

AN ABSTRACT OF THE DISSERTATION OF

Shannon B. Wanless for the degree of Doctor of Philosophy in Human Development and Family Sciences presented on December 1, 2008.

Title: Measuring Behavioral Regulation in Young Children.

Abstract approved:

Megan M. McClland

Alan C. Acock

Early behavioral regulation has been identified as an important component of school readiness, yet few easy-to-administer measures exist that directly assess this skill in young children outside of the United States (U.S.). In this dissertation, two studies were conducted to examine the reliability and validity of a direct measure of behavioral regulation in preschoolers. Study 1 examined the psychometric properties of the Head-to-Toes Task (HTT) for 3.5-year-old Taiwanese children, including relations to early math and vocabulary skills. Results indicated that the HTT captured maturational differences, and significantly predicted early math and vocabulary skills when controlling for child age, parent education level, and teacher ratings of behavioral regulation. The task however, was not significantly related to teacher ratings of behavioral regulation. Study 2 examined the reliability and validity of a more complicated version of the HTT the Head-Toes-Knees-Shoulders Task (HTKS) for preschoolers in the U.S., Taiwan, South Korea, and China. Consistency was found on the task across all four samples where the HTKS

captured maturational differences and significantly predicted early academic skills. Some differences were also found in the ages that the task was best-suited for, and in its relations to teacher ratings of behavioral regulation which were significant in the U.S. and South Korea but not in Taiwan and China. Overall, these studies lay the foundation for the use of the HTT and HTKS in four societies as measures that predict early achievement and can be used to help children to be successful in early school settings.

© Copyright by Shannon B. Wanless

December 1, 2008

All Rights Reserved

Measuring Behavioral Regulation in Young Children

by

Shannon B. Wanless

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented December 1, 2008
Commencement June 2009

Doctor of Philosophy dissertation of Shannon B. Wanless presented on December 1, 2008.

APPROVED:

Co-Major Professor, representing Human Development and Family Sciences

Co-Major Professor, representing Human Development and Family Sciences

Chair of the Department of Human Development and Family Sciences

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Shannon B. Wanless, Author

ACKNOWLEDGMENTS

I would like to thank my advisors, Megan M. McClelland and Alan C. Acock for the time and dedication they have generously given to me. Without exception, Megan had high expectations for my work and provided careful and unwavering guidance that raised my standard of success. In particular, I appreciated the integrity with which she conducted her research and her cultivation of the same value in me. Megan worked with unconditional devotion to develop my skills so that my career goals could become realities. I truly owe my successes to her mentorship. Alan has provided wisdom and patient mentoring. He taught me that a higher level of learning can only be attained when we are not afraid to make mistakes. I thank Alan for tirelessly helping me work through research challenges and giving me the courage to face the future with confidence. Alan has inspired me to pass on his patience and passion for learning. Megan and Alan gracefully allowed their different academic styles to complement each other, and the result was an unbeatable mentoring duo.

I would also like to thank my committee members, Sharon E. Rosenkoetter and Lizbeth A. Gray. I am grateful for the many opportunities that Sharon has given me to work on research that directly benefits children and families. Sharon has challenged me to hear the voices of the people we study and to use our findings to give back to their communities. I thank her for helping me to keep the needs and strengths of children and families at the forefront of my research. I thank Liz for sharing her infectious passion for international work. She has given me opportunities to travel abroad that have opened doors for a lifetime of international research and teaching. Seeing the world through

different cultural lenses has fundamentally changed the way that I understand children and families.

Finally I would like to thank my husband, Berry, for being everything that I need as I pursue my professional aspirations. His strength and belief in me, his willingness to be a part of my international adventures, and his loving care for our daughter Maryella have given me the freedom to succeed as a student, a wife, and a mother.

CONTRIBUTION OF AUTHORS

Megan McClelland assisted with securing funding for Study 1 and Study 2, influenced the design of the studies, contributed U.S. sample data to Study 2, advised in the analysis of data for Study 1 and Study 2, and assisted with writing the manuscripts for Study 1 and Study 2. Alan Acock advised in the data analysis for Study 1 and Study 2, and provided feedback regarding the design of the studies and the writing of the manuscripts. Fu-Mei Chen and Jo-Lin Chen assisted with the logistics of collecting data in Taiwan for Study 1 and Study 2. Claire Ponitz and Frederick Morrison contributed U.S. sample data to Study 2. Seung-Hee Son, Kangyi Lee, and Miyoung Sung contributed South Korea data to Study 2. Xuezhao Lan and Su Li contributed China sample data to Study 2.

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION.....	1
BEHAVIORAL REGULATION AND EARLY ACADEMIC ACHIEVEMENT IN TAIWAN.....	16
MEASURING BEHAVIORAL REGULATION IN FOUR SOCIETIES.....	49
CONCLUSION.....	88
BIBLIOGRAPHY.....	104

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1.1 Descriptive Statistics for Predictor and Outcome Variables	45
1.2. Head-to-Toes Task Scores	46
1.3 Correlations between Background, Predictor, and Outcome Variables.....	47
1.4 Effects of HTT Scores on Early Math and Vocabulary Skills.....	48
2.1 Descriptive Statistics by Society for Original Data and Imputed Data, M (SD), and Percent Missing in Original Data.....	80
2.2 Correlation Matrix.....	82
2.3 HTKS Direct Measure and CBRS Teacher-Rated Measure Descriptives.....	84
2.4 HTKS Score Distribution.....	85
2.5 Coefficients (Standard Errors) and Betas from Regressions of Early Math, Vocabulary, and Early Literacy Skills on Behavioral Regulation by Society.....	86

DEDICATION

I dedicate this dissertation to my favorite world travelers: my husband Berry and our daughter Maryella.

Measuring Behavioral Regulation in Young Children

INTRODUCTION

In the past decade there has been increased international recognition of the importance of early childhood education as a foundation for the future success of individuals and societies (Organization for Economic Cooperation & Development, 2001; Tobin, 2007). Some young children, however, are not successful in early childhood classrooms, and there is evidence that these children may be unlikely to catch up to their more successful peers (Phillips, Crouse, & Ralph, 1998; Rimm-Kaufman, Pianta, & Cox, 2000). Improving children's behavioral regulation (including attention, working memory, and inhibitory control) is one important way to help children to be more academically successful in early childhood classrooms in the United States (U.S.) and abroad (Barnett et al., 2008; McClelland, Cameron, Wanless, & Murray, 2007; McClelland & Wanless, 2008). In fact, behavioral regulation relates to academic achievement in preschool and elementary school, and evidence suggests that it predicts the likelihood that a child will graduate high school and college (McClelland, Acock, & Morrison, 2006; McClelland & Piccinin, 2008; Pagani et al., 2008).

Research outside of the U.S. on behavioral regulation and its links to academic success is limited, however, due in part to a lack of reliable and valid direct measures of behavioral regulation that are appropriate across cultures. For example, cross-cultural measures are needed to understand behavioral regulation in high-achieving societies such as Taiwan, South Korea, and China, where preliminary research suggests that children may have particularly strong behavioral regulation (Oh & Lewis, 2008; Sabbagh, Xu,

Carlson, Moses, & Lee, 2006). Based on the need for an efficient, practical, and direct measure of behavioral regulation that can be used cross-culturally, this dissertation includes two studies that examined the psychometric properties of a new measure of behavioral regulation in the U.S., Taiwan, South Korea, and China.

