–7.48
Inhibitory control	5.07	4.99	0.83	0.95	2.17–7	2.33–6.67	174	56	5.01	0.87	2.17–7.18
W scores											
Math											
Fall	433.39	437.13	14.48	17.66	373–467	382–417	272	62	434.07	15.06	374–471
Spring	455.59	450.34	15.18	21.29	405–485	388–494	269	59	446.54	16.56	388–494.23
Literacy											
Fall	374.59	380.03	28.94	38.88	305–508	310–528	272	62	375.42	30.81	305–528
Spring	404.42	415.19	29.62	39.61	326–515	320–548	269	59	406.58	32.05	320–548
Vocabulary											
Fall	478.50	474.27	11.51	18.86	409–513	423–509	272	62	477.78	13.19	409–513
Spring	484.59	482.49	10.23	17.73	435–517	435–520	269	59	484.22	11.93	435–520
Spring teacher-rated											
Behavioral regulation	3.94	4.17	0.72	0.62	1.4–5	2.5–5	158	56	3.98	0.70	1.40–5.68
Interpersonal skills	3.98	3.93	0.65	0.58	1.86–5	2.67–4.78	157	56	3.98	0.64	1.82–5.66

Note. HTKS = head-toes-knees-shoulders task; CBQ = Child Behavior Questionnaire; W scores = conversion of raw scores similar to the Rasch ability scales (Rasch, 1960).
[a] Estimates calculated by averaging aggregated results from 10 imputed data sets (Yuan, 2005).

examiner administer the task, and practicing administration. Interrater reliability has been established in Oregon on the Head-to-Toes task (HTT), a simple version of the HTKS. In one study, two examiners coded 42 children on the HTT and showed strong agreement in total scores ($\kappa = .90$) compared with what would be expected by chance (McClelland, 2007). Other recent work has established strong interreliability and validity on the Head-to-Toes task and the HTKS (Connor et al., 2007; Ponitz et al., 2008). However, it was also important to evaluate scoring consistency for the extended version in this study.

Scoring consistency. We assessed consistency in obtained scores across examiners for both sites of children on HTKS total score and self-corrected responses. At the Michigan site, controlling for child age, we found no significant differences by examiner in the fall in HTKS scores or average number of self-corrects ($p > .05$). In the spring, examiner differences emerged for total score, $F(11, 169) = 4.58$, $p < .01$, related to two examiners testing children with extremely low scores. In addition, two examiners gave significantly higher scores compared with the other examiners. Thus, for the 12 testers, cross-examiner consistency was 66%. In the spring, three examiners scored significantly more (one examiner) or fewer (two examiners) self-corrects, $F(11, 169) = 5.25$, $p < .01$, so consistency across examiners for self-corrects was 75%. At the Oregon site, after we controlled for child age and being a Spanish-speaker, no significant differences across examiners were found in overall HTKS scores in the fall or spring. Similarly, there were no significant differences by examiners at Oregon in the average number of self-correct responses in the fall or spring (all $ps > .05$).

Missing data analysis. For the achievement and behavioral regulation outcome variables, between 314 and 334 participants had complete data. Parent education data were available for 278 participants. For the CBRS and CBQ, data were available for 214 and 228 participants, respectively. In addition, there were 138 children with fall data but who were missing at least one spring variable. Variables with greater than 5% missing data points were spring HTKS score, parent education, child care experience, CBQ ratings of attentional focusing and inhibitory control, and CBRS ratings of behavioral regulation and interpersonal skills. We first conducted independent sample *t* tests comparing children with versus without our main spring outcomes. Children missing spring achievement data did not differ from those with complete spring data on fall achievement, fall HTKS, CBQ ratings, or CBRS ratings. Children missing a spring HTKS score did not differ from children with spring HTKS data on these variables either, except that children without spring HTKS scores had significantly lower fall HTKS scores ($M = 23.5$) than children with spring HTKS scores, ($M = 27.7$), $t(325) = 2.14$, $p < .05$, indicating potential nonrandom missingness of spring HTKS scores.

Next, we tested the missing data assumption (missing at random, or MAR), which assumes that missing data be explained by variables included in analyses and that remaining missingness is at random (Schafer & Graham, 2002). There is no definitive way to test whether remaining missingness is random, but we were able to test whether other auxiliary variables not included in final analyses predicted missingness. For those variables with greater than 5% missing data, we conducted logistic regression analyses using dummy variables indicating whether data were missing on a variable. For each case, the variable was coded 1 if missing and 0 if complete. Missingness on these variables was not predicted by any of our possible auxiliary variables (gender, parent education level, preschool experience, or minority status) except in one case: Missingness on teacher ratings of behavioral regulation and interpersonal skills was associated with children having less child care experience. These results are not easily interpretable, however, and given the number of analyses conducted, it is reasonable to assume that at least some, if not most, of the missing data in our analyses are MAR. However, although we included auxiliary variables, the MAR assumption cannot be fully verified, so conclusions must be interpreted cautiously.

Because traditional methods for handling missingness (e.g., listwise deletion or mean imputation) can produce biased parameter estimates (Acock, 2005; McCartney, Burchinal, & Bub, 2006), we used multiple imputation to create 10 data sets with the PROC MI command in SAS. These data sets were used in all further analyses. The PROC MIANALYZE command combines descriptives and correlation results from the imputed data (SAS Institute, 2008). We also conducted all reported analyses with the original data using listwise deletions; patterns were highly similar to results with the multiply imputed data. For consistency, we used the 10 imputed data sets for all further results ($N = 343$). Descriptive statistics by site for all raw data are shown in Table 1, along with the descriptives for the combined sample from the multiple imputed data sets. Table 2 shows correlations for the combined sample from the imputed data.

Results

Three research questions were pursued in this paper. First, we assessed variability and gains in children's performance on a structured observation of behavioral regulation, and construct validity with parent and teacher ratings. Second, we examined predictive validity, assessing whether initial levels of behavioral regulation predicted spring levels of achievement and teacher-ratings of kindergarten functioning at the end of the school year. Third, we assessed whether initial behavioral regulation predicted kindergarten gains in three academic domains.

