Published in final edited form as:
Psychol Assess. 2011 June ; 23(2): 364–378. doi:10.1037/a0021768.

Measuring behavioral regulation in four societies

Shannon B. Wanless,
Curry School of Education, University of Virginia

Megan M. McClelland,
Human Development & Family Sciences, Oregon State University

Alan C. Acock,
Human Development & Family Sciences, Oregon State University

Claire C. Ponitz,
Curry School of Education, University of Virginia

Seung-Hee Son,
Department of Child Development and Family Studies, Purdue University

Xuezhao Lan,
Graduate School of Education, Harvard University

Frederick J. Morrison,
Combined Program in Education and Psychology, University of Michigan

Jo-Lin Chen,
Department of Child & Family Studies, Fu Jen Catholic University, Taiwan

Taiwan Fu-Mei Chen,
Department of Child & Family Studies, Fu Jen Catholic University, Taiwan

Taiwan Kangyi Lee,
Department of Child Development and Family Studies, Seoul National University, South Korea

Miyoung Sung, and
Department of Child Studies, Seokyeong University, South Korea

Su Li
Institute of Psychology, Chinese Academy of Sciences, China

Abstract

The present study examined the psychometric properties of scores from a direct measure of behavioral regulation, the Head-Toes-Knees-Shoulders task (HTKS) with 3- to 6-year-old children in the U.S., Taiwan, South Korea, and China. Specifically, we investigated (1) the nature and variability of HTKS scores including relations to teacher-rated classroom behavioral regulation, and (2) relations between the HTKS and early mathematics, vocabulary, and literacy skills. Higher HTKS scores were significantly related to higher teacher ratings of classroom behavioral regulation in the U.S. and South Korea but not in Taiwan and China. Also, higher HTKS scores

Correspondence concerning this article should be addressed to: Shannon Wanless, Center for Advanced Study of Teaching & Learning, University of Virginia, 350 Old Ivy Way, Suite 300, Charlottesville, VA 22903.

were significantly related to higher early mathematics, vocabulary, and literacy skills beyond the influence of demographic variables and teacher-rated classroom behavioral regulation. These initial findings suggest that HTKS scores may be interpreted as reflecting early behavioral regulation in these four societies, and that behavioral regulation is important for early academic success in the U.S. and in Asian countries.

Keywords

Behavioral regulation; academic achievement; measurement; Taiwan; South Korea; China; preschool

The United States (U.S.) repeatedly scores below average on international academic assessments, while societies including Taiwan, South Korea, and China, consistently score at or near the top (Baldi, Jin, Skemer, Green, & Herget, 2007; Mullis, Martin, Gonzalez, & Chrostowski, 2004; Stevenson, et al., 1990). In an effort to better understand how to improve U.S. children's academic skills, many researchers have examined factors predicting achievement in high-achieving societies (Huntsinger, Jose, Liaw, & Ching, 1997; Stevenson, Chen, & Lee, 1993). These studies, however, have not measured behavioral regulation (the integration of attention, working memory, and inhibitory control), which significantly predicts academic achievement in the U.S. throughout preschool and elementary school (Blair & Razza, 2007; McClelland, Acock, & Morrison, 2006; McClelland, et al., 2007; Ponitz, McClelland, Matthews, & Morrison, 2009). Moreover, few reliable and valid measures of behavioral regulation have been developed and evaluated outside of the United States (U.S.). In the present study, scores from a direct measure of behavioral regulation, the Head-Toes-Knees-Shoulders task (HTKS; Ponitz, et al., 2009), were examined with children in the U.S., Taiwan, South Korea, and China, to investigate reliability and validity including relations to early achievement.

Children in Taiwan, South Korea, and China are intently focused on educational success and their daily schedules are typically filled with extracurricular academic classes and long hours of studying to improve their exams scores and class rankings (Bao, 2004; Dwyer, 2004; Yi & Wu, 2004). Despite this emphasis on school success, reliable and valid behavioral assessments for young children in these societies are limited. Early childhood experts in Asia have recently called for the development of culturally appropriate tools with scores that reliably and validly reflect learning behaviors for these samples (Kim, Lee, Suen, & Lee, 2003; Tsai, McClelland, Pratt, & Squires, 2006). Given the importance of behavioral regulation for children's academic success in the U.S., identifying a measure with scores that reflect behavioral regulation would be especially valuable in societies such as Taiwan, South Korea, and China where academic success is a heavily emphasized cultural expectation.

Defining Behavioral Regulation

In the present study, we define behavioral regulation, an aspect of self-regulation, as the integration of cognitive processes including attention, working memory, and inhibitory control (McClelland, et al., 2007; McClelland & Wanless, 2008). Behavioral regulation is especially relevant in school contexts. For example, a child with strong behavioral regulation can remember and follow a classroom rule, such as waiting for their turn at the water fountain, rather than using a more dominant response, such as cutting in line. Attention, working memory, and inhibitory control individually and collectively contribute to behavioral regulation and to the school success of young children.

Attention is defined as the ability to focus on a task while ignoring distractions (Rothbart & Posner, 2005; Rueda, Posner, & Rothbart, 2005). Research in the U.S. suggests that young children need to attend to activities amidst distractions in order to succeed in vocabulary acquisition and particularly in mathematics (Dixon Jr. & Salley, 2007;G. J. Duncan, et al., 2007; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; McClelland, 2009). Working memory is the ability to remember and apply information while encountering and processing new stimuli (Gathercole & Pickering, 2000). This skill has been positively related to mathematics and language skills (Adams, Bourke, & Willis, 1999; Espy, et al., 2004) and is particularly important during the transition to formal schooling (Senn, Espy, & Kaufmann, 2004). Finally, inhibitory control is the ability to suppress one response in favor of a more non-dominant behavior (Dowsett & Livesey, 2000). Preschoolers are often expected to use inhibitory control to follow classroom rules and this skill has been linked to mathematics and early literacy outcomes (Blair & Razza, 2007; van der Schoot, et al., 2004). Two aspects of behavioral regulation, attention and inhibitory control, have also been included in the construct of effortful control in the temperament literature, which includes emotional aspects of regulation (Rothbart & Sheese, 2007). Further, effortful control has been found to relate positively to academic achievement and grade point average (Checa, Rodríguez-Bailón, & Rueda, 2008; Deater-Deckard, Mullineaux, Petrill, & Thompson, 2009; Eisenberg, Sadovsky, & Spinrad, 2005; Liew, McTigue, Barrois, & Hughes, 2008; Zhou, Main, & Wang, 2010) as well as social development (Kochanska, Murray, & Harlan, 2000). We focus on the behavioral aspects of regulation (e.g., attention, working memory, inhibitory control) because they have been found to be most relevant to children's academic success (Howse, et al., 2003; McClelland & Ponitz, 2010; McClelland, Ponitz, Messersmith, & Tominey, in press).

In the present study, we focus on behavioral regulation as the integration of attention, working memory, and inhibitory control because children in early learning settings are expected to orchestrate these skills for learning. This may be particularly true in Asia where classroom characteristics such as low teacher-child ratios and frequent teacher lectures place high demands on behavioral regulation (Hsieh, 2004; Kim, et al., 2003; Pang & Richey, 2007). In the present study, we examined the HTKS task of behavioral regulation which integrates attention, working memory, and inhibitory control in a simple game.

Behavioral Regulation and Early Academic Achievement

Behavioral regulation predicts early academic success in preschool and early elementary school concurrently and over time in the U.S. and in Asia (Blair & Razza, 2007; Howse, et al., 2003; McClelland, et al., 2007; McClelland, Morrison, & Holmes, 2000; Ponitz, et al., 2009). For example, in the U.S., one study found that gains in behavioral regulation, measured by the HTKS, over the prekindergarten year significantly predicted improvement in early mathematics, vocabulary, and early literacy skills (McClelland, et al., 2007). In Taiwan, four-year-olds' behavioral regulation on the HTKS also significantly predicted early mathematics and vocabulary skills (Wanless, McClelland, Acock, Chen, & Chen, in press). Higher effortful control, which taps overlapping components of behavioral regulation, was also found to relate to higher grade point average in China (Zhou, et al., 2010). Moreover, in some studies, the relation between behavioral regulation and early academic achievement has been particularly pronounced for early mathematics skills (Blair & Razza, 2007;G. J. Duncan, et al., 2007; Ponitz, et al., 2009).

In the U.S., research using the HTKS to measure behavioral regulation has found that higher kindergarten behavioral regulation in the fall significantly predicted higher mathematics, vocabulary, and early literacy scores in the spring (Ponitz, et al., 2009). For U.S. first graders, similar relations were present for scores on the HTKS predicting reading comprehension and vocabulary skills (Connor, et al., in press). Taken together, these studies

suggest that behavioral regulation measured by the HTKS is a unique and significant predictor of early school success, in Western and perhaps in Eastern societies. This research, however, has not systematically examined the use of the HTKS in multiple societies by using participants of similar ages and including similar outcomes.

In addition, when examining links between direct measures of behavioral regulation and achievement, it is important to control for teacher's perceptions of children's classroom behavioral regulation (Matthews, Ponitz, & Morrison, 2009; Ponitz, et al., 2009). This is based on prior research suggesting that teachers' expectations can influence children's academic outcomes (Rosenthal, 2002). For example, in one meta-analysis of studies conducted with first through seventh grade teachers, those who had very little previous experience with children were found to be more easily influenced by proposed expectations, which in turn had an average effect of one quarter standard deviation on children's IQ scores assessed months later (Raudenbush, 1994). This effect was particularly strong in first and second grades. Since teacher perceptions can influence children, they were controlled for in the present study.