The Role of Behavioral Regulation for Early School Success

Early childhood experts in the U.S. and abroad have pointed to the need for an increased focus on the social aspects of learning, such as behavioral regulation, in order to more effectively help young children succeed in school (Kim, Lee, Suen, & Lee, 2003; Zero to Three, n.d.). In fact, U.S. research has shown that behavioral regulation is a particularly relevant element of school readiness because of its strong links to academic achievement, especially math skills, and because of the many demands that early childhood classrooms often place on children's behavioral regulation (Blair & Razza, 2007; McClelland, Cameron, Wanless et al., 2007). For example, preschoolers are frequently asked to pay attention to the teacher, to remember what the teacher said, and to inhibit behaviors in favor of different behaviors that the teacher asks the child to do. This interest in behavioral regulation may be particularly relevant in Taiwan, South Korea, and China due to the high teacher-child ratios (1:45 or more) and the predominant use of teacher lectures in these societies, which challenge children's behavioral regulation (Hsieh, 2004; Kim et al., 2003; Pang & Richey, 2007). Because of their ability to navigate these demands, children with strong behavioral regulation are more successful in school both concurrently, and over time. Specifically, in studies of U.S. children, behavioral regulation in kindergarten uniquely predicted reading, vocabulary, and math

between kindergarten and second grade (McClelland, Morrison, & Holmes, 2000). Gains in behavioral regulation over the prekindergarten year also significantly predicted improvement in literacy, vocabulary, and math skills in the U.S. (McClelland, Cameron, Connor et al., 2007). Further, a related study found that these skills in U.S. kindergarteners were important for achievement over time and predicted reading and math skills between kindergarten and sixth grade and growth in reading and math between kindergarten and second grade (McClelland et al., 2006). This evidence points to the importance of behavioral regulation for school readiness, but also for school success throughout elementary school and beyond.

Although behavioral regulation is important for academic success, it is just beginning to be examined in Asian societies that have a strong cultural emphasis on academic achievement (Tamis-LeMonda, Wang, Koutsouvanou, & Albright, 2002). Preliminary research has found that South Korean and Chinese children may have relatively high executive function skills which are the foundation for behavioral regulation (McClelland, Cameron, Wanless et al., 2007; Oh & Lewis, 2008; Sabbagh et al., 2006). Moreover, children in Asian societies including Taiwan, South Korea, and China have relatively high academic skills, particularly in math (Baldi, Jin, Skemer, Green, & Herget, 2007; Stevenson, Chen, & Lee, 1993). It is possible that these high academic skills are a function of strong behavioral regulation. Identifying children who are struggling with behavioral regulation may be particularly important in Taiwan, South Korea, and China where parents and teachers are intently focused on their children's academic achievement, but as of yet there are few screening tools available to identify

children that may be struggling (Sang-Hun, 2008; Sue & Okazaki, 1990; Tsai, McClelland, Pratt, & Squires, 2006). Finding an assessment tool that is practical for teachers to use and is reliable and valid across cultures is important for early intervention and research in these societies.

Measuring Behavioral Regulation as an Integrated, Contextually-Relevant Construct

Behavioral regulation is defined as “independently creating and acting on plans for behavior” draws on executive function skills including attention, working memory, and inhibitory control (McClelland, Cameron, Wanless et al., 2007, p91). In the early childhood classroom, behavioral regulation may include following a classroom rule to keep your hands off of other children during circle time, or waiting in line to wash your hands before sitting down to eat your favorite snack. Successfully following these rules requires children to integrate aspects of executive function including attention, working memory, and inhibitory control to produce optimal behaviors for early childhood learning contexts (Blair, 2002; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; McClelland, Cameron, Wanless et al., 2007). In other words, successfully integrating these executive function components allows for the translation of skills into behaviors that help children to be successful in school.

Because measures of behavioral regulation are not available in Taiwan, South Korea, or China, we reviewed U.S. measures of behavioral regulation to find an assessment that might be useful in these societies. Our review, however, revealed that most measures of behavioral regulation have been limited by the use of teacher reports of children’s skills, the aggregation of scores from executive function component measures,

and the use of measures designed for clinical or laboratory settings. First, although teacher-rated measures such as the Child Behavior Rating Scale (Bronson, Tivnan, & Seppanen, 1995) can provide valuable information, they may also be limited by rater bias, and impracticality for teachers. Specifically, a number of studies have highlighted the disagreement between teacher ratings and direct measures (Loo & Rapport, 1998; Mahone & Hoffman, 2005) demonstrating that these two types of assessment may be capturing different skills or that teacher-ratings may be biased by teacher, classroom, cultural, or child characteristics (Mashburn, Hamre, Downer, & Pianta, 2006). Previous research in Asia has also found differences in teacher-rated scores of social skills and direct assessments (Jose, Huntsinger, Huntsinger, & Liaw, 2000). In cross-cultural research, teacher-rated measures that use likert scales may be particularly limited for comparisons due to differences in teachers' culturally-based expectations for how each score may translate into behaviors (Heine, Lehman, Peng, & Greenholtz, 2002). Also, teachers may find that completing rating scales for all of their students is a substantial burden. Teacher-rated scales range in length, with some scales having 80 or more items. If teachers are asked to complete these for each child in their class, outside of their normal workload, they may not be able to complete the scales, may have to take time away from the children to complete them, or they may complete the measures quickly without paying close attention to the questions. In sum, developing a quick and easy-to-administer direct assessment that can be used without burdening teachers is an attractive and practical alternative to teacher-rated measures of behavioral regulation.

Another form of measurement, combining scores from executive function component tasks (measuring attention, working memory, inhibitory control), originates from clinical psychology (Carlson, 2005; Smith-Donald, Raver, Hayes, & Richardson, 2007). However, using these tasks to measure behavioral regulation is problematic for two reasons. First, it is possible that behavioral regulation is substantively different from the sum of these three skills. Integrating attention, working memory, and inhibitory control skills is a challenging task for young children and may reflect unique demands of using all three skills at the same time which may not be captured when measuring each skill individually. Moreover, evidence from previous research, suggests that these three skills are sometimes only weakly correlated (Archibald & Kerns, 1999; Espy & Bull, 2005; Oh & Lewis, 2008). A task that directly measures the integration of attention, working memory, and inhibitory control, is more likely to be a useful measure of behavioral regulation in preschool classrooms because it may reflect the behaviors that children need in order to be successful in school.

Clinical or laboratory tasks, such as those that measure attention, working memory, and inhibitory control are also limited because they are not always practical for classroom use. They may require a quiet environment, special and sometimes expensive materials, and a substantial amount of time to complete. These characteristics of clinical tasks make it less feasible to administer these assessments in classroom settings. In addition, clinical tasks are also limited in their ecological validity, meaning that they may not capture behaviors that truly reflect how children might behave in classrooms (Schmuckler, 2001). Thus, a direct measure that is quick, inexpensive, practical, and

demonstrates strong ecological validity would address some of the limitations of clinical tasks.