Variability and Kindergarten Gains in Behavioral Regulation: Relations to Observer Ratings

Descriptive statistics demonstrated considerable individual variability in HTKS scores (and all kindergarten outcomes; see Table 1). For the entire sample, children made significant fall–spring HTKS gains, improving from 27.5 points on average in the fall to 31.3 points in the spring, $t = -8.55$, $p < .01$. Although scores tended to fall above the halfway mark (20 points out of 40), skewness and kurtosis did not indicate a nonnormal distribution (Kline, 2005). Further, score distribution indicated that most children scored within the task range (i.e., not at floor or ceiling levels), and this was consistent across site. In the fall, 95% of scores in Michigan and 100% in Oregon fell between 1 and 39 points; in the spring, 94% of scores from both sites were in this range.

We explored construct validity for the HTKS, including scores for both the English and Spanish versions of the task, using parent

Table 2
Correlations of Background and Fall Independent Variables With Spring Outcomes (Combined Sample: N = 343)

	Spring outcome					
					Teacher-rated classroom	
Variable	HTKS	Mathematics	Literacy	Vocabulary	Behavioral regulation	Interpersonal skills
Fall child age	$.14^{**}$	$.23^{**}$	$.16^{**}$	$.16^{**}$	$.17^{**}$	$.19^{**}$
Gender (girl = 0, boy = 1)	$-.17^{**}$	.04	−.00	−.08	$-.23^{**}$	$-.10^{\dagger}$
Fall prior child care experience	.09	$.13^{*}$	.06	.06	−.06	$-.20^{**}$
Parent education level	$.26^{**}$	$.29^{**}$	$.14^{**}$	$.29^{**}$	.09	.03
Spanish-version HTKS (0 = English version, 1 = Spanish version)	$-.31^{**}$	$-.20^{**}$	$.13^{*}$	$-.26^{**}$	.04	.00
Site (Oregon = 0, Michigan = 1)	$.19^{**}$	$-.11^{*}$	$-.14^{**}$	.06	$-.11^{*}$	.04
Fall						
HTKS	$.59^{**}$	$.49^{**}$	$.29^{**}$	$.29^{**}$	$.29^{**}$	.07
Mathematics	$.41^{**}$	$.68^{**}$	$.52^{**}$	$.44^{**}$	$.33^{**}$	.06
Literacy	$.26^{**}$	$.48^{**}$	$.86^{**}$	$.30^{**}$	$.29^{**}$	$.15^{**}$
Vocabulary	$.37^{**}$	$.44^{**}$	$.29^{**}$	$.72^{**}$	$.13^{*}$	.07

Note. HTKS = head-toes-knees-shoulders task.
$^{\dagger} p < .10$. $^{*} p < .05$. $^{**} p < .01$.

ratings of attentional focusing and inhibitory control, and teacher ratings of behavioral regulation. Parent ratings, completed in the spring and summer before kindergarten, were compared with fall HTKS scores. Children with higher scores on the HTKS in fall of kindergarten received higher ratings on attentional focusing, $r = .25$, $p < .01$; and inhibitory control, $r = .20$, $p < .01$. Teacher ratings, completed in the spring of kindergarten, were compared with spring HTKS scores. Children rated higher in the spring by their kindergarten teachers on the behavioral regulation scale of the CBRS earned higher HTKS scores in the spring, $r = .20$, $p < .01$. Interpersonal skills and spring behavioral regulation were significantly but more weakly related, $r = .11$, $p < .05$. Teacher-rated classroom behavioral regulation and interpersonal skills were significantly correlated, $r = .51$, $p < .01$.

Initial Behavioral Regulation Predicting End-of-Year Achievement and Classroom Functioning

Because children were nested within classrooms, we built models of achievement and teacher-reported classroom functioning using the multiple imputation option in hierarchical linear modeling or HLM (Raudenbush & Bryk, 2002; Raudenbush, Bryk, Cheong, & Congdon, 2004). In HLM, the variance in a child-level outcome is partitioned into child (Level-1) and classroom (Level-2) sources. We first calculated the intraclass coefficient (ICC) to assess the percentage of the variance in each measure at the classroom level. ICCs were calculated for spring achievement in the three academic outcomes (mathematics, literacy, and vocabulary) and for spring teacher ratings of classroom behavioral regulation and interpersonal skills.

Variance at the classroom level was statistically significant at $p < .01$ for achievement: literacy spring level (13%), mathematics gain (14%); and vocabulary gain (14%). For teacher ratings, there was classroom-level variance in classroom behavioral regulation (18%) and interpersonal skills (12%), but not for mathematics and vocabulary level (i.e., the Level-2 variance was not significantly different from 0). This pattern of variance distribution indicates classroom differences in gains for some outcomes and in overall performance levels for others. We proceeded with HLM analyses because we sought to maintain analytic consistency and because classroom differences were prevalent. Further, although on average the number of child participants per classroom was small (3 children per classroom in Michigan; 5 in Oregon), the assumption of independence of observations was violated for participants who shared kindergarten classroom membership. Accounting for nesting with HLM enabled us to obtain accurate standard error estimates, even when the Level-1 units (children) within each Level-2 unit (classroom) were relatively few and varied by classroom (Raudenbush & Bryk, 2002).

To assess predictive validity of the initial behavioral regulation for markers of kindergarten success, we used HLM to model the unique contribution of fall HTKS behavioral regulation to spring levels of achievement and teacher-rated classroom functioning and interpersonal skills. The average spring score on each of the five outcomes was modeled, with child-level predictors (age, gender, parent education, and Spanish-speaking status) and site (Michigan or Oregon, a Level-2 dummy variable) being controlled. The final model is described for mathematics in the appendix. HLM results for spring outcomes are presented in Table 3.