Considering Cultural Context

One theoretical tenet that guides this study is the belief that patterns of development are contextually specific and a skill such as behavioral regulation must be examined within each society in order to understand its unique properties and meaning (Cole, 1996; Shweder, et al., 1998). Culturally specific parenting and teaching practices vary across the U.S., Taiwan, South Korea, and China, and may lead to children having different levels of behavioral regulation. For example, one study of effortful control in the U.S. and China found that although similar subscales of effortful control were found in both cultures, children's experiences may have lead to differences found by the researchers such as differences by gender and with relations among subscales (Ahadi, Rothbart, & Ye, 1993). Further research in the U.S. and China found differences in aspects of children's effortful control as early as infancy (Gartstein, et al., 2006). In the Taiwanese culture, parents focus on teaching children to regulate their behaviors, especially around elderly family members (Hsieh, 2004). Children are taught to eat quietly, sit still, and to not speak when elderly family members are speaking. In contrast, parents in the U.S. often encourage children to speak up and participate in family conversations at the dinner table (Martini, 1996). In China, cultural beliefs that may influence behavioral regulation are reflected in large class sizes and teaching practices where most lessons involve group-activities rather than individual learning (Pang & Richey, 2007). Young Chinese children are expected to regulate themselves and follow the behaviors of the group, even when these behaviors may conflict with more dominant responses. In South Korea, on the other hand, a child-centered approach to early childhood education has become increasingly popular (McMullen, et al., 2005). As these cultural differences suggest, studies of behavioral regulation conducted in the U.S. may lead us to new research questions about other societies, but they may not generalize to other groups. Thus, in the present study, we were especially interested in analyses within each society that addressed how the HTKS functioned as an assessment of behavioral regulation.

Measuring Behavioral Regulation

Research has traditionally measured behavioral regulation with teacher or parent ratings, as an aggregated score of individual attention, working memory, and inhibitory control tasks, or with direct measures (Bronson, Tivnan, & Seppannen, 1995; Carlson, 2005; Howse, et al., 2003; Rothbart, Ahadi, Hersey, & Fisher, 2001). Although these methods have been useful for understanding perceptions of children's behavior, individual components of behavioral regulation, and for assessing specific populations of children, they also have

limitations. For example, teacher ratings have often differed from scores on other assessments including child and parent ratings, and direct assessments of behavioral regulation (Kunter & Baumert, 2006; Loo & Rapport, 1998; Mahone, et al., 2002; Mahone & Hoffman, 2005; Wall & Paradise, 1981). In particular, teacher ratings in Asia have not always corresponded to directly measured child behaviors (Jose, Huntsinger, Huntsinger, & Liaw, 2000; Wanless, et al., in press). In addition, parent ratings provide useful information but may also be biased (Rothbart, et al., 2001). Further, teacher-rated assessments often use Likert scales, which may have limitations for cross-cultural comparisons because they rely on culturally-based teacher expectations for children's behaviors (Heine, Lehman, Peng, & Greenholtz, 2002). As a result, direct assessments offer a unique perspective on child behaviors that may not be captured by teacher ratings.

Aggregating scores from individual assessments of attention, working memory, and inhibitory control can also be problematic. Researchers in the U.S. and Asia have found relatively weak relations among the different tasks measuring attention, working memory, and inhibitory control (Archibald & Kerns, 1999; Espy & Bull, 2005; Oh & Lewis, 2008). In the present study, we measured the integration of these skills because children's ability to orchestrate these three skills has been shown to be a key predictor of academic success (McClelland, et al., 2007; Ponitz, et al., 2009).

Although there are a number of reliable and valid direct measures of self-regulation, many have been designed for the laboratory or with clinical populations, or have limited usefulness in classroom settings because of practical constraints (Fahie & Symons, 2003; Pickering & Gathercole, 2004). For example, tasks such as the Go/No-Go Task require a computer to administer (Simpson & Riggs, 2006), and tasks such as the Attention Network Task or the Working Memory Test Battery for Children take a relatively long time to administer (Pickering & Gathercole, 2004; Rueda, et al., 2005). A number of tasks such as the day/night stroop task (Gerstadt, Hong, & Diamond, 1994) and the Dimensional Change Card Sort (DCCS) (Müller, Dick, Gela, Overton, & Zelazo, 2006) require less time but children often successfully pass them by age 5 (Carlson, 2005). As noted by Ponitz and her colleagues (2009), one task, Luria's Hand Game, is available that does not require special materials or substantial amount of time to administer, but this task is designed for early elementary school children and may not be appropriate for younger children (Fahie & Symons, 2003).

The HTKS task is a direct measure designed to capture the integration of attention, working memory, and inhibitory control. In the task, children are instructed to perform the opposite of four paired commands: "touch your head" and "touch your toes"; "touch your knees" and "touch your shoulders." The first pair of commands is used in the first half of the task, and the second pair is incorporated in the second half of the task. The task is unique because it is a direct measure of children's ability to use attention, working memory, and inhibitory control skills simultaneously to produce a behavior in a social interaction, which is similar to how children often use these skills in classroom settings. Specifically, the HTKS aims to tap attention skills by requiring children to focus on the task and instructions, working memory by remembering multiple task instructions while applying each of them (i.e. Remembering to touch your head instead of your toes, while responding to a command to "touch your knees"), and inhibitory control by requiring children to stop their dominant response (to touch their toes when asked to touch their toes) and replace this response with the more adaptive response (to touch their head when asked to touch their toes, as instructed at the beginning of the task).

Research demonstrates that the HTKS is significantly related to individual measures of attention, working memory and inhibitory control. For example, higher HTKS scores have

been significantly related to higher parent ratings of attention and inhibitory control in the U.S. (Ponitz, et al., 2009), and to higher directly measured attention and working memory in Beijing (Lan & Morrison, 2008). Moreover, recent research has found evidence that the HTKS is significantly related to direct measures of attention and inhibitory control in the U.S. (R. Duncan & McClelland, 2010). The measure is also useful because it is quick to administer, does not require special materials, requires limited training, and scores have been reliable when the assessment was administered in classroom settings (McClelland, et al., 2007; Ponitz, et al., 2008; Ponitz, et al., 2009).

Goals of the Present Study

In the present study, the psychometric properties of scores on a direct measure of behavioral regulation, the HTKS, were analyzed for young children in the U.S., Taiwan, South Korea, and China by examining (1) the nature and variability of behavioral regulation measured with the HTKS in each society, including relations with teacher-rated classroom behavioral regulation, and (2) relations to early mathematics, vocabulary, and early literacy skills.

We hypothesized that the HTKS would capture substantial variability in behavioral regulation in the samples from all four societies, as has been found with previous research in the U.S. (Ponitz, et al., 2009). In other words, we hypothesized that scores would reflect the full range of the task. It was somewhat unclear whether scores on the HTKS would relate to teacher-rated classroom behavioral regulation in the Asian samples because results from previous research have been mixed. In the U.S., teacher-rated classroom behavioral regulation was significantly and positively related to the HTKS, but research using a simpler form of the HTKS (the Head-to-Toes task; HTT) in Taiwan did not find significant relations with teacher-rated classroom behavioral regulation (Ponitz, et al., 2009; Wanless, et al., in press). Similarly, in previous research in Taiwan, teachers rated children differently than did research assistants who directly observed children's behavior (Jose, et al., 2000). Therefore, it was plausible that HTKS and teacher-rated classroom behavioral regulation would be positively related in the present study, but that the relation would be weaker in the Asian samples than in the U.S. sample. Finally, based on previous studies using the HTKS in the U.S. and the HTT in Taiwan, we anticipated that scores on the HTKS would be positively related to early mathematics, vocabulary, and early literacy skills.

Method

Participants

Data for the present study were collected in four societies: the U.S., Taiwan, South Korea, and China and consisted of parents, teachers, and children who volunteered to participate. Combining all four samples, 814 children, 695 parents, and one teacher from each of 73 classrooms participated in the present study (see Table 1 for descriptive statistics). The ages of the children ranged from 3.12 to 6.50 years old, with each sample's age range overlapping with one another. Taiwan had the youngest mean age (4.56 years) and the U.S. had the oldest mean age (5.48 years). China had the largest age range (3.12 to 6.45 years) and Taiwan had the smallest (3.89 to 5.00 years). Overall, the majority of children (91%) in the present study were either four or five years old.

Participants in Each Society

United States: Participants from the U.S. were recruited from two geographic locations: Michigan and Oregon. Data from both sites were combined for the present study (see Ponitz, et al., 2009 for a description of each U.S. site).

The children in the U.S. sample (N = 310) were between 4.14 and 6.24 years old (M = 5.48, SD = .33; see Table 1) and were in 40 kindergarten classrooms. About half of the children were girls (51%, n = 159), and the average mother's education level was some college. A small subgroup of the children had mothers with a high school degree or less (11%, n = 35). Some of the children (4%, n = 13) from Oregon spoke Spanish as their first language, and were given assessments in Spanish. In the overall U.S. sample, the majority of children were Caucasian (74%), and the remaining children were Asian (7%), Hispanic (6%), or another ethnicity (13%) including children identified as biracial.

Taiwan: Ages of the Taiwanese participants (N = 158) ranged from 3.89 to 5.00 years old (M = 4.56, SD = .29; see Table 1). Children were from ten preschool classrooms in Taipei (the capital city of Taiwan), and about half of the children were girls (48%, n = 76). On average, mothers had between a high school and college degree, and slightly more than half of children had mothers with a high school degree or less (51%, n = 80). Most parents were born in Taiwan (77% of mothers, 100% of fathers), with the remaining mothers originally from China (4%), Vietnam (4%), Indonesia (1%), or the Philippines (1%).

South Korea: Children in the South Korean sample (N = 227) were between 3.58 and 6.50 years old (M = 5.05, SD = .85; see Table 1) and were from 16 preschool classrooms housed in three childcare centers in Seoul (the capital city of South Korea) and Kyonggi province. Slightly less than half the participants were girls (40%, n = 91). Mothers had an average education level between a high school and college degree and almost half of the children had mothers with a high school degree or less (43%, n = 98). All of the children in this sample were of South Korean descent.