One direct measure of behavioral regulation, the Head-to-Toes Task (HTT) and its more complicated version, the Head-Toes-Knees-Shoulders Task (HTKS) assess children's ability to integrate attention, working memory, and inhibitory control to produce self-regulated behaviors. These tasks address the limitations of other behavioral measures by being direct measures, by specifically measuring the integration of executive function skills, and by being affordable, efficient, and ecologically valid for preschool classrooms. Specifically, the tasks can be administered in 5-10 minutes in classroom settings and require no special materials. These characteristics make the tasks attractive for potential use as academic screening tools in early childhood classrooms. This dissertation includes preliminary investigations of the psychometric properties of the HTT and the HTKS for young children in the U.S., Taiwan, South Korea, and China.

Origin and Description of the HTT and HTKS

The HTT and HTKS are tasks adapted from a measure called the Head-Feet task by McCabe, Rebello-Britto, Hernandez, and Brooks-Gunn (2004). The Head-Feet task was designed to be easy-to-administer in home or classroom settings without special materials or technology. The authors hoped that the Head-Feet task would minimize cultural bias due to its straightforward use of language, and lack of materials that may be alien to children of some cultural or socioeconomic statuses. Findings about the properties of the Head-Feet task, however, were limited because they came from analyses

on a sample of less than 30 children. To build on this work, the HTT was directly adapted from the Head-Feet task.

The HTT is a game played one-on-one with a research assistant and a young child. The assistant takes the child to a quiet corner of a classroom or to a school multipurpose room and stands up to begin the game. The child then mimics the research assistant who asks him/her to touch his/her head or toes. The research assistant repeats these commands until the child responds correctly. Now the researcher tells the child that instead of following the commands, the child should do the opposite of the command. For example, instead of touching their head when asked to touch their head, the child should touch their toes. There are two training questions so that the researcher can determine whether the child understands the instruction to do the opposite. Then the game continues with four practice questions that are the same as the training questions. Throughout the training and practice questions, if the child does not respond correctly, the research assistant reminds them of the instruction to do the opposite, up to three times. At this point, whether the child is responding correctly or not, the researcher stops repeating the instructions and administers 10 testing questions that continue using the opposite rule. Throughout the entire game, the researcher records 2 points if the child responds correctly to the commands, 0 points if they respond incorrectly, and 1 point if the child initially responds incorrectly, but then corrects him/herself for a final correct response. The entire task takes about 5 minutes to administer.

The HTKS is an extended version of the HTT. The HTT is considered to be Part 1 of the HTKS, and Part 2 adds two new commands. In Part 2 of the HTKS, the child is

asked to touch his/her knees and shoulders. The children complete training and practice questions that are the same as outlined for the HTT, but these questions only ask about the new commands involving knees and shoulders. Again, the child is reminded to do the opposite up to three times in the training and practice sections. Then, there are 10 testing questions that include all four commands that the child has practiced throughout earlier sections of the HTKS (head, toes, knees, shoulders). Scoring for items on the HTKS is the same as the HTT. The HTKS takes about 10 minutes to administer.

Theoretical Framework

In the present studies, cultural psychology was used to provide a framework for studying behavioral regulation in multiple societies. This theory was chosen because it directly addresses issues that arise when studying one phenomenon in multiple contexts. One major theoretical tenet that guides this dissertation is the belief that patterns of development are contextually specific and behavioral regulation, for example, must be examined in each society separately in order to understand its unique properties and meanings (Cole, 1996; Shweder et al., 1998). Although studies of behavioral regulation conducted in the U.S. may contribute to creating research questions about other societies, U.S. findings may not generalize to other groups. Thus both studies in the present dissertation examined behavioral regulation within individual societal contexts. A review of previous U.S. behavioral regulation research provided a structure for creating our research questions and study designs. Ultimately, however, experts in each society were consulted, and culturally appropriate measures, such as measures of academic skills, were used whenever possible. Behavioral regulation measures were not available in Taiwan,

South Korea, and China, so the present study focused on how the HTT and HTKS (designed in the U.S.) may be adapted to fit the needs of each context. Moreover, the principle investigators on the studies were either natives of the society studied or spent an extended period of time living within the society to gain familiarity with the culture. Thus, the present studies adhered to the assumptions of cultural psychology by balancing the need for a similar behavioral regulation measure across societies so that comparisons could be made, with the need to conduct analyses within each society.

Understanding child outcomes through the lens of cultural psychology. Cultural psychology posits that child outcomes are, in part, a function of cultural mentalities including cultural beliefs, thoughts, and knowledge (Shweder et al., 1998). This belief underlies the reason that the present dissertation studies were conducted separately in each society, even though these societies have cultural similarities. In Taiwan, South Korea, and China, cultural mentalities are strongly rooted in Confucianism which stresses the importance of the group over the individual, as well as the values of hard work and education (Lu & Kao, 2002; Sorensen, 1994; Zhu & Zhang, 2008). These cultural mentalities are inextricably linked with child outcomes via the human behaviors that result from them, which cultural psychologists refer to as custom complexes (Shweder et al., 1998). In other words, the outcome that we focus on in the present studies, behavioral regulation, may have unique features in Taiwan, South Korea, and China due to the practices of children, teachers, and parents, and the cultural beliefs that inform those practices. Although cultural beliefs and mentalities can directly influence child outcomes,

it is the behaviors of individuals within a society that that mediate this relation (Shweder et al., 1998; Sorensen, 1994).

The based behaviors of parents and teachers in Taiwan, South Korea, and China that are grounded in Confucianism may contribute to children's high behavioral regulation (Oh & Lewis, 2008; Sabbagh et al., 2006). For example, in the Taiwanese culture, parents focus on teaching children to control their behaviors, especially around elderly family members (Hsieh, 2004). Children are taught to eat quietly, sit still, or to not speak when elderly family members are speaking. In China, cultural beliefs are also reflected in teaching behaviors where most lessons involve group-activities during which children are expected to all do the same thing and independence is discouraged (Pang & Richey, 2007). In these examples, cultural behaviors which are rooted in Confucius beliefs, lead to parenting and teaching practices that give children many opportunities to practice behavioral regulation. Children are expected to regulate their behaviors and follow the behaviors of the group, even when these behaviors may be in conflict with more dominant responses. Thus far research has not examined the behavioral regulation of preschoolers in Taiwan, South Korea, or China, with a measure that has been evaluated in each society.

Measurement equivalence. The present dissertation studies centered on questions of measurement equivalence, a methodological concern in cultural psychology, which asks whether assessments actually tap the same constructs across contexts (Hughes, Seidman, & Williams, 1993; Ratner & Hui, 2003). Answering this question is essential because it offers insight into whether the data collected with such assessments can

provide meaningful findings in each context. Although our research questions suggest that it may be possible to use one direct measure of behavioral regulation (the HTT for younger children and the HTKS for older children) in multiple societies, we analyzed HTT and HTKS scores within each society separately so that culturally specific aspects of the tasks, such as the ages that the task was most appropriate for, could emerge.