Following recent recommendations about framing educational results to maximize practical interpretability (Hill, Bloom, Black, & Lipsey, 2008), we also calculated effect sizes for significant predictors. For this analysis, we computed *d* by multiplying the coefficient for each predictor by its standard deviation (*SD*) and then dividing by the *SD* of the outcome (Cohen, 1988). Values do not take into account classroom variance components and should

Table 3
Hierarchical Linear Modeling Results for Fall Head-Toes-Knees-Shoulders Task Predicting Spring Levels of Achievement and Classroom Functioning (N = 343)

										Teacher-rated classroom					
	Mathematics			Literacy			Vocabulary			Behavioral regulation			Interpersonal skills		
Variable	Coeff	*df*	*t*	Coeff	*df*	*t*	Coeff	*df*	*t*	Coeff	*df*	*t*	Coeff	*df*	*t*
Fixed effect															
Average level, β_0, γ_{00}	452.35	110	195.50**	409.87	110	90.80**	485.77	110	249.56**	4.18	110	35.54**	3.92	110	34.59**
Michigan, γ_{01}	**−8.69**	**110**	**3.75****	−7.47	110	−1.61	0.76	110	−0.69	−0.11	110	−0.85	0.16	110	1.41
Fall age, β_1, γ_{10}	**7.69**	**336**	**3.33****	8.31	336	1.70†	**6.03**	**336**	**3.23****	**0.29**	**53**	**2.11***	**0.42**	**105**	**3.36****
Male, β_2, γ_{20}	**3.52**	**336**	**2.37***	1.42	336	0.42	−1.18	336	−0.89	**−0.28**	**52**	**−3.16****	−0.12	31	−1.23
Parent education, β_3, γ_{30}	**1.36**	**72**	**3.11****	**3.88**	**336**	**4.87****	**0.88**	**205**	**2.31***	0.04	19	1.19	0.01	30	0.33
Spanish version, β_4, γ_{40}	**−9.96**	**187**	**−1.98***	**48.33**	**336**	**3.05****	**−9.41**	**336**	**−2.39***	0.40	40	1.17	0.11	52	0.34
Fall HTKS, β_5, γ_{50}	**0.74**	**336**	**8.36****	**0.87**	**336**	**4.77****	**0.24**	**336**	**3.87****	**0.02**	**123**	**4.03****	0.00	23	0.21
	Var	*df*	χ^2	Var	*df*	χ^2	Var	*df*	χ^2	Var	*df*	χ^2	Var	*df*	χ^2
Random effects															
Level-2 intercept, u_{0j}	5.14	110	120.77	86.56	110	147.91**	0.21	110	109.51	0.07	110	178.56**	0.04	110	151.77**
Level-1 effects, r_{ij}	170.86			750.23			116.40			0.34			0.34		

Note. Varying *df*s are due to models being estimated from multiple imputed data sets. Significant predictors of the intercepts (average level, γ_{00}) are in bold type. Dummy variables coded as follows: Michigan = 1, Oregon = 0; male = 1, female = 0; Spanish-version HTKS = 1, English version = 0. HTKS = head-toes-knees-shoulders behavioral regulation task; coeff = fixed effect coefficient estimate; var = variance component.
† $p < .10$. * $p < .05$. ** $p < .01$.

be interpreted cautiously. We also calculated the contribution of fall HTKS scores to spring outcomes, on the basis of average time between testing (5.8 months) and average monthly point increases associated with each achievement outcome.

Compared with participants with low fall HTKS scores, children with higher scores earned relatively higher scores on four of the five outcomes. These findings emerged with all other variables held constant at their means or reference values. Several other predictors reached significance or marginal significance, but to streamline the results, we highlight HTKS findings reaching a level of $p < .05$. We also tested whether fall HTKS predicted achievement similarly by site, but none of these interactions were statistically significant.

For mathematics level in the spring, fall HTKS was a significant predictor ($d = .56$), as was parent education ($d = .26$). In practical terms, children scoring 1 *SD* above the mean on fall HTKS, or earning 37 instead of 27 points, were expected to score almost 7 points higher than the mean on the spring Applied Problems subscale, or 459 instead of 452 points. Further, on average, spring scores were about 13 points higher than fall scores (or 13/5.8 = 2.4 points per month); thus, 7 points was the gain associated, with about 3.4 additional months of mathematics learning in kindergarten. This model explained 36% of the variance in Applied Problems level in the spring.

For spring literacy level, again both fall HTKS ($d = .27$) and parent education ($d = .29$) were significant predictors. The literacy gain of 4.2 points associated with scoring 1 *SD* above the mean on fall HTKS was associated with about 1.7 months of kindergarten attendance. The model explained 18% of the variance in literacy.

For vocabulary level in the spring, increased fall HTKS ($d = .19$) and child age ($d = .16$) were significant positive predictors. Children scoring 1 *SD* above the mean fall HTKS were expected to earn vocabulary scores almost 2 points higher than the spring average of 486 points. This increase was associated with the vocabulary learning associated with about 1.9 months of kindergarten. The model explained 18% of the variance in vocabulary.

For teacher-rated classroom behavioral regulation, fall HTKS was the sole statistically significant predictor ($d = .28$). Children with stronger fall HTKS scores (e.g., 37 instead of 27 points) were rated 0.2 points higher than the spring classroom behavioral regulation mean of 4.2. The model explained 17% of the variance in classroom behavioral regulation.

Finally, for spring interpersonal skills, older children received higher ratings than did younger children ($d = .22$); HTKS did not predict interpersonal skills. The model explained 7% of the variance in interpersonal skills.

Initial Behavioral Regulation Predicting Achievement Gains

The final research question addressed whether children entering kindergarten with strong behavioral regulation as assessed by the HTKS would make greater achievement gains. We built three final HLM models, one for each gain score, in which pretest in that domain (to assess whether children's initial status predicted gains over the kindergarten year), and fall HTKS behavioral regulation were controlled. We retained age, gender, parent education, the Spanish-speaking variable, and site because of their theoretical importance and predictive utility. See Table 4 for HLM results for achievement gains.