China: Children in the sample from China (N = 119) were between 3.12 to 6.45 years old (M = 5.03, SD = .62; see Table 1), were from seven preschool classrooms in Beijing (the capital city of China), and about half of the children were girls (46%, n = 55). Information about mother's education level was not available for the Chinese sample. All of the children were originally from China.

Procedure

Behavioral regulation and academic data were collected from children and teachers in all four samples and additional background information was collected from parents in three samples (the U.S., Taiwan, and South Korea). The Head-Toes-Knees-Shoulders (HTKS) and academic tasks (early mathematics, vocabulary, and early literacy) were given to children in the fall of the school year. The school year in South Korea and China, however, begins in the spring and ends in the winter, so the fall assessment represents the middle of the school year for these children. The school year in the U.S. and Taiwan begins in the fall and ends in the spring, so the fall assessment represents the beginning of the school year for these children. All three academic domains were assessed in each sample except vocabulary in China and early literacy in Taiwan (see Table 1). In each sample, research assistants visited the schools and assessed the children in unused classrooms, multi-purpose rooms, or other quiet spaces. Assessments were given in two sessions, with each session lasting between 15 and 40 minutes. In addition, in all four samples, teachers completed a questionnaire rating classroom behavioral regulation for each child in their classroom in the fall of the school year, at the same time as the HTKS assessment, except in the U.S. where it was completed in the spring of the school year. In other words, in the U.S. sample, the HTKS and academic assessments were administered in the beginning of the school year and the teacher ratings in the end of the school year. In the South Korean and Chinese samples, the HTKS, academic assessments, and teacher ratings were collected in the middle of the school year. And in the

Taiwan sample, the HTKS, academic assessments, and teacher ratings were collected in the beginning of the school year.

Measures

All measures that were not previously used in each society were translated and/or back-translated by professors who were native speakers and also fluent in English, as well as bilingual graduate students from the society where the assessments were used. In the United States (Oregon only), assessments that were not previously translated into Spanish were translated and back-translated by bilingual research assistants and a professor of Spanish, and used with participants who were identified by teachers as having Spanish as their first language.

Background Questionnaire—In the U.S., Taiwan, and South Korea, parents completed background questionnaires asking about parent education level, prior child care experience, child age, gender, and ethnicity. In China, some background information was collected from school records.

Direct Measure of Behavioral Regulation—The Head-Toes-Knees-Shoulders task (HTKS) was used to measure behavioral regulation and children were given the task in their native language. Each of the 20 items were scored with 0 for an incorrect response (touching his/her head when asked to touch his/her head), 1 for a self-correct (initially responding incorrectly, but correcting him/herself), or a 2 for a correct response (touching his/her toes when asked to touch his/her head). Total scores on the HTKS range from 0 to 40 points. There are two forms of the HTKS. Form A starts with head-toes commands and Form B starts with knees-shoulders commands (items 1 – 10). In items 11–20, all four body parts are used in both forms of the task. In the present study, there were no significant differences between scores on the two forms in the U.S., Taiwan, or China when controlling for age, (*ps* > .05), which is consistent with previous research in the U.S. and Taiwan (Ponitz, et al., 2009; Wanless, et al., in press). In South Korea, only one version of the task (Form A) was used.

Research assistants were trained on the HTKS by studying the task forms, watching videos of trained research assistants giving the task to children, and practicing with other research assistants. In the U.S., scores on the HTKS and the Head-to-Toes task (HTT), a simple version of the HTKS, have shown strong inter-rater reliability (Connor, et al., in press; McClelland, 2009). Further, previous research of scores from the Head-to-Toes task in Taiwan has demonstrated strong inter-rater reliability (Wanless, et al., in press). In the U.S., Taiwanese, and Chinese samples in the present study, examiners scored different children, so traditional methods of inter-rater reliability could not be calculated. However, there were no significant differences in these samples between examiners in children's HTKS scores, after controlling for child age: U.S., F (141, 299) = 1.25, $p > .05$; Taiwan, F (40, 155) = 1.08, $p > .05$; China F (28, 114) = 1.28, $p > .05$. In the South Korean sample, two research assistants rated the same children for a subsample of the participants ($n = 72$) and had good consistency for each HTKS item (ICC = .71, $p < .001$).

Teacher-Rated Measure of Classroom Behavioral Regulation—In all four samples, the Child Behavior Rating Scale (CBRS) was used to measure teacher-rated classroom behavioral regulation (Bronson, et al., 1995). Items on the CBRS ask teachers to rate the child's typical behaviors when using materials, interacting with peers, and completing tasks, using a scale of 1 (never) to 5 (usually/always). To determine whether a classroom behavioral regulation factor was present in each of the four societies, CBRS scores were analyzed using principal axis factor analysis with a promax rotation. In each of

the four samples, the same 10-item classroom behavioral regulation factor emerged that has been found in other research in the U.S. (α = .94 – .95; Matthews, et al., 2009; Ponitz, et al., 2009), and in Taiwan (α = .94; Wanless, et al., in press). In the present study, scores on the CBRS classroom behavioral regulation factor in each society had strong inter-item reliability (U.S.: α = .94; Taiwan: α = .94; South Korea: α = .94; China: α = .95). This factor included items such as "Concentrates when working on a task; is not easily distracted by surrounding activities," and "Completes learning tasks involving two or more steps (e.g., cutting and pasting) in an organized way." The mean score on the CBRS classroom behavioral regulation factor ranged from 1 to 5 with higher scores demonstrating higher levels of classroom behavioral regulation.

Academic Achievement

United States: Mathematics, literacy, and vocabulary subtests from the Woodoock Johnson Psycho-Educational Battery-III Tests of Achievement (WJ-III; Woodcock & Mather, 2000) was used for the English-speaking U.S. children and the Batería Woodcock-Muñoz-R (Batería-R; Woodcock & Muñoz-Sandoval, 1996) was used for the Spanish-speaking U.S. children. There were 13 Spanish-speaking children out of the sample of 310 children. Each child's score was based on the norms for the specific version of the test. The Batería-R standard scores are considered to be equivalent to the Woodcock Johnson Psycho-Educational Battery-R (Woodcock & Johnson, 1989) which is the English version of the assessment prior to the one used for the English speaking children in our sample (WJ-III). Unfortunately, the Spanish version of the WJ-III was not available at the time of testing, so the Batería-R was used. Children were tested on the same subtests (applied problems, letter-word id, and picture vocabulary) for the WJ-III and the Batería-R. The sample to standardize the items on the Batería-R consisted of approximately 2,000 native Spanish-speaking individuals (Woodcock & Muñoz-Sandoval, 1996). For the WJ-III, the sample consisted of a nationally representative U.S. sample of 8,800 individuals (McGrew & Woodcock, 2001). When both the Batería-R and the Spanish equivalent of the WJ-III (Batería III; Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005) were administered to children in previous research, scores on both versions were highly correlated at r = .95 (Liew, et al., 2008). This evidence suggests that the English WJ-III and Spanish Batería-R scores were comparable and could be used in our analyses, as has been done in previous research. The Applied Problems subtest assessed early mathematics and includes questions about quantity, time, money, and word problems. The Letter-word Identification subtest measured early literacy and asked children to name letters and read actual words, and the Picture Vocabulary subtest uses pictures to assess expressive vocabulary. W-scores were used to take into account age at the time of assessment and to allow for comparison of performance of children across a range of ages. In the United States, inter-rater reliability of scores from these subtests is greater than .85 (Woodcock & Mather, 2000).

Taiwan: Early mathematics and vocabulary skills were assessed in Taiwan with measures that had previously been translated into traditional Chinese and used in Taiwan. The Test of Early Mathematics Ability-2 (TEMA-2) measured relative magnitude, counting, calculation, and enumeration by asking children to count objects on a page, determine greater than and less than, and to correctly identify numbers (Ginsburg & Baroody, 1990). In previous research in Taiwan, scores from the TEMA-2 had high internal consistency (.89 – .90) and test-retest reliabilities between .91 and .94 (Hsu, 1998, 2000). Early vocabulary was measured with the Peabody Picture Vocabulary Test-Revised (PPVT-R), which asked children to point to pictures that were named by research assistants. Scores on the PPVT-R had a split-half reliability of .90 – .97 in Taiwanese samples (Lu & Liu, 1998).

South Korea: Early mathematics and early vocabulary were measured with subtests of the Korean-Wechsler Preschool and Primary Scale of Intelligence (K-WPPSI; Park, Kwak, & Park, 1989). The mathematics subtest measured relative magnitude, counting, and calculation. In previous research, scores on the Korean mathematics subtest had a split-half reliability of .82 – .87 and test-retest reliability of .68 for 4 – 6 year olds (Park, et al., 1989). The vocabulary subtest required children to identify pictured objects, and define words. Scores from this subtest had a split-half reliability of .78 – .86 and a test-retest reliability of . 63 for 4 – 6 year olds (Park, et al., 1989). Early literacy skills were assessed with the Test of Hangul Word Reading in which children are asked to decode two-syllable Korean words and pseudo-words (Choi & Yi, 2007). The internal consistency of scores from the test was .99, split-half reliability was .98 – .99, and test-retest reliability was .93 – .97 in previous research (Choi & Yi, 2007).

China: Early mathematics and early literacy skills were assessed in China with the Zareki-KP task (von Aster, 2001) and the Character Recognition task (Chow, McBride-Chang, Cheung, & Chow, 2008), respectively. Using subtests of the Zareki-KP task, counting and calculation (addition and subtraction) skills were separately assessed and their scores were added together to make a mathematics composite score. Counting and calculation scores in the present study were significantly correlated ($r = .44$, $p < .001$). The task had previously been translated into Simple Chinese (Liu, 2007). In Liu's research, scores on the counting and calculation tests had a reliability of .84 and .87, respectively, and were correlated with teacher reports and cognitive tasks. For the literacy task, all traditional characters were translated into simplified Chinese, and the children were asked to read the characters aloud.