In addition to examining the measurement equivalence of the HTT and HTKS in the present studies, we also considered this issue for other measures used in the studies. Specifically, all samples used the same U.S. measure for teacher-rated behavioral regulation (the Child Behavior Rating Scale; CBRS) which asked teachers to rate the extent to which examples of behavioral regulation reflected each child's behavior, on a scale of one to five. It was important to use a teacher-rated measure to get another perspective on children's behavioral regulation, while keeping in mind that previous research has shown that teacher ratings and direct measures do not always agree. In Taiwan, for example, teacher ratings of children's social skills, and researchers' ratings based on videotaped observations were different (Jose et al., 2000). A teacher-rated measure with documented psychometric properties was not available in Taiwan, South Korea, or China so we examined how the measure functioned in these societies. It was possible, for example, that teachers in cultures that emphasize high expectations for behavioral regulation might have been reluctant to assign the top score to any child. In the present study, we addressed the measurement equivalence of the CBRS by conducting a factor analysis within each society to determine which items held together to represent a behavioral regulation construct.

Overview of Longitudinal Study

A longitudinal study conducted in Taiwan contributed the sample for Study 1, and the Taiwanese sample for Study 2 which also included samples from the U.S., South Korea, and China. The longitudinal study consisted of two time points that were six months apart. Time 1 data collection occurred in March 2007 which was the spring of the children's preschool year when they were 3.5 to 4.5 years old. Time 2 was in September 2007 which marked the beginning of the children's prekindergarten year when the children were 4.0 to 5.0 years old. These two time points were scheduled at the end of one school year and the beginning of the next due to practical constraints. Although gathering data over a six month time period offered important information about growth in children's skills, the timing of these data collection points did not offer information about the effects of being in school over the course of one academic year. Moreover, none of the children attended preschool during the summer months between data collection points (July and August 2007). The longitudinal study included 191 children, 152 parents, and 10 teachers in 10 different classrooms (between Time 1 and Time 2, some of the classrooms combined and others separated leaving a total of nine classrooms at Time 2). All of the classrooms were located in Taipei City (the capital of Taiwan) or Taipei Country. Nine of the ten original participating classrooms were public, did not charge tuition, and although they were open to all children, prioritized admittance of children from diverse backgrounds, such as children of aboriginal ethnicity, children with special needs, and children from families that immigrated from South-East Asia. Overall, the nine public preschool classrooms in our study had 20-25% children from these

categories. The tenth classroom that participated in the longitudinal study was in a university preschool that did not have any children from the diverse backgrounds defined above.

The goals of the longitudinal study in Taiwan were to use the HTT and HTKS (1) to understand the nature, variability, and development of behavioral regulation in Taiwanese preschoolers, (2) to examine how behavioral regulation relates to early academic achievement in Taiwan, and (3) to determine whether teacher and parent beliefs influenced the development of behavioral regulation. Data collected to address these goals include child, parent, and teacher measures (see Figure 1). From children, we used direct assessments to measure their behavioral regulation, early math, and vocabulary skills. Parents completed self-reported questionnaires about their parenting beliefs, background information, and the types of summer activities their children participated in. Teachers reported on their teaching beliefs, rated children's behavioral regulation, and provided teacher background information. Adult measures were gathered at one time point because previous research has documented the stability of parenting and teaching beliefs over a six month period (McClelland, 2002). Measures of children's skills, however, were collected at both points to obtain information about the rate of change in skills that typically develop rapidly during this age period (Berk, 2008). For example, previous research on behavioral regulation in the U.S. has shown that measuring preschoolers at six month intervals showed a substantive change in skill level (McClelland, Cameron, Connor et al., 2007).

Study 1. The studies included in this dissertation addressed the first two goals of the longitudinal study in Taiwan by examining the psychometric properties of two direct measures of behavioral regulation, including relations to early academic achievement and teacher-rated behavioral regulation. Specifically, we examined the nature and variability of behavioral regulation scores and their relations to early academic skills including early math and vocabulary. Study 1 examined the psychometric properties of the HTT, which has been established in the U.S. as a reliable and valid measure of behavioral regulation that is ecologically valid for research in preschool settings (Ponitz, McClelland, Jewkes et al., 2008). Moreover, the HTT has been a useful indicator of academic success in the U.S., and may be a valuable predictor of early achievement in Taiwan (McClelland, Cameron, Connor et al., 2007).

The HTT was designed for use with U.S. 4- to 5-year-olds but previous research has documented that Asian children may have stronger behavioral regulation than their Western counterparts (Oh & Lewis, 2008; Sabbagh et al., 2006). Therefore, task scores from Time 1 (in the spring of the preschool year when children were 3.5 to 4.5 years old), were used to examine the nature and variability of behavioral regulation in Taiwanese preschoolers, and relations to early math, vocabulary, and teacher-rated behavioral regulation.

Study 2. The behavioral regulation measure used in Study 2 was a more complicated version of the HTT, called the HTKS, which included ten additional items and two additional rules to increase the demands of the task. The HTKS has also been found to be reliable and valid for samples in the U.S. and to predict early achievement

outcomes (Ponitz, McClelland, Matthews, & Morrison, 2008). Study 2 extended Study 1 by using a measure designed for older children, and by studying four societies: the U.S., Taiwan, South Korea, and China. The U.S. was included as a sample in which the psychometric properties of the HTKS had already been established, providing a set of information by which to evaluate the usefulness of the task in the other societies. Taiwan, South Korea, and China were included in Study 2 because of the high academic achievement of all three societies, which provided a unique environment for understanding links between behavioral regulation and early academic achievement. Although the goals of Study 2 did not include direct comparisons across societies, patterns of similarities and differences were analyzed as a way to understand cultural nuances in the reliability and validity of the task.

The goals of Study 2 were to examine the nature and variability of behavioral regulation in each society, and investigate relations with early math, vocabulary, and early literacy skills, and with teacher-rated behavioral regulation in each society. The HTKS was designed for older children, so data from Time 2 in the fall of the children's prekindergarten year when the children were 4 to 5 years old were used to examine the psychometric properties of this task. This data was combined with data from related samples in the U.S., South Korea, and China to examine the use of the HTKS for multiple societies. Our goal was to understand the utility of the HTKS across high-achieving Asian cultures.

Although the third goal of the overall longitudinal study in Taiwan was to examine the influences of parents and teachers on behavioral regulation skills, it fell

beyond the scope of the present dissertation. Thus, the two studies presented here lay the groundwork for future research regarding how parents and teacher influences are related to children's early behavioral regulation. Both studies provided a preliminary look at behavioral regulation in Asia and at the use of the HTT and HTKS for behavioral regulation assessment outside of the U.S. Thus, this research provides the foundation for future research on the development of a behavioral regulation screening tool that may identify children that are need of additional academic support.