After including respective fall achievement scores, few predictors continued to explain additional variance in gains. Children with higher initial scores on each assessment made fewer gains, relative to children with lower initial scores. There remained site differences, with Oregon participants making greater gains in mathematics and literacy than Michigan children. Older children and English-speaking participants made greater mathematics gains, relative to younger children and Spanish speakers. Girls made greater vocabulary gains than did boys.

Children's initial HTKS behavioral regulation upon kindergarten entry significantly predicted gains only in mathematics after background variables (see Table 4) were controlled. Participants

Table 4
Hierarchical Linear Modeling Results For Fall Head-Toes-Knees-Shoulders Task Predicting Gains in Achievement (N = 343)

	Mathematics			Literacy			Vocabulary		
Variable	Coeff	*df*	*t*	Coeff	*df*	*t*	Coeff	*df*	*t*
Fixed effect									
Average gain, β_0, γ_{00}	16.11	110	9.74**	35.03	110	14.87**	7.02	110	4.74**
Michigan, γ_{01}	**−5.10**	**110**	**−2.78****	**−5.30**	**110**	**−2.08***	0.31	110	0.21
Fall age, β_1, γ_{10}	**4.84**	**335**	**2.27***	−2.27	335	−0.88	1.27	335	0.95
Male, β_2, γ_{20}	1.82	335	1.58	−0.18	335	−0.10	**−2.25**	**335**	**−2.20***
Parent education, β_3, γ_{30}	0.56	157	1.47	0.61	335	1.19	0.04	130	0.11
Spanish version, β_4, γ_{40}	**−8.24**	**137**	**−2.12***	13.62	335	1.68†	3.62	198	1.06
Initial score, β_5, γ_{50}	**−0.43**	**335**	**−7.86****	**−0.13**	**335**	**−5.56****	**−0.34**	**207**	**−7.83****
Fall HTKS, β_6, γ_{60}	**0.40**	**335**	**5.35****	0.14	335	1.53	0.08	335	1.66
	Var	*df*	χ^2	Var	*df*	χ^2	Var	*df*	x^2
Random effects									
Level-2 intercept, u_{0j}	11.00	110	143.48*	3.96	110	117.12	6.75	110	161.34**
Level-1 effects, r_{ij}	115.64			246.61			59.94		

Note. Significant predictors of the intercept (average gain, γ_{00}) are in bold type. Dummy variables coded as follows: Michigan = 1, Oregon = 0; male = 1, female = 0; Spanish-version HTKS = 1, English version = 0. HTKS = head-toes-knees-shoulders behavioral regulation task; coeff = fixed effect coefficient estimate; var = variance component.
† $p < .10$. * $p < .05$. ** $p < .01$.

scoring 1 *SD* above the fall average on the HTKS task gained almost 4 additional points in mathematics. That is, having strong fall HTKS behavioral regulation predicted greater mathematics gains, after accounting for achievement at the beginning of the year (d = .31).

Discussion

This initial examination of the HTKS structured observation of behavioral regulation revealed significant variability and construct validity with parent ratings of kindergarteners' attentional focusing and inhibitory control and teacher ratings of behavioral regulation. Participants with higher levels of behavioral regulation in the fall of kindergarten reached higher levels of mathematics, literacy, and vocabulary skills and teacher-rated classroom behavioral regulation in the spring. Finally, compared with children entering kindergarten with lower behavioral regulation, those with stronger initial performance showed greater gains in mathematics but not in literacy or vocabulary.

Evaluating a New Measure of Behavioral Regulation

The HTKS structured observational assessment successfully extended the HTT task, established with 3- to 6-year olds (Ponitz et al., 2008). With its requirement to remember and respond to two rules, the simpler HTT task proved most appropriate for 4- and 5-year-olds, consistent with developmental research on the age at which children can cognitively manage multiple rules (Diamond et al., 2002; Müller, Zelazo, Hood, Leone, & Rohrer, 2004; Rennie et al., 2004). In this study, we examined 5- and 6-year-olds' performance on a complex version of the task introducing two additional commands. Unlike many tasks with this age group, the HTKS showed adequate reliability and variability, and performance was corroborated by two observer reports of behavior (S. M. Carlson, 2005; Rothbart et al., 2006).

Children earning higher scores on the HTKS received higher parent ratings of attentional focusing and inhibitory control, two cognitively based skill components proposed to be tapped by the HTKS. Teachers also reported children with high HTKS performance as showing self-directed task-related behavior in the classroom. Correlations with parent and teacher ratings were modest, although consistent with those found in other studies that compared caregiver questionnaires evaluating behavior with direct assessments of behavior (Eisenberg et al., 2000; Howse, Calkins, et al., 2003; Smith-Donald, Raver, Hayes, & Richardson, 2007). Situational or contextual differences in behavior may have contributed to these patterns, pointing to the importance of multiple assessments of school readiness (Mashburn et al., 2006; Meisels, 2006).

We also found HTKS performance improved significantly from fall to spring. Scores tended to cluster above the mean, yet the majority of children at both sites scored below ceiling at each time point. It is possible that some of this improvement could have been due to children's familiarity with the task and the HTT they were administered in preschool, although other research with the HTT indicates that development, rather than number of administrations, predicts improvement on the task (Ponitz et al., 2008). HTKS performance also predicted levels of achievement and teacher-rated classroom behavioral regulation beyond the contributions of age, gender, parent education, ethnicity, and site. Taken together, results suggest the potential of the HTKS as a viable measure of behavioral regulation in children through 6.5 years of age, to be used in concert with observer reports.

Contributions of Behavioral Regulation to Academic Aspects of Kindergarten Functioning

Our results reinforced prior work that connects behavioral regulation with achievement and specific aspects of classroom functioning during early childhood (Fantuzzo, Perry, & McDermott, 2004; Gathercole, Alloway, Willis, & Adams, 2006; Howse, Calkins, et al., 2003; McClelland et al., 2000; Ready et al., 2005). Children's performance on the HTKS positively predicted their mathematics, literacy, and vocabulary achievement in the spring. The magnitude of contributions of HTKS were twice that of parent education for mathematics and were similar to the effect sizes of parent education and chronological age for literacy and vocabulary, respectively. Kindergarteners who excelled on the HTKS at school entry also earned higher ratings of classroom behavioral regulation, but fall HTKS scores did not predict interpersonal skills ratings. This pattern of findings adds to accumulating evidence that links strong early childhood performance on similar tasks to higher achievement (Blair & Razza, 2007; McClelland, Cameron, Connor, et al., 2007). This superiority likely arises because the skills required by tasks like the HTKS (attentional focusing, working memory, and inhibitory control) also help children perform well scholastically and on academic classroom tasks (Bull & Scerif, 2001; Howse, Calkins, et al., 2003; McClelland et al., 2000).