Results

The present study investigated the psychometric properties of scores on a direct measure of behavioral regulation in four societies. Specifically, we examined (1) the nature and variability of behavioral regulation measured with the HTKS in each society, including relations with teacher-rated classroom behavioral regulation, and (2) relations to early mathematics, vocabulary, and early literacy skills.

Missing data and multiple imputation—The HTKS had missing data for just one child out of the 814 children that participated across the four societies (see Table 1). The academic outcomes had less than 2% of missing data in each society, except for South Korea which had 34% missing data for each academic domain including early mathematics, vocabulary, and early literacy. Information on the amount of missing data for all variables is presented in Table 1. In all four samples, the majority of participants did not have missing data for more than 1 variable.

In order to deal with missing data, multiple imputation was utilized for all final models (Acock, 2005). When using multiple imputation, data are assumed to be missing at random (MAR), meaning that the pattern of missingness can be explained by variables that are included in the models or by auxiliary variables (variables not included in our models that are theoretically related to missing data on the variables of interest), and any remaining missingness is random (Schafer & Graham, 2002). First, we examined relations between missingness and the variables in our models. Some of the predictors (such as child age and gender) were significantly related to the missingness of variables that were imputed in all four societies. Second, we explored possible auxiliary variables. For variables with more than 5% missing data, logistic regressions were conducted between possible auxiliary variables and dummy variables indicating whether data were missing. Additional culturally relevant auxiliary variables not included in our models but that might explain missingness

were chosen for each sample and included the amount of experience in preschool, number of hours in preschool each week, family income, whether the child spoke Spanish, or whether the child was an ethnic minority[1]. Although the MAR assumption cannot be fully tested, these findings suggest that the missing data in the present study may be missing at random so results are presented using the imputed data (Acock, 2005; Meng, 1995; Rubin, 1996).

To obtain unbiased parameter estimates, multiple imputation (using Stata) was used to create 10 imputed datasets for each of the four samples (Acock, 2005; StataCorp, 2007). Descriptive statistics for each society using the original and imputed data are presented in Table 1. Descriptive statistics were highly similar for the original and imputed data, which gave us more confidence in using the imputed data. Descriptive statistics, correlations, and regression coefficients were also based on analyses on the 10 imputed datasets for each sample. Correlation significance levels were adjusted to account for the inflated sample size (see Table 2).

Analytic Strategy

We began our analyses by examining descriptive statistics, including intraclass correlation coefficients (ICCs) using multilevel models to account for the nested structure of the data (children nested in classrooms) (Raudenbush & Bryk, 2002). Multilevel analyses were then used to analyze the data, acknowledging the non-independent nature of the children due to classroom membership and to ensure that our standard errors were not biased (Peugh, 2010).

Nature and Variability of Behavioral Regulation Measured by the HTKS

Variability in HTKS scores in each society was found with China showing low and Taiwan showing high variability in scores (see Table 3). In all four societies, children's HTKS scores ranged from 0 to 40, utilizing the entire range of the task. Taiwan (n = 1, 0.6%) and South Korea (n = 3, 1.3%) had very few children reach ceiling levels of 40 points on the task. China (n = 8, 6.7%) and the U.S. (n = 7, 2.3%), had more children, but overall, a relatively small number earned 40 points on the task. Further, less than 10% of the children in each society, except for Taiwan (N = 36, 23%), scored at floor level. Distributions of task scores within each society were somewhat skewed, but skewness and kurtosis values did not exceed accepted levels for normal distributions (Kline, 2005; see Table 3). China's HTKS score distribution was particularly skewed, with many high scoring children, and Taiwan's distribution had a positive skew, with the majority of children in Taiwan having low scores. Taiwanese children, however, were the youngest across all the samples.

Relations between child demographic variables and HTKS scores were also analyzed with correlations within each society (see Table 2). Child age and parent education level were positively and significantly correlated with HTKS scores. Child gender was also significantly related to HTKS scores in the U.S., but not in the Asian societies. This relation was negative, indicating that being a boy was related to lower HTKS scores in the U.S.

To better understand the psychometric properties of scores on the HTKS in the four samples, we compared relations between HTKS scores and teacher-rated classroom behavioral regulation scores. In the U.S. and in South Korea, higher scores on the HTKS were

[1]In South Korea, amount of preschool experience significantly predicted missingness for variables with more than 5% missing data (mathematics, early literacy, vocabulary, and mother's education), so it was included as an auxiliary variable in the multiple imputation model. In the U.S. and Taiwan, the auxiliary variables did not significantly predict missingness on any variables that had more than 5% missing data (mother's education and teacher-rated classroom behavioral regulation), so only the predictor variables (child age, gender, mother's education, and teacher-rated classroom behavioral regulation) were used in the imputation model. In China, no variables had more than 5% missing data, so only the predictor variables (child age, gender, mother's education, and teacher-rated classroom behavioral regulation) were included in the imputation model.

significantly related to higher teacher-rated classroom behavioral regulation scores (U.S. r = .29, South Korea r = .23; see Table 2). In Taiwan and China, however, relations between HTKS and classroom behavioral regulation scores were weak and not significant (Taiwan r = .09, China r = .12).

We also compared the amount of variability in HTKS scores and classroom behavioral regulation within each society by calculating the coefficient of variation (standard deviation divided by the mean). In all four societies, the HTKS scores reflected greater variability than the teacher ratings of classroom behavioral regulation (see Table 3).

Finally, we calculated intra-class correlation coefficients (ICCs) using multilevel models for the HTKS and CBRS in each society to determine how much of the variation in these measures was due to classroom membership (see Table 3). In all four societies, more variance in CBRS scores than HTKS scores was attributable to classroom differences, controlling for the predictors used in the present regression analyses (see Table 4).

HTKS Scores and Relations to Mathematics, Vocabulary, and Early Literacy

To investigate our second research question, we used correlations and multilevel modeling within each sample to determine whether HTKS scores related to early academic outcomes. In all samples that measured mathematics (U.S., Taiwan, South Korea, China) and in those that measured vocabulary (U.S., Taiwan, South Korea), higher HTKS scores were significantly correlated with higher mathematics skills, and higher vocabulary scores (see Table 2). In addition, higher HTKS scores were significantly related to higher early literacy scores in South Korea, the U.S., and China (early literacy was not assessed in Taiwan).

To consider the effect of children being nested in classrooms, ICCs were calculated for all academic outcomes in each sample. Less than 9% of the variance in mathematics, reading, and vocabulary was between classrooms. ICCs in South Korea and Taiwan reached statistical significance (ps<.05), but none of the ICCs in the U.S. and China were statistically significant (p > .05). Since some of the ICCs were significant, multilevel analyses using Hierarchical Linear Modeling software (HLM), were conducted for all models to maintain a consistent analytical approach (Raudenbush, Bryk, Cheong, & Congdon, 2004).

Our final models examined the relations between directly measured behavioral regulation (HTKS) scores and early mathematics, vocabulary, and early literacy skills for each sample controlling for child age, gender, mother's education (except in China), teacher perceptions of children's behavioral regulation, and whether the data were collected in Oregon or Michigan (only in the U.S. analyses) (see Table 4). Overall, HTKS scores were significantly related to all achievement outcomes with a few differences based on the sample. In all four societies, higher HTKS scores were significantly related to higher mathematics skills, and patterns were similar in the three samples that assessed early literacy (U.S., China, and South Korea) and vocabulary (U.S., Taiwan, and South Korea). Specifically, higher scores on the HTKS significantly predicted higher early literacy in the U.S., China, and South Korea, and higher vocabulary skills in the U.S. and Taiwan, but not in South Korea.

Discussion

Overall, evidence for the psychometric properties of HTKS scores was found across the four samples, but some differences also emerged. HTKS scores demonstrated variability in behavioral regulation in each society but were differentially related to teacher perceptions of classroom behavioral regulation with significant relations present in the U.S. and South Korea but not in Taiwan or China. Finally, higher behavioral regulation scores on the HTKS were significantly related to higher mathematics, vocabulary, and early literacy (with the

exception of vocabulary in South Korea), beyond the influence of control variables and teacher-rated classroom behavioral regulation.

Nature and Variability of Behavioral Regulation

HTKS scores reflected individual variation in behavioral regulation in all four samples which supports research in the U.S. finding that HTKS scores capture variability in children's behavioral regulation in early childhood (McClelland, et al., 2007; Ponitz, et al., 2008; Ponitz, et al., 2009). Specifically, scores in all four cultures covered the entire range of possible HTKS scores. This suggested that the HTKS differentiated children and was not too easy or difficult for the majority of children assessed in each culture.

In each society, child age and parent education was significantly and positively related to HTKS behavioral regulation scores. Child gender, however, was not significantly related to HTKS scores in the Asian samples. It may be that behavioral regulation is promoted more strongly, regardless of gender, in the Asian societies. For example, previous research has found that teachers gave more instructions regarding behavioral regulation in Chinese classrooms compared to U.S. classrooms (Lan, et al., 2009). Moreover, recent research in the U.S. has documented that boys scored significantly worse on the HTKS than girls in the fall and spring of the kindergarten year (Matthews, et al., 2009). Thus, boys may be more at-risk for poor behavioral regulation in the U.S. but not in Asian samples, which could be due to a greater tolerance for aggressive, unregulated behavior in boys in the U.S. (Entwisle, Alexander, & Olson, 2007). Taken together, these results extend previous research by finding that HTKS scores demonstrate variability in behavioral regulation across four samples in the U.S., Taiwan, South Korea, and China. They also suggest that variability in scores is related to child demographic factors like age and parent education but cultural differences exist in how gender is related to HTKS scores.