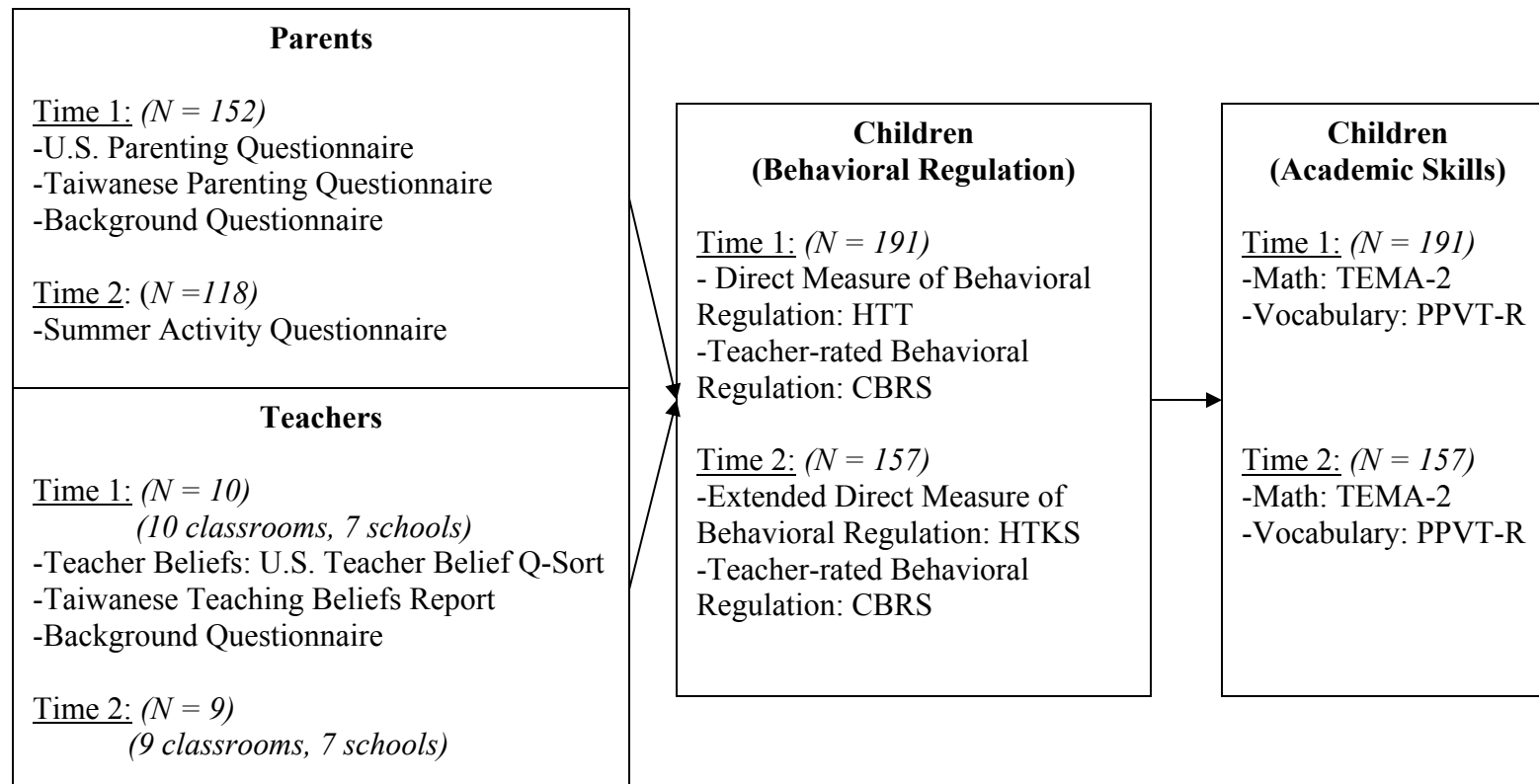


Figure 1. Overview of Longitudinal Study in Taiwan: Study 1 (Time 1) and Taiwanese Sample in Study 2 (Time 2). Time 1 data was collected in spring of the preschool year when the children were 3.5 to 4.5 years old. Time 2 data was collected six months later in fall of the prekindergarten year when the children were 4.0 to 5.0 years old.

BEHAVIORAL REGULATION AND
EARLY ACADEMIC ACHIEVEMENT IN TAIWAN

Shannon B. Wanless^a

Megan M. McClelland^a

Alan C. Acock^a

Fu-Mei Chen^b

Jo-Lin Chen^b

^a Oregon State University, Department of Human Development and Family Sciences. ^b Fu Jen Catholic University, Department of Child and Family Studies.

The authors thank the participants, research assistants, Colleen Hsieh and Claire Cameron Ponitz 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, and the Oregon State University Department of Human Development and Family Sciences.

Abstract

Behavioral regulation (the integration of attention, working memory, and inhibitory control) is critical for school readiness and early academic achievement. In Taiwan, however, where academic success is highly valued, there is a dearth of assessments available to measure young children's behavioral regulation. In the present study, the psychometric properties of a direct measure of behavioral regulation, the Head-to-Toes Task (HTT), were examined with Taiwanese 3- and 4-year-olds by (1) investigating the nature and variability of HTT scores, and by (2) exploring relations between HTT scores and early math and vocabulary skills, and teacher-rated behavioral regulation skills in the spring of the preschool year. Results indicated that the HTT captured substantial variability, was significantly related to early math and vocabulary skills after controlling for age, mother's education level, and teacher-rated behavioral regulation, but was not significantly related to teacher ratings of behavioral regulation. These findings suggest that behavioral regulation is important for early academic success in Taiwan, and the HTT may be a useful measure of behavioral regulation for Taiwanese preschoolers. The HTT and teacher ratings of behavioral regulation, however, may have tapped different aspects of behavioral regulation.

Accumulating evidence suggests that behavioral regulation is a strong predictor of school success. Specifically, children with poor self-regulation are at significantly greater risk of academic difficulty in preschool and throughout elementary school (Blair & Razza, 2007; McClelland et al., 2006; Ponitz, McClelland, Matthews et al., 2008), and are less likely to graduate from high school and college (McClelland & Piccinin, 2008; Pagani et al., 2008). Uncertainty remains, however, about the most appropriate way to measure this skill in young children (Blair, Zelazo, & Greenberg, 2005; Carlson, 2005; Smith-Donald et al., 2007). Teachers need a reliable and valid assessment of behavioral regulation that can be administered quickly and inexpensively in classroom settings, to identify children that would benefit from additional support. In Taiwan, where parents are intently focused on their children's academic success, such an assessment of behavioral regulation would be particularly useful (Stevenson et al., 1990; Yi & Wu, 2004). Unfortunately, most measures of behavioral regulation are limited by their use of observer report or their reliance on clinical and laboratory settings. In the present study, we examined the psychometric properties of a behavioral regulation measure called the Head-to-Toes Task with Taiwanese preschoolers. We investigated the nature and variability of Head-to-Toes Task scores in Taiwan, and relations to early math and vocabulary skills and teacher-ratings of behavioral regulation.

The Taiwanese culture places a strong emphasis on academic success. This cultural value leads to Taiwanese parents being particularly focused on their children's education and to Taiwan having the highest math achievement out of 57 countries tested (Baldi et al., 2007; Kotchick & Forehand, 2002; Ogbu, 1981). In fact, one study found

that Taiwan's cultural emphasis on academic achievement was so strong that Taiwanese students studied eight times as many hours outside of class as their American counterparts, and twice as many hours as their Japanese counterparts ("What the numbers say," 1998). Although Taiwanese parents put a substantial amount of family resources into helping their children succeed in school (Sharma, 1997), there is a lack of culturally appropriate screening tools, and as a result, identification of children who may need additional support (Tsai et al., 2006). Finding a psychometrically sound measure of behavioral regulation for Taiwan is necessary so that children can get the help that they need to be successful in school.