Prior work has suggested that interpersonal expertise taps the same or similar regulatory processes involved in behavioral regulation (i.e., effortful control), so we expected links between the HTKS and teacher ratings of interpersonal skills (Calkins, 2007; Fox & Calkins, 2003; Kiley-Brabeck & Sobin, 2006; Patrick, 1997; Pettit, Harrist, Bates, & Dodge, 1991). In contrast to our expectations, fall HTKS scores did not predict teachers' ratings of children's social interactions in the classroom after we controlled for age and demographic factors. Whereas interpersonal skills require regulating emotions and impulses in social situations, such as not lashing out at a peer who accidentally knocks over one's blocks, modulating intense emotionality is not a distinct feature of the HTKS (Howse, Calkins, et al., 2003; McClelland, Cameron, Wanless, et al., 2007; Trentacosta & Izard, 2007). With its four rules to remember, the HTKS may instead draw heavily on working memory and attentional focusing in kindergarten, executive function components not as clearly tied to interpersonal skills as inhibitory control but important for achievement (Duncan et al., 2007; Kochanska & Knaack, 2003; Pickering & Gathercole, 2004). A recent report suggests children's attention skills more strongly predict later achievement compared with interpersonal social skills, but this research relied heavily on observer ratings of children's behavior (Duncan et al., 2007). A study in which the contributions of the three cognitive components to HTKS performance, achievement, and interpersonal skills are compared would help illuminate these results.

Behavioral Regulation and Mathematics Gains

Our results contrasted with our prior work showing the HTT measure of behavioral regulation predicted gains in multiple achievement domains in prekindergarten (McClelland, Cameron, Connor, et al., 2007). Our kindergarten-age participants with strong initial behavioral regulation made gains in mathematics, but beginning-of-year behavioral regulation did not significantly predict literacy or vocabulary gains. During preschool, strong behavioral regulation may facilitate learning in diverse aspects of early achievement. Growing evidence suggests these relations become domain specific as children enter kindergarten, where certain content may differentially challenge regulatory skills. For example, one recent study found kindergarteners with strong inhibitory control exhibited higher performance than peers with low inhibitory control on concurrent measures of mathematics, letter knowledge, and phonemic awareness (Blair & Razza, 2007). Strong prior inhibitory control (measured in preschool) independently predicted kindergarten mathematics achievement but did not make unique contributions to performance on either literacy task. Our results provide further evidence for the generalizability of these domain-specific patterns of association, until now found primarily with disadvantaged samples (Blair & Razza, 2007; Espy et al., 2004). The present study included many children from middle-SES families, suggesting robust connections between behavioral regulation and mathematics learning.

Why should behavioral regulation be more important for making progress in mathematics compared to literacy? One possibility is that by kindergarten, the cognitive processes necessary for successful reading and letter recognition become more automatic for literacy than for mathematics, in part because children are exposed to larger amounts of literacy instruction relative to other subjects (Connor, Morrison, & Slominski, 2006; Miller, Kelly, & Zhou, 2005; NICHD ECCRN, 2002). Strong behavioral regulation at school entry may help children learn from relatively few learning opportunities provided in mathematics. In line with this explanation, behavioral regulation did not contribute as strongly to literacy, in which children may have received more instruction; this would reduce the variance attributable to prior child factors like behavioral regulation. Features of the assessments that we used may also help explain these results: Whereas advanced items on the Letter–Word Identification and Picture Vocabulary subtests increased in content-related difficulty but not in regulatory demand, progressing on the Applied Problems subtest required children to hold more information in mind and inhibit the tendency to attend to incorrect solutions (Blair et al., 2007; Blair & Razza, 2007; Bull & Scerif, 2001). Thus, the cognitive and behavioral skills required by the HTKS, such as working memory, may be especially important for developing mathematics skills as measured by the Applied Problems subscale (Espy et al., 2004).

When other factors were controlled, non-Spanish-speaking children in Oregon (the reference group for analyses) showed greater mathematics and literacy gains relative to Michigan participants. In contrast, Spanish-speaking participants, who were from especially disadvantaged backgrounds compared with other children in Oregon (even other Head Start children), showed significantly lower gains in mathematics when site was controlled. Two possibilities may explain these differences. First, the non-Spanish-speaking families in Oregon were highly educated, and many were recruited from the nearby university, whereas families in Michigan with college degrees worked at diverse occupations. Although we did not measure family learning environments, it is possible that the more educated, academically oriented subsample of Oregon families provided more stimulating learning environments for their kindergarteners, compared with Michigan families with the same levels of parent education. Second, there was greater socioeconomic diversity found in the Oregon sample overall, with half the sample drawn from Head Start preschools and half from non–Head Start preschools. Greater variability in the Oregon sample may have contributed to greater gains among those participants, relative to Michigan children.

Implications and Limitations of the Study

Considering the importance of kindergarten for multiple developmental outcomes, these results provide initial evidence for the HTKS as a reliable and ecologically valid assessment of behavioral regulation during this time (Aram, 2005; McClelland et al., 2006; Vitaro et al., 2005). However, the HTKS needs further construct validation against other established measures, including data on problem behaviors arising from regulatory deficits (Shaw et al., 2003). Future research could provide more robust evidence through examination of whether the HTKS predicts achievement beyond the contribution of other assessments, including teacher reports or individualized assessments such as prekindergarten HTT scores.[1] Hence, much work remains in the area of creation and systematic evaluation of tasks appropriate for young children that are sensitive to individual differences and meaningfully predictive of development throughout the early years of school (Hongwanishkul, Happaney, Lee, & Zelazo, 2005; McCabe, Cunnington, & Brooks-Gunn, 2004; McCabe, Hernandez, Lara, & Brooks-Gunn, 2000; Meisels, 2006; Raver, 2002). Future studies should also incorporate multiple time points to explore child and classroom differences in growth rate (with three time points) and acceleration (with four or more time points) in achievement domains (Raudenbush & Bryk, 2002; Rogosa, Brandt, & Zimowski, 1982; Willett, 1997).