Relations between the HTKS and Teacher-Rated Classroom Behavioral Regulation

Results of this study also point to culturally-specific relations between HTKS scores and teacher ratings of classroom behavioral regulation. In the U.S. and South Korea, higher HTKS scores were significantly related to higher classroom behavioral regulation ratings. In Taiwan and China, however, HTKS scores were not significantly related to teacher ratings. Previous research has also found inconsistent relations between the HTKS scores and teacher-rated classroom behavioral regulation, indicating the need for further research into this issue (Ponitz, et al., 2009; Wanless, et al., in press). One explanation for differences in relations between classroom behavioral regulation and HTKS scores may be the familiarity of the teachers with the children. In the U.S. and South Korea where the HTKS was significantly related to teacher-ratings, teacher ratings were collected after the teachers had one semester with children in their classes. In Taiwan, teacher ratings were collected at the beginning of the school year, but HTKS was not significant related to teacher ratings. It is possible that the Taiwanese teachers had less experience with the children in their classes before completing the teacher ratings. However, although teachers in China rated children after a semester, HTKS scores were not significantly related to teacher ratings. This lack of relation may be a function of the particularly large class sizes in China which may have limited how familiar the teachers were with the children, even after one semester (Stevenson & Stigler, 1994). Overall, it is possible that teacher familiarity influenced teacher ratings but we were unable to directly test this in the present study, and future research should examine it. Based on the differential relations between the HTKS and teacher ratings, our results suggest that in addition to teacher reports, HTKS scores may independently contribute useful information for differentiating children's levels of behavioral regulation skills.

Finally, classroom behavioral regulation also differed from HTKS scores in predictability of academic outcomes. In all four samples and for all academic outcomes assessed, HTKS scores were uniquely predictive of academic skills beyond classroom behavioral regulation. Two possibilities may explain the different information provided by these two assessments. First, HTKS scores may reflect different aspects of regulation than classroom behavioral regulation which are differentially relevant for academic success. An accumulation of classroom experiences with the children, with skills other than behavioral regulation, may be included in teachers' ratings, in contrast to a directly administered measure. For example, children may have difficulty with persistence or cooperation which may influence teachers' ratings of the children's behavioral regulation. These skills, however, are not aspects of behavioral regulation as measured by the HTKS task. Second, teacher-ratings may capture different child, classroom, teacher, or cultural characteristics compared to what the HTKS assesses (Mashburn, Hamre, Downer, & Pianta, 2006). One cross-cultural study of likert scales, for example, suggested that cultural values shaped teachers expectations for the child characteristics represented on a likert scale, which led to the *reference-group effect*, which is the lack of a common reference group. This made cross-cultural comparisons of likert scale scores problematic (Heine, et al., 2002). In the present study, teachers within a society may have also had difficulty rating children with a similar reference group in mind, based on teachers' different years of experience or different range of ability levels represented in their classes. It is also possible that the regulation measures were not reliable in Asian cultures; although evidence has been found for the measures in U.S. samples (Matthews, et al., 2009; McClelland, et al., 2007; McClelland & Morrison, 2003; Ponitz, et al., 2009) and for a simplified version of the HTKS in Taiwan (Wanless, et al., in press). Future research is needed to further examine this issue. Together, results comparing HTKS scores and teacher-ratings suggest that in all four samples, HTKS scores provide a unique window on children's academic achievement, beyond that of teacher ratings of classroom behavioral regulation.

Relations between Behavioral Regulation on the HTKS and Early Mathematics, Vocabulary, and Early Literacy

In the present study, HTKS scores predicted children's academic achievement in all samples. Specifically, HTKS scores significantly predicted early mathematics in all four samples, early literacy in all three samples that it was measured (U.S., South Korea, and China), and early vocabulary in the samples that it was measured (U.S. and Taiwan), with the exception of vocabulary for South Korean children. These significant relations were present after controlling for child age, gender, mother's education (except in China), teacher-rated classroom behavioral regulation, and whether the data were collected in Oregon or Michigan in the U.S. analyses. It is likely that integrating attention, working memory, and inhibitory control helps children perform better on achievement tests and in academic settings. Early childhood classrooms place demands on children to pay attention to instructions, to remember classroom and activity rules, and to inhibit impulses to behave in ways that conflict with these instructions and rules. Children who are able to master these skills, and use them in tandem in classroom settings, are more likely to acquire new skills and knowledge which will translate into increased mathematics, vocabulary, and reading scores.

This relation was present beyond the influence of classroom behavioral regulation and important background variables. These findings align with previous research relating behavioral regulation and academic achievement (Howse, et al., 2003; McClelland, Piccinin, & Stallings, 2010) as well as effortful control, which includes aspects of behavioral regulation, and academic achievement (Checa, et al., 2008; Zhou, et al., 2010). The predictive utility of HTKS scores is noteworthy given that these effects were found after taking classroom behavioral regulation into account. This suggests that HTKS behavioral

regulation is important for early academic success in all four societies, which supports a growing body of evidence found with U.S. samples (Ponitz, et al., 2009).

Of all of the academic domains tested in each sample, the only domain not significantly related to HTKS behavioral regulation scores was vocabulary in the South Korean sample. This lack of a significant relation may also reflect differences in the vocabulary measures used in each sample. Specifically, the vocabulary measures used in the U.S. and Taiwan required children to know the names of words, but the South Korean measure included an additional component for children to define words. This difference should be examined in future research.

Within all four samples, high behavioral regulation scores on the HTKS were most strongly related to high early mathematics compared to vocabulary or literacy. This fits other findings in the U.S. (Blair & Razza, 2007;G. J. Duncan, et al., 2007; Ponitz, et al., 2009) and researchers have suggested that behavioral regulation may be particularly important for mathematics skills because the HTKS and math assessments may require similar cognitive processes (Ponitz, et al., 2009). Thus, the findings from the present study, and from previous research, suggest that the skills children need to be successful on the HTKS (attention, working memory, and inhibitory control) help children achieve in early mathematics, even more so than in vocabulary or early literacy.

Two issues regarding model specification warrant consideration. First, analyses in the present study examine behavioral regulation as a predictor of academic achievement, even though both measures were collected concurrently. Conversely, it is possible that children with better academic skills subsequently develop stronger behavioral regulation. Previous research, however, supports a directional relation and has found that early behavioral regulation significantly predicts later achievement (even after accounting for child IQ; McClelland, et al., 2006; McClelland, et al., 2000). Moreover, other research has found that early behavioral regulation significantly mediates relations between family risk factors and later academic achievement (Sektnan, McClelland, Acock, & Morrison, 2010) and between emotion regulation and achievement (Howse, et al., 2003). Together, this research supports a directional relation between behavioral regulation and achievement including the view that children who can pay attention, remember instructions, and inhibit certain behaviors in classroom settings are more likely to take advantage of learning opportunities and do better on achievement tests. Although it is also possible that better behavioral regulation predicts stronger achievement, which then predicts better behavioral regulation, our data do not allow us to test this in the present study. This is an important avenue for future research.

Second, it is possible that a spurious variable, such as intelligence, accounts for the relation between HTKS scores and academic achievement. In fact, general mental ability, as measured by psychometric intelligence tests, is the strongest predictor to date of academic achievement for children of any age (Gottfredson, 2005). Previous research in the U.S., however, has demonstrated that aspects of behavioral regulation uniquely predicted academic achievement beyond the influence of intelligence (Blair, 2006; McClelland, et al., 2006; McClelland, et al., 2000). We suggest that HTKS scores may positively relate to academic achievement over and beyond a measure of intelligence. Although we were not able to directly test this in the current study, future research using the HTKS to assess behavioral regulation should include a measure of intelligence.

Practical Implications

Early childhood professionals in the U.S. and South Korea have called for additional attention to be paid to the importance of social and behavioral skills for school readiness (Kim, et al., 2003; Rimm-Kaufman, Pianta, & Cox, 2000). Results from the present study

support this assertion and suggest that in the U.S., Taiwan, South Korea, and China, three of which have particularly high academic outcomes, strong behavioral regulation in preschool predicts children's early achievement. Thus, promoting the development of behavioral regulation in early childhood settings in these societies may help children to be more successful in school. This is especially relevant for children who struggle with self-regulation, including an increasing number of children in the U.S. (Gilliam & Shahar, 2006). This is also relevant in Asian societies, where although compliance is a cultural norm, there are variations among children and not all children exhibit strong self-regulation (Ahadi, et al., 1993).

These results also suggest that promoting early behavioral regulation in the U.S., Taiwan, South Korea, and China may have significant benefits for young children. A number of behavioral regulation interventions used in the U.S. have demonstrated effectiveness for improving children's skills. For example, the *Tools of the Mind* curriculum that incorporates sociodramatic play, private speech, and drawing strategies that help children to pay attention, has increased children's attention, working memory, and inhibitory control skills, and achievement skills (Barnett, et al., 2008; Diamond, Barnett, Thomas, & Munro, 2007). In addition, another recent study found that preschool games designed to help children practice attention, working memory, and inhibitory control skills significantly improved children's behavioral regulation, especially for children low in these skills (Tominey & McClelland, 2008). Early childhood professionals in the U.S., Taiwan, South Korea, and China may be able to help children succeed in school by incorporating similar games, especially for children who have low behavioral regulation.

Limitations

The present study revealed a number of findings about the use of the HTKS for measuring behavioral regulation across four societies. There are some limitations, however, that should be noted. First, all of the participants in this study from Asian societies lived in urban areas. Including children from multiple geographic areas would provide a more accurate picture of the presence and variability of behavioral regulation in these societies. Second, although there was overlap in child age across samples, and age was controlled for in the analyses, the age groups and ranges were not the same. This limitation made it difficult to compare behavioral regulation means, distributions, and floor and ceiling effects across samples. In future research, age ranges should be expanded in order to more clearly define the upper and lower age limits of the HTKS task in each society. Third, some variables in some samples (mother's education, early mathematics, vocabulary, and early literacy in South Korea; teacher-rated behavioral regulation in the U.S.) had a relatively high amount of missing data which suggests that results regarding analyses using these variables may need to be interpreted with caution. Comparison of means and standard deviations of these variables before and after imputation, however, suggests that overall descriptive scores are relatively similar after missing data have been replaced. Fourth, teacher-rated classroom behavioral regulation data were collected in the U.S., South Korea, and China in the middle of the school year and in Taiwan at the beginning of the school year. Future cross-cultural studies should align the time of year that data is collected so that society effects and timing effects can be disentangled. Finally, future studies should use more than one direct measure of behavioral regulation and teacher-rated measure within each society, and use academic measures with maximum cross-cultural consistency across societies to better understand the how much the assessment accounts for nuances in relations among these variables.