Links between Behavioral Regulation and Early Achievement

Finding an assessment of behavioral regulation that is reliable and valid in Taiwan is particularly important because of behavioral regulation's strong link to school success. A number of studies in the United States (U.S.) have documented this relation concurrently and over time (Blair & Razza, 2007; Gathercole, Pickering, Knight, & Stegmann, 2004; Howse, Calkins et al., 2003). In one study, gains in behavioral regulation over the prekindergarten year significantly predicted improvement in literacy, vocabulary, and math skills (McClelland, Cameron, Connor et al., 2007). In another, children's kindergarten learning-related skills, which included behavioral regulation and social competence, predicted reading and math skills between kindergarten and sixth grade and growth in reading and math between kindergarten and second grade (McClelland et al., 2006). Finally, children with poor behavioral regulation skills had low academic skills in a study of prekindergarteners (Bronson et al., 1995). Less research,

however, has been conducted on links between behavioral regulation and early achievement in Taiwan which was a focus of the present study.

Definition and Components of Behavioral Regulation

Although researchers continue to debate self-regulation definitions, the present study defines behavioral regulation, one aspect of self-regulation, as the integration and behavioral manifestation of cognitive processes including attention, working memory, and inhibitory control (McClelland, Cameron, Connor et al., 2007; McClelland, Cameron, Wanless et al., 2007). A child with strong behavioral regulation can remember and follow a classroom rule, such as washing his or her hands before snack time, rather than using a more prepotent response, such as immediately reaching for the food.

In some previous studies, researchers have individually measured attention, working memory, and inhibitory control to investigate aspects of executive function that underlie behavioral regulation (Blair et al., 2005; Brock, Rimm-Kaufman, & Nathanson, 2008; Carlson, 2005). In the present study, we argue that behavioral regulation is the integration of these skills and a unique predictor of academic success. Each executive function component contributes individually and collectively to early school success. Specifically, attention refers to the ability to voluntarily focus on a particular task while simultaneously ignoring environmental distractions (Rothbart & Posner, 2005; Rueda, Posner, & Rothbart, 2005). Attention is positively related to math in the U.S. and abroad, and is related to vocabulary acquisition when distractions are present such as in classrooms (Dixon Jr. & Salley, 2007; Duncan et al., 2007; Howse, Lange, Farran, & Boyles, 2003). Working memory helps children remember and apply information while

encountering and processing new stimuli (Gathercole & Pickering, 2000). Research has found this skill to be particularly salient for children between 4 and 6 years old, when children enter formal schooling (Senn, Espy, & Kaufmann, 2004), and has been significantly related to stronger math achievement and language comprehension (Adams, Bourke, & Willis, 1999; Espy et al., 2004). Finally, inhibitory control is characterized by the ability to stop one response in favor of a more adaptive behavior (Dowsett & Livesey, 2000) and has been linked to kindergarten math skills and to reading skills in children with dyslexia (van der Schoot et al., 2004).

The integration of attention, working memory, and inhibitory control is relevant in the preschool context where demands for using these skills in tandem are often very high (McClelland, Cameron, Wanless et al., 2007; Rimm-Kaufman & Chiu, 2007). For example, preschoolers often need to make a plan for their behavior and carry out the intended plan, such as when they are building a house with blocks, or setting the table for lunch time amidst potentially distracting activities and peers. Children with strong behavioral regulation are able to integrate attention, working memory, and inhibitory control to successfully plan, organize, and control their behavior in early social settings.

Behavioral Regulation in Taiwan

Research on behavioral regulation has not previously been conducted in Taiwan, but findings from other Asian countries suggest that Asian preschoolers may have particularly strong behavioral regulation. In one study, preschoolers in China consistently scored about 6 months ahead of their counterparts in the U.S. on tasks that individually measured attention, working memory, and inhibitory control (Sabbagh et al., 2006). In

another study, South Korean 3-year-olds had stronger inhibitory control than British 5-year-olds (Oh & Lewis, 2008). Understanding behavioral regulation may be particularly important in Taiwan where classroom characteristics such as high teacher child ratios may place greater demands on children's behavioral regulation (Hsieh, 2004). Research in this area, however, remains limited due to the few studies that have been conducted in Asia and the use of measures that do not directly assess behavioral regulation as an integration of skills necessary for learning in classroom contexts.

Measuring Behavioral Regulation

Behavioral regulation has traditionally been measured by teacher ratings, aggregating scores from tasks that measure components of behavioral regulation such as attention, working memory, and inhibitory control, or by assessments that are designed for use in clinical or laboratory settings. These types of assessments, however, have limitations. First, teacher ratings provide valuable information, but often differ from parent ratings (Wall & Paradise, 1981), child reports (Kunter & Baumert, 2006), and direct assessments (Loo & Rapport, 1998), including direct assessments of executive functioning (Mahone et al., 2002; Mahone & Hoffman, 2005). In one study, cross-cultural differences were found on a teacher-rated measure but no differences were found using an observational assessment. These researchers proposed that the teacher ratings of children's behaviors may reflect cultural influences and expectations more so than direct observations (Jose et al., 2000).

A second method for measuring behavioral regulation is measuring components of children's executive function separately and then combining the scores (Carlson, 2005;

Smith-Donald et al., 2007). Although component assessments are valuable for assessing attention, working memory, and inhibitory control, aggregating their scores may be problematic because these separate measures are often only weakly correlated (Archibald & Kerns, 1999; Espy & Bull, 2005). For example, in one study of Asian preschoolers, measures of attention, working memory, and inhibitory control were not significantly related (Oh & Lewis, 2008). Examining these skills individually does not tap children's ability to use attention, working memory, and inhibitory control simultaneously, which may be a key component of academic success in learning contexts.

Assessments of behavioral regulation that are designed for use in clinical or laboratory settings can also be problematic for school-based research. First, these tasks may require special equipment or technology. For example, in two recent reviews of executive function assessments almost all of the tasks required specialized materials and or a computer, and some required lengthy testing sessions (Carlson, 2005; Ponitz, McClelland, Jewkes et al., 2008). These tasks are also often given as part of a longer battery of assessments that do not easily lend themselves administration in schools (Fahie & Symons, 2003; Pickering & Gathercole, 2004). Second, tasks designed for laboratory settings may have limited ecological validity for research in schools. For example, tasks that ask preschoolers to respond on computers or to press a button may not reflect demands seen in classrooms. In order for a task to be useful to teachers, the materials should be minimal and inexpensive, the administration time should be brief, and the assessment should be ecologically valid.

One measure called the Head-to-Toes Task (HTT) has been used to directly assess behavioral regulation in young children (McClelland, Cameron, Connor et al., 2007; Ponitz, McClelland, Matthews et al., 2008) and has the potential to be used in the future as a screening tool. In the HTT, children are asked to touch their head or toes, (or knees or shoulders in an alternate form of the task), and are instructed to do the opposite of what the experimenter says. For example, the experimenter instructs children to touch their head (or knees), and instead of touching their head (or knees), children are directed to do the opposite and touch their toes (or shoulders). The HTT directly measures behavioral regulation by requiring children to integrate attention, working memory, and inhibitory control, and apply these skills to their behavior. Finally, the task has ecological validity because it approximates tasks seen in early childhood settings.

In research conducted in the U.S., the HTT was a reliable and valid measure of behavioral regulation for a group of over 350 3- to 6.5-year-old children from two geographical locations, including a subset of Spanish-speaking children (Ponitz, McClelland, Jewkes et al., 2008). Results demonstrated variability in HTT scores, with significant effects for age and gender, such that older children did significantly better than younger children, and girls did significantly better than boys, although effects were small for gender (Ponitz, McClelland, Jewkes et al., 2008). The task was reliable, predicted children's early academic achievement, and was significantly correlated to teacher ratings of behavioral regulation (McClelland, Cameron, Connor et al., 2007; Ponitz, McClelland, Jewkes et al., 2008).