Findings provide preliminary support for teachers to promote behavioral regulation in their efforts to strengthen early achievement and school success (McClelland et al., 2006; Ponitz et al., 2008; Rimm-Kaufman & Chiu, 2007). However, because our study is correlational, a third, unmeasured variable or singular cognitive skill, such as processing speed, could explain the HTKS's predictive power for outcomes (Demetriou, Christou, Spanoudis, & Platsidou, 2002; Luna, Garver, Urban, Lazar, & Sweeney, 2004). Unmeasured auxiliary variables underlying whether children were missing data may also play a role in our findings. Random assignment within the context of an intervention to teach children behavioral regulation would help address these issues and could establish causal links between regulatory competence and gains in academic skills like mathematics in kindergarten. One such study has had notable success in teaching the executive function skills underlying behavioral regulation to preschoolers (Diamond, Barnett, Thomas, & Munro, 2007).

Exploring causal links between behavioral regulation and academic achievement seems especially important in light of recent reports that document the difficulty that many children have with the school transition (Gilliam & Shahar, 2006; McClelland et al., 2000; Rimm-Kaufman, Pianta, & Cox, 2000). Ecological perspectives can guide efforts to design learning environments that better meet children's needs as they enter formal schooling, rather than focusing responsibility for a successful school transition solely on the child (Kemp & Carter, 2000; Pianta, Cox, & Snow, 2007; Stipek, 2002). Mounting

[1] We thank an anonymous reviewer for this point.

evidence has shown that understanding the multiple ecologies surrounding development (e.g., parenting, home–school partnership, and teacher–child relationships) is more effective than concentrating on isolated skills and skill deficits within individuals alone (Pianta & Rimm-Kaufman, 2006). Translating these suggestions into practice, such as supporting children's emerging regulatory competence with appropriate opportunities to exercise behavioral regulation in the classroom, would be an important step toward improving developmental trajectories in school adaptation and achievement.

A final limitation is the lack of observational data on classroom time devoted to different subjects, so we could not examine whether differential mathematics versus literacy and language experiences contributed to findings about gains in these domains. Nonetheless, classrooms were drawn from two socioeconomically and geographically distinct areas. They represented a range of typical American early elementary settings, which tend to devote little time to mathematics (Hamre & Pianta, 2007; Miller et al., 2005). Furthermore, associations between behavioral regulation and outcomes were consistent across sites.

Summary

Results of the present study suggest looking more deeply at the role of behavioral regulation in kindergarten and further exploring its relation to achievement, especially in mathematics (Duncan et al., 2007). Adaptive, healthy development in our society increasingly means completing higher education; working jobs requiring technological, organizational, and interpersonal expertise; and managing multiple career and family commitments (Heckman, Stixrud, & Urzua, 2006; Miller et al., 2005; Muraven & Baumeister, 2000). Children's behavioral regulation at the start of formal schooling is uniquely associated with success in academic domains and also predicts improvement over the kindergarten year in mathematics, an instructionally underemphasized subject in the United States. Helping young students successfully manage their behavior in school will undoubtedly have benefits beyond mathematics achievement.

References

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

Adams, A-M., Bourke, L., & Willis, C. (1999). Working memory and spoken language comprehension in young children. *International Journal of Psychology, 34,* 364–373.

Adams, J. W., & Snowling, M. J. (2001). Executive function and reading impairments in children reported by their teachers as "hyperactive." *British Journal of Developmental Psychology, 19,* 293–306.

Aram, D. (2005). Continuity in children's literacy achievements: A longitudinal perspective from kindergarten to school. *First Language, 25,* 259–289.

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

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

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., Knipe, H., Cummings, E., Baker, D. P., Gamson, D., Eslinger, P., et al. (2007). A developmental neuroscience approach to the study of school readiness. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 149–174). Baltimore: Brookes.

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

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45,* 743–754.

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

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.

Calkins, S. D. (2007). The emergence of self-regulation: Biological and behavioral control mechanisms supporting toddler competencies. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and transformations* (pp. 261–284). New York: Guilford.

Carlson, M. J., & Corcoran, M. E. (2001). Family structure and children's behavioral and cognitive outcomes. *Journal of Marriage and Family, 63,* 779–792.

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

Christian, K., Bachman, H. J., & Morrison, F. J. (2001). Schooling and cognitive development. In *Environmental effects on cognitive abilities* (pp. 287–335). Mahwah, NJ: Erlbaum.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology, 98,* 665–689.

Connor, C. M., Ponitz, C. C., Phillips, B., Travis, M., Glassney, S., & Morrison, F. J. (2007). *First graders' literacy and self-regulation gains: The effect of individualizing student instruction*. Manuscript in preparation.

Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society for Research in Child Development, 67,* vii–154.

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., et al. (2003). Preschool emotional competence: Pathway to social competence. *Child Development, 74,* 238–256.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007, November 30). Preschool program improves cognitive control. *Science, 318,* 1387–1388.

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.

Diener, M. L., & Kim, D.-Y. (2004). Maternal and child predictors of preschool children's social competence. *Journal of Applied Developmental Psychology, 25,* 3–24.

Dobbs, J., Doctoroff, G. L., Fisher, P. H., & Arnold, D. H. (2006). The association between preschool children's socio-emotional functioning and their mathematical skills. *Journal of Applied Developmental Psychology, 27,* 97–108.

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.

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.