Conclusion

This preliminary study examined the psychometric properties of scores on a direct measure of behavioral regulation, the HTKS, in four societies. Findings suggest that HTKS scores

were reliable, captured variability, and predicted academic achievement in the U.S., Taiwan, South Korea, and China. Cultural differences emerged in relations between HTKS directly measured behavioral regulation and teacher-rated classroom behavioral regulation. Overall, results suggest that HTKS scores may be reliable and validly interpreted as a reflection of behavioral regulation for preschoolers in the U.S., Taiwan, South Korea, and China. These findings extend previous research on behavioral regulation in the U.S. and provide the foundation for continued research beyond the U.S. on behavioral regulation in high-achieving societies including Taiwan, South Korea, and China to help ensure that all children are successful in school.

Acknowledgments

The authors would like to thank all of the participants, collaborators, and research assistants for their contributions. The Taiwan Social Skill Development Study was supported by the U.S. Department of State Fulbright Student Scholarship, the Ryoichi Sasakawa Young Leaders Fellowship Fund, the Oregon Sports Lottery, the Kappa Omicron Nu Honor Society, Oregon State University Department of Human Development and Family Sciences, and Fu Jen Catholic University Department of Child and Family Studies. The University of Michigan Pathways to Literacy Project was funded by the National Institute of Child and Human Development and the National Science Foundation under grant numbers R01 HD27176 and 0111754.

References

Acock AC. Working with missing values. Journal of Marriage and Family. 2005; 67:1012–1028.

Adams AM, Bourke L, Willis C. Working memory and spoken language comprehension in young children. International Journal of Psychology. 1999; 34(5/6):364–373.

Ahadi SA, Rothbart MK, Ye R. Children's temperament in the US and China: similarities and differences. European Journal of Personality. 1993; 7:359–377.

Archibald SJ, Kerns KA. Identification and description of new tests of executive functioning in children. Child Neuropsychology. 1999; 5(2):115–129.

Baldi, S.; Jin, Y.; Skemer, M.; Green, PJ.; Herget, D. Highlights from the PISA 2006: Performance of US 15-year-old students in sciences and mathematics literacy in an international context. National Center for Education Statistics; 2007.

Bao, L. China. In: Slater, JJ., editor. Teen life in Asia. Westport, CT: Greenwood Press; 2004. p. 19-33.

Barnett WS, Jung K, Yarosz DJ, Thomas J, Hornbeck A, Stechuk R, et al. Educational effects of the Tools of the Mind curriculum: A randomized trial. Early Childhood Research Quarterly. 2008; 23(3):299–313.

Blair C. How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. Behavioral and Brain Sciences. 2006; 29(2):109–125. [PubMed: 16606477]

Blair C, Razza RP. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child Development. 2007; 78(2):647–663. [PubMed: 17381795]

Bronson MB, Tivnan T, Seppannen PS. Relations between teacher and classroom activity variables and the classroom behaviors of prekindergarten children in Chapter 1 funded programs. Journal of Applied Developmental Psychology. 1995; 16:253–282.

Carlson SM. Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology. 2005; 28(2):595–616. [PubMed: 16144429]

Checa P, Rodríguez-Bailón R, Rueda M. Neurocognitive and Temperamental Systems of Self-Regulation and Early Adolescents' Social and Academic Outcomes. Mind, Brain, and Education. 2008; 2(4):177–187.

Choi NY, Yi SH. The effects of alphabet knowledge on Korean kindergarteners' reading of Hangul words. Journal of Korean Home Management Association. 2007; 25(3):151–168.

Chow B, McBride-Chang C, Cheung H, Chow C. Dialogic reading and morphology training in Chinese children: Effects on language and literacy. Developmental Psychology. 2008; 44(1):233–244. [PubMed: 18194022]

Cole, M. Cultural psychology: A once and future discipline. Cambridge, MA: The Belknap Press of Harvard University Press; 1996.

Connor CM, Ponitz CC, Phillips BM, Travis QM, Glasney S, Morrison FJ. First graders' literacy and self-regulation gains: The effect of individualizing student instruction. Journal of School Psychology. (in press).

Deater-Deckard K, Mullineaux P, Petrill S, Thompson L. Effortful Control, Surgency, and reading skills in middle childhood. Reading and Writing. 2009; 22(1):107–116. [PubMed: 20526377]

Diamond A, Barnett WS, Thomas J, Munro S. Preschool program improves cognitive control. Science. 2007; 318:1387–1388. [PubMed: 18048670]

Dixon WE Jr, Salley BJ. Shhh! We're tryin' to concentrate": Attention and environmental distracters in novel word learning. The Journal of Genetic Psychology. 2007; 167(4):393–414. [PubMed: 17645230]

Dowsett SM, Livesey DJ. The development of inhibitory control in preschool children: Effects of "Executive skills" training. Developmental Psychobiology. 2000; 36(2):161–174. [PubMed: 10689286]

Duncan GJ, Dowsett CJ, Claessens A, Magnuson K, Huston AC, Klebanov P, et al. School readiness and later achievement. Developmental Psychology. 2007; 43(6):1428–1446. [PubMed: 18020822]

Duncan, R.; McClelland, MM. Relations among individual measures of self-regulation. 2010. Manuscript in preparation

Dwyer, E. South Korea. In: Slater, JJ., editor. Teen life in Asia. Westport, CT: Greenwood Press; 2004. p. 205-222.

Eisenberg N, Sadovsky A, Spinrad T. Associations of emotion-related regulation with language skills, emotion knowledge, and academic outcomes. New Directions for Child and Adolescent Development. 2005; 109:109. [PubMed: 16342899]

Entwisle DR, Alexander KL, Olson LS. Early schooling: The handicap of being poor and male. Sociology of Education. 2007; 80:114–138.

Espy KA, Bull R. Inhibitory processes in young children and individual varation in short-term memory. Developmental Neuropsychology. 2005; 28(2):669–688. [PubMed: 16144432]

Espy KA, McDiarmid MM, Cwik MF, Stalets MM, Hamby A, Stern TE. The contribution of executive functions to emergent mathematic skills in preschool children. Developmental Neuropsychology. 2004; 26(1):465–486. [PubMed: 15276905]

Fahie CM, Symons DK. Executive functioning and theory of mind in children clinically referred for attention and behavior problems. Journal of Applied Developmental Psychology. 2003; 24(1):51–74.

Gartstein MA, Gonzalez C, Carranza JA, Ahadi SA, Ye R, Rothbart MK, et al. Studying cross-cultural differences in the development of infant temperament: People's Republic of China, the United States of America, and Spain. Child Psychiatry and Human Development. 2006; 37:145–161. [PubMed: 16874564]

Gathercole SE, Pickering SJ. Working memory deficits in children with low achievements in the national curriculum at 7 years of age. British Journal of Educational Psychology. 2000; 70:177–194. [PubMed: 10900777]

Gerstadt CL, Hong YJ, Diamond A. The relationship between cognition and action: Performance of children 3.5–7 years old on a Stroop-like day-night test. Cognition. 1994; 53:129–153. [PubMed: 7805351]

Gilliam WS, Shahar G. Preschool and child care expulsion and suspension rates and predictors in one state. Infants and Young Children. 2006; 19(3):228–245.

Ginsburg, HP.; Baroody, AJ. Test of early mathematics ability. 2. Austin, TX: Pro-Ed; 1990.

Gottfredson, L. Implications of cognitive differences for schooling within diverse societies. In: Frisby, C.; Reynolds, CR., editors. Comprehensive handbook of multicultural school psychology. New York: Wiley; 2005. p. 517-554.

Heine SJ, Lehman DR, Peng K, Greenholtz J. What's wrong with cross-cultural comparisons of subjective likert scales? The reference-group effect. Journal of Personality and Social Psychology. 2002; 82(6):903–918. [PubMed: 12051579]

Howse RB, Calkins SD, Anastopoulos AD, Keane SP, Shelton TL. Regulatory contributors to children's kindergarten achievement. Early Education and Development. 2003; 14(1):101–119.

Hsieh MF. Teaching practices in Taiwan's education for young children: Complexity and ambiguity of developmentally appropriate practices and/or developmentally inappropriate practices. Contemporary Issues in Early Childhood. 2004; 5(3):309–329.

Hsu, H-H. The concept of early learning mathematics seminar collection. Paper presented at the Life, games, mathematics in young children; Taipei, Taiwan. 1998.

Hsu, H-H. Child mathematical aptitude test TEMA-2. Taipei, Taiwan: Psychological Publishing Company Ltd; 2000.

Huntsinger CS, Jose PE, Liaw FR, Ching WD. Cultural differences in early mathematics learning: A comparison of Euro-American, Chinese-American, and Taiwan-Chinese families. International Journal of Behavioral Development. 1997; 21(2):371–388.

Jose PE, Huntsinger CS, Huntsinger PR, Liaw FR. Parental values and practices relevant to young children's social development in Taiwan and the United States. Journal of Cross-Cultural Psychology. 2000; 31(6):677–702.

Kim J, Lee Y-s, Suen H, Lee GS. A Delphi study of young children's readiness in Korea: Challenges and implications for early childhood schooling. Educational Research and Evaluation. 2003; 9(4): 345–355.