Preliminary research in Asian cultures on an extended version of the HTT designed for older children, called the Head-Toes-Knees-Shoulders Task (HTKS; Ponitz, McClelland, Matthews et al., 2008) has found mostly similar results. The HTKS contains the same items as the HTT as well as 10 additional related items to increase difficulty. Findings from China, Taiwan, and South Korea suggest that the task captured variability, age differences (Son et al., 2008; Wanless, McClelland, Son et al., 2008), and was significantly related to early math, reading, and vocabulary skills (Wanless, McClelland, Son et al., 2008), and to teacher ratings of behavioral regulation in South Korea, but not in China or Taiwan (Wanless, McClelland, Son et al., 2008). No significant gender differences were found on the HTKS in these Asian cultures (Xuezhao Lan et al., 2008). These findings provide preliminary evidence of the psychometric properties of an extended version of the task for older children, in China, Taiwan, and South Korea, but it is unclear if similar patterns are present in Taiwan for the HTT, a simpler version of the task which is designed for younger children. If the HTT is reliable and valid in Taiwan, it may have the potential for future use as a screening tool to identify young preschool children who might need additional help with behavioral regulation.

Goals of the Present Study

The present study examined the psychometric properties of the HTT including (1) the nature and variability of scores, and (2) relations to early math and vocabulary, and teacher-rated behavioral regulation with Taiwanese preschoolers. Regarding our first goal, we hypothesized that there would be substantial variability in the behavioral regulation of Taiwanese preschoolers, with older children performing significantly better on the task

than younger children. In previous work, substantial variability and age differences were found in the U.S., and in self-reliance and inhibition in Taiwan (Jose et al., 2000; Livesey & Morgan, 1991; Ponitz, McClelland, Jewkes et al., 2008; Rothbart, Posner, & Kieras, 2006). In addition, based on research on an extended version of the Head-to-Toes Task with older children in China, Taiwan, and South Korea, we anticipated that there would be no significant gender differences on the HTT (Xuezhao Lan et al., 2008). For our second research goal, based on previous U.S. findings, we predicted that HTT behavioral regulation would significantly predict early math and vocabulary skills, after controlling for child age, mother's education level, and behavioral regulation rated by teachers (Howse, Calkins et al., 2003; McClelland, Cameron, Connor et al., 2007). Finally, we expected that the task would be related to teacher ratings of behavioral regulation in Taiwanese 3- and 4-year-olds, as found in the U.S. and South Korea, but that relations would be weak as was found in other studies with older children in China and Taiwan (McClelland, Cameron, Connor et al., 2007; Ponitz, McClelland, Jewkes et al., 2008; Wanless, McClelland, Son et al., 2008). Similarly weak relations have been found in previous research in Taiwan where Taiwanese teachers had different cultural expectations for behavior and rated children differently than observations of behavior (Jose et al., 2000). Thus, it seems plausible that weak relations would be found between teacher ratings and the direct measure of behavioral regulation in the present study.

Method

Participants

Children, parents, and teachers volunteered to participate in a study on the development of preschoolers' behavioral regulation and academic skills through their preschools in Taipei City (the capital of Taiwan) and Taipei County, Taiwan. Six of the seven participating preschools were public, located in an urban setting, and did not charge tuition. All of the schools except for the university preschool were open to all children, but children from diverse backgrounds, such as children of aboriginal ethnicity, children with special needs, and children from families that immigrated from South-East Asia were given priority admittance. Overall, the public preschools in our study had between 20-25% children from these categories and the university preschool had none. Five sites each had one classroom participating in the study, one site had two classrooms, and one site had three classrooms, for a total of ten classrooms participating in the study. Overall, 81.5% ($N=189$) of the recruited families participated in the study. Eighty percent ($N=152$) of the participating families returned parenting questionnaires which limited the sample in the current analyses to 152 children. Controlling for child age, children included in the present analyses ($N=152$) did not significantly differ from excluded children ($N=37$) on HTT scores or vocabulary scores, but excluded children did have significantly lower math scores.

The present study examined children 3.58 to 4.58 years old ($M=4.09$ years, $SD=.29$ years) in the spring of the preschool year (see Table 1). Approximately half of the participants were girls (51%), and girls and boys were evenly distributed within age groups (3.58- to 4.00-year-olds, 31 boys and 30 girls; 4.01- to 4.58-year-olds, 44 boys and 47 girls). Participants attended preschool for an average of 37.64 hours per week (SD

= 6.93) with attendance ranging from 17 to 45 hours each week. Nearly all children (91%), however, attended preschool for 36 to 45 hours per week. The group of children who attended preschool for less than 36 hours per week (9%) was too small to reliably compare to the rest of the children. Preliminary t-tests revealed that children who went to school 17 hours per week were not significantly different from children who went to school 36 to 45 hours per week, in terms of HTT scores, math scores, or vocabulary scores. Further, regression analyses controlling for age and mother education level indicated that the number of hours spent in the classroom each week was not a significant predictor of any of the outcome variables used in the present investigation including HTT scores, math scores, or vocabulary scores. Therefore, all children were included in the present analyses. The average education level for mothers and fathers was between a high school and college degree. Three percent of parents did not have a high school degree, and 5% of parents held a graduate degree. The majority of parents were born in Taiwan (89% of mothers; 100% of fathers), with the remaining mothers originally from China (6%), Vietnam (4%), or Indonesia (1%).

Procedure

In March of the children's preschool year, Taiwanese research assistants assessed child participants' behavioral regulation, math, and vocabulary in two sessions. Assessments were given to children individually in a quiet, unused classroom or multipurpose room with most sessions lasting fewer than 30 minutes. In most cases there were 2 or 3 data collectors individually assessing children in different areas of the same room, but otherwise, there were no parents or teachers observing the assessment. After

completing the three tasks, children were given stickers. During the same period of time, teachers completed a questionnaire measuring children's behavioral regulation and parents completed a background questionnaire. Teachers and parents were also given small gifts for their participation in the study.

Measures

All measures that were not previously used in Taiwan were translated into Mandarin Chinese by two native Taiwanese professors of Child and Family Studies at a university in Taiwan. Both professors are fluent in English and Mandarin Chinese. Measures were back translated by bilingual graduate students living in Taipei.

Background questionnaire. Parents completed a short background questionnaire including questions about their education level, country of origin, status level of jobs, ages, and the number of people in the child's household. Background questionnaires were completed by mothers (77%), fathers (11%), mothers and fathers together (11%), and one grandfather (1%).

Direct measure of behavioral regulation. The HTT was used to measure behavioral regulation. The 5-minute task requires children to use attention, working memory, and inhibitory control skills to do the opposite of given commands (Ponitz, McClelland, Jewkes et al., 2008). After hearing the instructions, children are asked to respond to four practice questions to make sure that they understand the instructions. If they answer a practice question incorrectly, they are given up to three reminders of the instructions and then the test is administered whether or not the child answered the practice questions correctly. Once the practice questions end, the 10 task items begin.