Eiden, R. D., Edwards, E. P., & Leonard, K. E. (2007). A conceptual model for the development of externalizing behavior problems among kindergarten children of alcoholic families: Role of parenting and children's self-regulation. *Developmental Psychology, 43,* 1187–1201.

Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Losoya, S., Murphy, B. C., et al. (2000). Prediction of elementary school children's externalizing problem behaviors from attention and behavioral regulation and negative emotionality. *Child Development, 71,* 1367–1382.

Eisenberg, N., Smith, C. L., Sadovsky, A., & Spinrad, T. L. (2004). Effortful control: Relations with emotion regulation, adjustment, and socialization in childhood. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 259–282). New York: Guilford.

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.

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.

Fantuzzo, J., Perry, M. A., & McDermott, P. (2004). Preschool approaches to learning and their relationship to other relevant classroom competencies for low-income children. *School Psychology Quarterly, 19,* 212–230.

Foorman, B. R., Francis, D. J., Fletcher, J. M., & Lynn, A. (1996). Relation of phonological and orthographic processing to early reading: Comparing two approaches to regression-based, reading-level-match designs. *Journal of Educational Psychology, 88,* 639–652.

Fowler, M. G., & Cross, A. W. (1986). Preschool risk factors as predictors of early school performance. *Journal of Developmental & Behavioral Pediatrics, 7,* 237–241.

Fox, N. A., & Calkins, S. D. (2003). The development of self-control of emotion: Intrinsic and extrinsic influences. *Motivation and Emotion, 27,* 7–26.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A.-M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93,* 265–281.

Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology, 92,* 377–390.

Gilliam, W. S., & Shahar, G. (2006). Preschool and child care expulsion and suspension: Rates and predictors in one state. *Infants and Young Children, 19,* 228–245.

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

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49–83). Baltimore: Brookes.

Heckman, J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24,* 411–482.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2,* 172–177.

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology, 28,* 617–644.

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., White, A., Sharpen, J., & Dunn, J. (2000). Antisocial, angry, and unsympathetic: "Hard-to-manage" preschoolers' peer problems and possible cognitive influences. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41,* 169–179.

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.

Kagan, J. (1989). Temperamental contributions to social behavior. *American Psychologist, 44,* 668–674.

Kail, R. V. (2003). Information processing and memory. In M. H. Bornstein, L. Davidson, C. L. M. Keyes, & K. A. Moore (Eds.), *Crosscurrents in contemporary psychology.* Mahwah, NJ: Erlbaum.

Kemp, C., & Carter, M. (2000). Demonstration of classroom survival skills in kindergarten: A five year transition study of children with intellectual disabilities. *Educational Psychology, 20,* 393–411.

Kiley-Brabeck, K., & Sobin, C. (2006). Social skills and executive function deficits in children with the 22q11 deletion syndrome. *Applied Neuropsychology, 13,* 258–268.

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

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality, 71,* 1087–1112.

Kochanska, G., Murray, K., Jacques, T. Y., & Koenig, A. L. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67,* 490–507.

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.

Lawrence, V., Houghton, S., Tannock, R., Douglas, G., Durkin, K., & Whiting, K. (2002). ADHD outside the laboratory: Boys' executive function performance on tasks in videogame play and on a visit to the zoo. *Journal of Abnormal Child Psychology, 30,* 447–462.

Lengua, L. J., Honorado, E., & Bush, N. R. (2007). Contextual risk and parenting as predictors of effortful control and social competence in preschool children. *Journal of Applied Developmental Psychology, 28,* 40–55.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75,* 1357–1372.

Müller, U., Dick, A. S., Gela, K., Overton, W. F., & Zelazo, P. D. (2006). The role of negative priming in preschoolers' flexible rule use on the dimensional change card sort task. *Child Development, 77,* 395–412.

Müller, U., Zelazo, P. D., Hood, S., Leone, T., & Rohrer, L. (2004). Interference control in a new rule use task: Age-related changes, labeling, and attention. *Child Development, 75,* 1594–1609.

Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. *Journal of Psychoeducational Assessment, 24,* 367–380.

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

Matthews, J., Ponitz, C. C., & Morrison, F. J. (in press). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology.*

McCabe, L. A., Cunnington, M., & Brooks-Gunn, J. (2004). The development of self-regulation in young children: Individual characteristics and environmental contexts. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 340–356). New York: Guilford.

McCabe, L. A., Hernandez, M., Lara, S. L., & Brooks-Gunn, J. (2000). Assessing preschoolers' self-regulation in homes and classrooms: Lessons from the field. *Behavioral Disorders, 26,* 53–69.

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.

McCartney, K., Burchinal, M. R., & Bub, K. L. (2006). Best practices in quantitative methods for developmentalists. *Monographs of the Society for Research in Child Development, 71*(3).

McClelland, M. M. (2007). *Interrater reliability for the head-to-toes task.* Manuscript in preparation.

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., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43,* 947–959.

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). Greenwich, CT: Information Age.

McClelland, M. M., & Morrison, F. J. (2003). The emergence of learning-related social skills in preschool children. *Early Childhood Research Quarterly, 18,* 206–224.

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.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53,* 185–204.

Meisels, S. J. (2006). *Accountability in early childhood: No easy answers* (Occasional Paper No. 6). Chicago: Erikson Institute Herr Research Center for Children & Social Policy. Available from http://www.erikson.edu/research.asp?file=publications_topic#assess

Merrell, C., & Tymms, P. (2007). Identifying reading problems with computer-adaptive assessments. *Journal of Computer Assisted Learning, 23,* 27–35.

Miech, R., Essex, M. J., & Goldsmith, H. H. (2001). Socioeconomic status and the adjustment to school: The role of self-regulation during early childhood. *Sociology of Education, 74,* 102–120.

Miller, K. F., Kelly, M., & Zhou, X. (2005). Learning mathematics in China and the United States: Cross-cultural insights into the nature and course of preschool mathematical development. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 163–177). New York: Psychology Press.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126,* 247–259.

Murray, A., Sektnan, M., & McClelland, M. M. (2006, April). Relations between effortful control and social–emotional competencies across multiple measures. Paper presented at the Conference on Human Development. Louisville, KY.