Kline, RB. Principles and practice of structural equation modeling. 2. New York, NY: Guilford Press; 2005.

Kochanska G, Murray KT, Harlan ET. Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. Developmental Psychology. 2000; 36(2): 220–232. [PubMed: 10749079]

Kunter M, Baumert J. Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. Learning Environments Research. 2006; 9(3):231–251.

Lan, X.; Morrison, FJ. Inter-correlations among components fo self-regulation and their relationship with academic outcomes. Paper presented at the International Society for the Study of Behavioral Development; 2008.

Lan X, Ponitz CC, Miller KF, Li S, Cortina K, Perry M, et al. Keeping their attention: Classroom practices associated with behavioral engagement in first grade mathematics classes in China and the United States. Early Childhood Research Quarterly. 2009; 24:198–211.

Liew J, McTigue E, Barrois L, Hughes J. Adaptive and effortful control and academic self-efficacy beliefs on achievement: A longitudinal study of 1st through 3rd graders. Early Childhood Research Quarterly. 2008; 23(4):515–526. [PubMed: 19169387]

Liu, S. Early characteristics of number cognition of children with developmental dyscalculia. Beijing Normal University; Beijing, China: 2007.

Loo SK, Rapport MD. Ethnic variations in children's problem behaviors: A cross-sectional, developmental study of Hawaii school children. Journal of Child Psychology and Psychiatry. 1998; 39(4):567–575. [PubMed: 9599784]

Lu, L.; Liu, H. Xiu Ding Bi Pao De Tu Hua Ci Hui Ce Yan (Revised Peabody Picture Vocabulary Test: Mandarin Chinese Version). Taipei: Xin Li; 1998.

Mahone EM, Cirino PT, Cutting LE, Cerrone PM, Hagelthorn KM, Hiemenz JR, et al. Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. Archives of Clinical Neuropsychology. 2002; 17(7):643–663. [PubMed: 14591848]

Mahone EM, Hoffman JC. Behavior ratings of executive function among preschoolers with ADHD [Abstract]. International Neuropsychological Society. 2005; 11(S1):129.

Martini M. 'What's New?' at the dinner table: Family dynamics during mealtimes in two cultural groups in Hawaii. Early development and parenting. 1996; 5(1):23–34.

Mashburn AJ, Hamre BK, Downer JT, Pianta RC. Teacher and classroom characteristics associated with teachers' ratings of prekindergarteners; relationships and behaviors. Journal of Psychoeducational Assessment. 2006; 24(4):367–380.

Matthews JS, Ponitz CC, Morrison FJ. Early gender differences in self-regulation: A sixth-sense for academic achievement? Journal of Educational Psychology. 2009; 101(3):689–704.

McClelland, MM. Inter-rater reliability for the Head-to-Toes Task. 2009. Manuscript in preparation

McClelland MM, Acock AC, Morrison FJ. The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. Early Childhood Research Quarterly. 2006; 21:471–490.

McClelland MM, Cameron CE, Connor CM, Farris CL, Jewkes AM, Morrison FJ. Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. Developmental Psychology. 2007; 43:947–959. [PubMed: 17605527]

McClelland MM, Morrison FJ. The emergence of learning-related social skills in preschool children. Early Childhood Research Quarterly. 2003; 18:206–224.

McClelland MM, Morrison FJ, Holmes DL. Children at risk for early academic problems: The role of learning-related social skills. Early Childhood Research Quarterly. 2000; 15(3):307–329.

McClelland, MM.; Piccinin, AM.; Stallings, MC. Relations between preschool attention and sociability and later achievement outcomes. 2010. Manuscript in review

McClelland, MM.; Ponitz, CC. Self-regulation in early childhood: Improving conceptual clarity and developing ecologically-valid measures. 2010. Manuscript in review

McClelland, MM.; Ponitz, CC.; Messersmith, E.; Tominey, SL. Self-regulation: The integration of cognition and emotion. In: Overton, WF.; Lerner, RM., editors. Handbook of lifespan human development. Hoboken, NJ: Wiley & Sons; (in press)

McClelland, MM.; Wanless, SB. Touching your toes in four cultures: Developing a new measure of behavioral regulation for young children. Paper presented at the International Society for the Study of Behavioural Development; 2008.

McGrew, KS.; Woodcock, RW. Technical Manual, Woodcock-Johnson III. Itasca, IL: Riverside Publishing; 2001.

McMullen MB, Elicker J, Wang J, Erdiller Z, Lee SM, Lin CH, et al. Comparing beliefs about appropriate practice among early childhood education and care professionals from the U.S., China, Taiwan, Korea, and Turkey. Early Childhood Research Quarterly. 2005; 20(4):451–464.

Meng XL. Multiple imputation inferences with uncongenial sources of input. Statistical Sciences. 1995; 10:538–573.

Müller U, Dick AS, Gela K, Overton WF, Zelazo PD. The role of negative priming in preschoolers' flexible rule use on the Dimensional Change Card Sort task. Child Development. 2006; 77(2):395–412. [PubMed: 16611180]

Mullis, IVS.; Martin, MO.; Gonzalez, EJ.; Chrostowski, SJ. TIMSS 2003 International Mathematics Report. Chestnut Hill, MA: TIMSS & PIRLS International Study Center; 2004.

Muñoz-Sandoval, AF.; Woodcock, RW.; McGrew, KS.; Mather, N. Batería III: Woodcock-Muñoz: Pruebas de aprovechamiento. Itasca, IL: Riverside Publishing; 2005.

Oh S, Lewis C. Korean preschoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. Child Development. 2008; 79(1):80–99. [PubMed: 18269510]

Pang Y, Richey D. Preschool education in China and the United States: A personal perspective. Early Child Development and Care. 2007; 177(1):1–13.

Park, H.; Kwak, K.; Park, K. Lorean-Wechsler Preschool and Primary Scale of Intelligence. Seoul, South Korea: Special Education; 1989.

Peugh J. A practical guide to multilevel modeling. Journal of School Psychology. 2010; 48:85–112. [PubMed: 20006989]

Pickering SJ, Gathercole SE. Distinctive working memory profiles in children with special educational needs. Educational Psychology. 2004; 24(3):393–408.

Ponitz CC, McClelland MM, Jewkes AM, Connor CM, Farris CL, Morrison FJ. Touch your toes! Developing a direct measure of behavioral regulation in early childhood. Early Childhood Research Quarterly. 2008; 23:141–158.

Ponitz CC, McClelland MM, Matthews JS, Morrison FJ. A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. Developmental Psychology. 2009; 45:605–619. [PubMed: 19413419]

Raudenbush, SW. Random effects models. In: Cooper, H.; Hedges, LV., editors. The handbook of research synthesis. New York: Russell Sage Foundation; 1994.

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2. Thousand Oaks, CA: Sage Publications, Inc; 2002.

Raudenbush, SW.; Bryk, AS.; Cheong, YF.; Congdon, R. HLM 6.06: Hierarchical linear and nonlinear modeling. Lincolnwood, IL: Scientific Software International, Inc; 2004.

Rimm-Kaufman SE, Pianta RC, Cox MJ. Teachers' judgments of problems in the transition to kindergarten. Early Childhood Research Quarterly. 2000; 15(2):147–166.

Rosenthal, R. The Pygmalion effect and its mediating mechanisms. In: Aronson, J., editor. Improving academic achievement: Impact of psychological factors on education. San Diego, CA: Academic Press; 2002. p. 25-36.

Rothbart MK, Ahadi SA, Hersey KL, Fisher P. Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. Child Development. 2001; 72(5):1394–1408. [PubMed: 11699677]

Rothbart MK, Posner MI. Genes and experience in the development of executive attention and effortful control. New Directions for Child and Adolescent Development. 2005; 109:101–108. [PubMed: 16342898]

Rothbart, MK.; Sheese, B. Temperament and emotion regulation. In: Gross, JJ., editor. Handbook of emotion regulation. New York: Guilford; 2007. p. 331-350.

Rubin DB. Multiple imputation after 18+ years. Journal of the American Statistical Association. 1996; 91(434):473–489.

Rueda MR, Posner MI, Rothbart MK. The development of executive attention: Contributions to the emergence of self-regulation. Developmental Neuropsychology. 2005; 28(2):573–594. [PubMed: 16144428]

Schafer J, Graham J. Missing data: Our view of the state of the art. Psychological Methods. 2002; 7(2): 147–177. [PubMed: 12090408]

Sektnan M, McClelland MM, Acock AC, Morrison FJ. Relations between early family risk, children's behavioral regulation, and academic achievement. Early Childhood Research Quarterly. 2010

Senn TE, Espy KA, Kaufmann PM. Using path analysis to understand executive function organization in preschool children. Developmental Neuropsychology. 2004; 26(1):445–464. [PubMed: 15276904]

Shweder, RA.; Goodnow, J.; Hatano, G.; LeVine, RA.; Markus, H.; Miller, P. The cultural psychology of development: One mind, many mentalities. In: Lerner, RM., editor. Handbook of child psychology: Vol. 1. Theoretical models of human development. 5. New York: Wiley; 1998. p. 865-937.

Simpson A, Riggs KJ. Conditions under which children expereince inhibitory difficulty with a 'button-press' go/no-go task. Journal of Experimental Child Psychology. 2006; 94(1):18–26. [PubMed: 16325846]

StataCorp. Stata statistical software: Release 10. College Station, TX: StataCorp LP; 2007.

Stevenson HW, Chen C, Lee SY. Mathematics achievement of Chinese, Japanese, and American children: Ten years later. Science. 1993; 259(5091):53–58. [PubMed: 8418494]

Stevenson HW, Lee S-y, Chen C, Lummis M, Stigler JW, Fan L, et al. Mathematics achievement of children in China and the United States. Child Development. 1990; 61(4):1053–1066. [PubMed: 2209177]

Stevenson, HW.; Stigler, JW. The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education. New York: Simon and Schuster; 1994.