National Institute of Child Health and Human Development Early Childcare Research Network. (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behaviors. *The Elementary School Journal, 102,* 367–387.

National Institute of Child Health and Human Development Early Childcare Research Network. (2003). Do children's attention processes mediate the link between family predictors and school readiness? *Developmental Psychology, 39,* 581–593.

Patrick, H. (1997). Social self-regulation: Exploring the relations between children's social relationships, academic self-regulation, and school performance. *Educational Psychologist, 32,* 209–220.

Pettit, G. S., Harrist, A. W., Bates, J. E., & Dodge, K. A. (1991). Family interaction, social cognition and children's subsequent relations with peers at kindergarten. *Journal of Social and Personal Relationships, 8,* 383–402.

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

Pianta, R. C., & Rimm-Kaufman, S. (2006). The social ecology of the transition to school: Classrooms, families, and children. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development.* (pp. 490–507). Malden, MA: Blackwell.

Pickering, S. J., & Gathercole, S. E. (2004). Distinctive working memory profiles in children with special educational needs. *Educational Psychology, 24,* 393–408.

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.

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment, 87,* 102–112.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danish Institute for Educational Research.

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.

Raver, C. C. (2002). Emotions matter: Making the case of the role of young children's emotional development for early school readiness. *SRCD Social Policy Report, 16*(3).

Raver, C. C., Blackburn, E. K., Bancroft, M., & Torp, N. (1999). Relations between effective emotional self-regulation, attentional control, and low-income preschoolers' social competence with peers. *Early Education and Development, 10,* 333–350.

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

Rennie, D. A. C., Bull, R., & Diamond, A. (2004). Executive functioning in preschoolers: Reducing the inhibitory demands of the dimensional change card sort task. *Developmental Neuropsychology, 26,* 423–443.

Rimm-Kaufman, S. E., & Chiu, Y. J. I. (2007). Promoting social and academic competence in the classroom: An intervention study examining the contribution of the Responsive Classroom approach. *Psychology in the Schools, 44,* 397–413.

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., & Jones, L. B. (1998). Temperament, self-regulation and education. *School Psychology Review, 27,* 479–491.
Rothbart, M. K., & Posner, M. I. (2005). Genes and experience in the development of executive attention and effortful control. *New Directions for Child and Adolescent Development, 109,* 101–108.
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 Publishing.
Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42,* 1029–1040.
Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 14391–14936.
SAS Institute. (2008). The MIANALYZE procedure. In *SAS/STAT 9.2 user's guide* (pp. 3834–3883). Cary, NC: Author. Retrieved from http://support.sas.com/documentation/cdl/en/statugmianalyze/61806/PDF/default/statugmianalyze.pdf
Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177.
Schwebel, D. C. (2004). Temperamental risk factors for children's unintentional injury: The role of impulsivity and inhibitory control. *Personality and Individual Differences, 37,* 567–578.
Seppanen, P. S., Godin, K. W., Metzger, J. L., Bronson, M. B., & Cichon, D. J. (1993). *Observational study of Chapter 1– funded early childhood programs* (Draft final report). Hampton, NH: RMC Research.
Shaw, D. S., Gilliom, M., Ingoldsby, E. M., & Nagin, D. S. (2003). Trajectories leading to school-age conduct problems. *Developmental Psychology, 39,* 189–200.
Shonkoff, J. P., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development.* Washington, DC: National Academy Press.
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.
Snow, K. L. (2007). Integrative views of the domains of child function: Unifying school readiness. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 197–216). Baltimore: Brookes.
St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59,* 745–759.
Stipek, D. J. (2002). At what age should children enter kindergarten? *SRCD Social Policy Report, 16,* 3–16.
Trentacosta, C. J., & Izard, C. E. (2007). Kindergarten children's emotion competence as a predictor of their academic competence in first grade. *Emotion, 7,* 77–88.
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.
Wachs, T. D., Gurkas, P., & Kontos, S. (2004). Predictors of preschool children's compliance behavior in early childhood classroom settings. *Journal of Applied Developmental Psychology, 25,* 439–457.
Willett, J. B. (1997). Measuring change: What individual growth modeling buys you. In E. Amsel & K. A. Reninger (Eds.), *Change and development* (pp. 213–243). Mahwah, NJ: Erlbaum.
Woodcock, R. W., & Mather, N. (2000). *Woodcock–Johnson Psychoeducational Battery III.* Itasca, IL: Riverside.
Woodcock, R. W., & Muñoz-Sandoval, A. F. (1996). *Batería Woodcock–Muñoz—Revisada: Pruebas de Aprovechamiento* [Woodcock–Muñoz Review Battery: Tests of Achievement]. Itasca, IL: Riverside.
Yuan, Y. C. (2005). *Multiple imputation for missing data: Concepts and new developments* (P267–25). Rockville, MD: SAS Institute. Retrieved from http://support.sas.com/rnd/app/papers/multipleimputation.pdf
Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1,* 198–226.
Zelazo, P. D., Müller, U., & Goswami, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 445–469). Malden, MA: Blackwell.

Appendix

Final Hierarchical Linear Model for Spring Mathematics Level

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(\text{age}) + \beta_{2j}\,(\text{male}) + \beta_{3j}\,(\text{parent education}) + \beta_{4j}\,(\text{Spanish version}) + \beta_{5j}\,(\text{fall behavioral regulation}) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\,(\text{site}) + u_j$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

The mathematics level in the spring for child *i* in classroom *j* includes the intercept or overall average gain, plus the contributions of age, being male, parent education, receiving the Spanish version of the HTKS, and fall behavioral regulation, plus error for the individual child (*rij*). The intercept is further defined at Level 2 as the mean of the group means, plus the effect of site, plus error for classroom *j* (*uj*). Thus intercept differences were allowed to vary across classrooms. The effects of $\beta_1 - \beta_5$ were similar at Level 2, and so were fixed.

Received May 2, 2007
Revision received December 10, 2008
Accepted January 12, 2009 ■