Tominey, S.; McClelland, MM. Red light, purple light: Initial findings from an intervention to improve self-regulation over the pre-kindergarten year. 2008. Manuscript in preparation

Tsai HLA, McClelland MM, Pratt C, Squires J. Adaptation of the 36-Month Ages and Stages Questionnaire in Taiwan: Results from a preliminary study. Journal of Early Intervention. 2006; 28(3):213–225.

van der Schoot M, Licht R, Horsley TM, Aarts LT, van Koert B, Sergeant JA. Inhibitory control during sentence reading in dyslexic children. Child Neuropsychology. 2004; 10(3):173–188. [PubMed: 15590496]

von Aster, M. Die Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (ZAREKI). Frankfurt, Germany: Swets & Zeitlinger, Swets Test Services; 2001.

Wall SM, Paradise LV. A comparison of parent and teacher reports of selected adaptive behaviors of children. Journal of School Psychology. 1981; 19(1):73–77.

Wanless SB, McClelland MM, Acock AC, Chen F-M, Chen J-L. Behavioral regulation and early academic achievement in Taiwan. Early Education and Development. (in press).

Woodcock, RW.; Johnson, MB. WJ-R tests of cognitive ability. Itasca, IL: Riverside Publishing; 1989.

Woodcock, RW.; Mather, N. Woodcock Johnson Psycho-Educational Battery-III. Itasca, IL: Riverside; 2000.

Woodcock, RW.; Muñoz-Sandoval, AF. Batería Woodcock-Muñoz-R: Pruebas de aprovechamiento-Revisada, Supplemental Manual. Itasca, IL: Riverside; 1996.

Yi, C-C.; Wu, C-I. Taiwan. In: Slater, JJ., editor. Teen life in Asia. Westport, CT: Greenwood Press; 2004. p. 223-241.

Zhou Q, Main A, Wang Y. The Relations of Temperamental Effortful Control and Anger/Frustration to Chinese Children's Academic Achievement and Social Adjustment: A Longitudinal Study. Journal of Educational Psychology. 2010; 102(1):180–196.

Psychol Assess. Author manuscript; available in PMC 2012 June 1.

Table 1

Descriptive Statistics by Society for Original and Imputed Data, including Mean, Standard Deviation, and Percent Missing

	United States *N* = **310**		**Taiwan** *N* = **158**		**South Korea** *N* = **227**		**China** *N* = **119**	
	Original	**Imputed**	**Original**	**Imputed**	**Original**	**Imputed**	**Original**	**Imputed**
Child Age	5.48 (0.33)	5.48 (0.33)	4.56 (0.29)	4.56 (0.29)	5.06 (0.85)	5.05 (0.85)	5.02 (0.62)	5.03 (0.62)
	0.6%		0.6%		3.1%		0.0%	
Gender	0.49 (0.50)	0.49 (0.50)	0.52 (0.50)	0.52 (0.50)	0.60 (0.49)	0.60 (0.49)	0.54 (0.50)	0.54 (0.50)
	0.0%		0.0%		0.0%		0.0%	
Mother's Education	3.32 (0.70)	3.25 (0.77)	2.45 (0.45)	2.45 (0.61)	2.63 (1.10)	2.67 (1.10)	---	---
	19.4%		13.3%		28.6%			
HTKS	26.28 (11.05)	26.28 (11.03)	15.84 (13.17)	15.79 (13.12)	23.99 (12.96)	23.99 (12.94)	31.77 (8.81)	31.77 (8.77)
	0.0%		0.6%		0.0%		0.0%	
CBRS	4.02 (0.68)	4.00 (0.68)	3.90 (0.62)	3.89 (0.62)	3.75 (0.72)	3.75 (0.71)	3.76 (0.77)	3.76 (0.77)
	35.5%		13.9%		0.4%		1.7%	
Mathematics	434.00 (15.33)	434.04 (15.30)	18.11 (8.65)	18.10 (8.60)	12.77 (3.87)	12.85 (4.01)	5.95 (2.76)	5.95 (2.76)
	0.6%		1.3%		33.9%		0.8%	
Vocabulary	477.41 (13.47)	477.45 (13.45)	38.38 (15.76)	38.35 (15.74)	13.71 (6.07)	13.99 (6.22)	---	---
	0.6%		0.6%		33.9%			
Early Literacy	375.31 (30.18)	375.37 (30.16)	---	---	28.81 (30.49)	28.33 (31.77)	23.61 (20.27)	23.61 (20.20)
	0.6%				33.9%		0.0%	

Note. HTKS is the Head-Toes-Knees-Shoulders task. CBRS is the Child Behavior Rating Scale. Mother's education is coded as (1 = 0–8 years, 2 = >8–≤12 years, 3 = >12–≤16 years, 4 = >16 years)

Table 2

Correlation Matrix

US (*N* = 310)	HTKS	Vocabulary	Mathematics	Early Literacy
Child Age	.12*	.17**	.19***	.19***
Gender	−.16**	.01	.04	−.01
Parent Education	.17**	.37***	.26***	.20***
CBRS	.29***	.09	.28***	.23***
HTKS	---	.31***	.47***	.30***
Taiwan (*N* = 158)				
Child Age	.20*	.27***	.37***	---
Gender	−.06	−.13	−.17*	---
Parent Education	.18*	.19*	.10	---
CBRS	.09	.12	.33***	---
HTKS	---	.30***	.34***	---
South Korea (*N* = 227)				
Child Age	.53***	.51***	.68***	.55***
Gender	−.03	.03	−.06	−.11†
Parent Education	.22***	.20**	.26***	.11†
CBRS	.23***	.13*	.22***	.26***
HTKS	---	.36***	.59***	.50***
China (*N* = 119)				
Child Age	.24**	---	.42***	.23**
Gender	−.09	---	−.19*	.01
CBRS	.12	---	.15	.14
HTKS	---	---	.40***	.24**

Note. HTKS is the Head-Toes-Knees-Shoulders task. CBRS is the Child Behavior Rating Scale.

Gender is coded as a 1 for boys and 0 for girls.

† $p < .10$;

* $p < .05$;

** $p < .01$;

*** $p < .001$.

Table 3

Descriptives of HTKS direct behavioral regulation measure and CBRS teacher-rated behavioral regulation measure

		United States	Taiwan	South Korea	China
HTKS	Variance	121.72	172.26	167.38	76.95
	Coefficient of Variation	.42	.83	.54	.28
	Skewness	−.97	.18	−.67	−2.06
	Kurtosis	−.17	−1.33	−.94	4.32
	ICC	.06	.00	.09	.03
CBRS	Variance	.46	.38	.51	.60
	Coefficient of Variation	.17	.16	.19	.20
	Skewness	−.50	.04	−.40	−.03
	Kurtosis	3.25	2.56	2.77	1.97
	ICC	.15	.38	.04	.60

Note. HTKS is the Head-Toes-Knees-Shoulders task. CBRS is the Child Behavior Rating Scale. ICC is the intraclass correlation coefficient, and is reported controlling for age.

Table 4

Coefficients and standard errors from multi-level models of early mathematics, vocabulary, and early literacy skills on behavioral regulation by society

	United States			Taiwan		South Korea			China	
	Math	Vocab	Early Literacy	Math	Vocab	Math	Vocab	Early Literacy	Math	Early Literacy
	Unstand. Coeff. (SE) Stand. Coeff.			Unstand. Coeff. (SE) Stand. Coeff.		Unstand. Coeff. (SE) Stand. Coeff.			Unstand. Coeff. Stand. Coeff. (SE)	
HTKS	.56*** (.07)	.29*** (.07)	.63*** (.15)	.15** (.05)	.24** (.09)	.08** (.02)	.02 (.05)	.60* (.22)	.10*** (.03)	.43* (.21)
	.40	.24	.23	.23	.20	.26	.04	.24	.32	.19
Child Age (years)	4.10† (2.39)	7.32** (2.24)	11.66* (5.00)	7.39** (2.12)	11.01** (4.17)	2.31** (.53)	3.09*** (.79)	12.82† (6.63)	1.47*** (.37)	6.82* (3.05)
	.09	.18	.13	.25	.20	.49	.42	.34	.33	.21
Gender (boys=1, girls=0)	4.31** (1.51)	1.22 (1.42)	3.31 (3.27)	−1.43 (1.23)	−4.25† (2.39)	−.63 (.41)	.14 (.80)	−5.99 (5.38)	−.39 (.46)	4.32 (3.74)
	.14	.05	.05	−.08	−.14	−.08	.01	−.09	.07	.11
Mother's Education	4.20** (1.31)	5.84*** (1.02)	7.21** (2.36)	1.00 (1.06)	4.45* (2.02)	.65** (.22)	.86† (.47)	1.67 (2.13)	---	---
	.21	.33	.18	.07	.17	.18	.15	.06		
CBRS	3.52* (1.40)	−.11 (1.16)	6.00* (2.70)	4.20** (1.25)	.98 (2.29)	.41 (.47)	.68 (.90)	6.20* (2.98)	.33 (.29)	3.83 (2.42)
	.16	−.01	.14	.30	.04	.07	.08	.14	.09	.15
Site (MI=1, OR=0)	−4.70* (2.09)	2.73 (1.81)	−5.76 (4.50)	---	---	---	---	---	---	---
	−.12	.08	−.08							

Note. Unstand. Coeff. is the unstandardized coefficient. SE is the standard error. Stand. Coeff. is the standardized coefficient. HTKS is the Head-Toes-Knees-Shoulders task. CBRS is the Child Behavior Rating Scale. Parent Education is scored as 1 = 0–8 years, 2 = 9–12 years, 3 = 13–16 years, 4 = greater than 16 years. In the U.S. sample, site was also controlled because data were collected in Michigan (MI) and Oregon (OR).

† < .10;

* $p < .05$;

** $p < .01$;

*** $p < .001$