There are two versions of the HTT. In the Head-Toes version, children are asked to touch their head (or their toes) but instead, are instructed to do the opposite and touch their toes (or their head). In the Knees-Shoulders version, children are asked to touch their knees (or their shoulders) but instead, are instructed to do the opposite and touch their shoulders (or their knees). Half of the children were given the Head-Toes version of this task, and half were given the Knees-Shoulders version of the HTT. Commands on the task were consistently randomly ordered and scoring was 0 (incorrect), 1 (self-correct), or 2 (correct) points. Self-corrects were scored if children made any discernable motion to the incorrect response, but then corrected themselves. The sum of scores for the 10 items was computed and possible scores ranged from 0 – 20 points. In the present study, there were no significant differences between the two versions of the task after controlling for child age $F(1, 151) = .19, p > .05$ (Head-Toes version $M = 7.65, SD = 6.82$; Knees-Shoulders version $M = 8.14, SD = 6.93$).

Inter-rater agreement of the HTT in Taiwan was also investigated in order to determine the reliability of the task. Two research assistants rated fifteen videotapes of the HTT given to children who did not participate in the present study. Weighted kappa (.85, $p < .001$) analyses indicated that inter-rater agreement for overall scores on the HTT was very strong (Landis & Koch, 1977). These results are similar to a recent study conducted in the U.S., where two examiners coded 42 children on the HTT and showed strong agreement in total scores with a weighted kappa of .90 (McClelland, 2007). The HTT was also evaluated for face validity by four Taiwanese experts currently living in Taiwan. Evaluators included two professors of Child and Family Studies, a researcher

specializing in teaching and curricula, and a professor of Child Care. These experts rated the HTT as an easy to understand, culturally appropriate, and useful measure of behavioral regulation for Taiwanese preschoolers.

Teacher-rated measure of behavioral regulation. The Child Behavior Rating Scale (CBRS) was used to obtain teacher ratings of behavioral regulation in the classroom (Bronson et al., 1995). Teachers rated children's behaviors with tasks, peers, and materials. Responses on the CBRS range from one (never) to five (usually/always). Previous research in the U.S. found a behavioral regulation factor consisting of ten CBRS items tapping attention (e.g., "Concentrates when working on a task; is not easily distracted by surrounding activities,"), working memory (e.g., "Observes rules and follows directions without requiring repeated reminders") and inhibitory control (e.g., "Completes learning tasks involving two or more steps (e.g., cutting and pasting) in an organized way.") In previous research conducted in the U.S., the internal reliability coefficient for the behavioral regulation items on the CBRS in the spring of kindergarten ranged from .94 - .95 (Ponitz, McClelland, Matthews et al., 2008). In the present study, the same behavioral regulation factor emerged in factor analyses and items demonstrated similar loadings (.65 - .86) and internal reliability ($\alpha = .94$) to studies in the U.S., which suggested that the measure may have utility in Taiwan (Ponitz, McClelland, Matthews et al., 2008). Items from the CBRS are useful for measuring behavioral regulation because they ask teachers to think about contextually relevant behaviors that require the child to tap multiple components of behavioral regulation, including attention, working memory, and inhibitory control. Previous U.S. research provides evidence of the validity of the

CBRS which was correlated with the amount of time teachers spent on math and language activities, and with tests of children's intelligence (Bronson et al., 1995). In the present study, teachers were asked to complete the CBRS, but were not given specific training on how to define each score category (on a scale of 1-5).

Early math skills. The Test of Early Mathematics Ability-2 (TEMA-2) was used to measure formal and informal math concepts including relative magnitude, counting, calculation, and enumeration (Ginsburg & Baroody, 1990). Children were asked to perform tasks such as counting the number of objects on a page, determining greater than and less than from pictures of dots, and correctly naming numbers. In previous research with Taiwanese 4- and 5-year-olds, the TEMA-2 had internal consistency between .89 and .90 and test-retest reliability between .91 and .94 (Hsu, 1998, 2000). Previous research in Taiwan found that TEMA-2 scores related to parents' math beliefs and behaviors (Huntsinger, Jose, Liaw, & Ching, 1997). Raw scores from the TEMA-2 were used in the present analysis.

Early vocabulary skills. The Peabody Picture Vocabulary Test-Revised (PPVT-R) measures receptive language and was used in the present study to assess children's early vocabulary skills. Children were asked to point to the correct picture as it was named by the research assistants. The PPVT-R has been previously translated into Mandarin Chinese, has been used in previous studies with Taiwanese preschoolers (Sheng, McGregor, & Marian, 2006), and had a split-half reliability from .90-.97 for Taiwanese samples (Lu & Liu, 1998). Previous studies using the PPVT-R in Taiwan have found that scores were significantly related to mothers' language-related interactions with their

children (Wu, 2007) and significant differences in PPVT-R scores for children with and without autism (I.-C. Lee, 2004). Raw PPVT-R scores were used in the present analysis.

Results

The present study examined the psychometric properties of the HTT with a sample of Taiwanese preschoolers by (1) investigating the nature and variability of HTT scores and by (2) examining relations to early math skills and vocabulary skills, and teacher-rated behavioral regulation.

Nature and Variability of Head-to-Toes Task Scores

Results indicated variability in children's overall HTT scores, with a mean of 7.88 points ($SD = 6.86$) out of 20 possible points. Most children had total scores above floor and below ceiling levels (74%; see Table 2), with 21% of children earning 0 points on the task (floor level) and 5% reaching ceiling levels on the task. Item-level variability was also found with individual items ranging from 0 to 2 points and standard deviations ranging from .89 to .97. Although there was a somewhat bimodal distribution for total task scores, skewness (.38) and kurtosis (-1.29) values did not indicate a non-normal distribution (Kline, 2005). In addition, the difficulty of the HTT increased as the task progressed with an average of 47% ($SD = .43$) of the children receiving 0 points on the first three items, 56% ($SD = .40$) receiving 0 points on the middle four items, and 66% ($SD = .43$) receiving 0 points on the last three items. Similarly, the percentage of children receiving 2 points decreased as the task progressed, with 41% ($SD = .40$) receiving 2 points on the first three items, 34% ($SD = .36$) receiving 2 points on the middle four items, and 30% ($SD = .41$) receiving 2 points on the last three items.

In addition to score variability, the HTT captured significant maturational differences in behavioral regulation $F(1, 59) = 6.20, p < .05$, with the oldest children (at least 1 SD above the mean age; $n = 31$) having total scores significantly higher ($M = 10.87, SD = 6.79$) than the youngest children (at least 1 SD below the mean age; $n = 30$; $M = 6.17, SD = 6.73$). There were, however, no significant gender differences on the task, $t(150) = 1.80, p > .05$ (girls $M = 8.86, SD = 7.43$; boys $M = 6.87, SD = 6.09$).

Relations between Behavioral Regulation on the Head-to-Toes Task and Early Math and Vocabulary

Next we examined the relations between HTT scores and early academic outcomes. Correlations indicated that HTT scores were significantly positively related to early math and vocabulary skills (see Table 3). Demographic variables including child age and mother's education were also positively related to early math and vocabulary, but child gender was not.

Due to the nesting of children in classrooms, intra-class correlation coefficients were calculated for early math and vocabulary (.02 for both outcomes). These coefficients were not statistically significant and indicated that only 2% of the variance in math and vocabulary ($\chi^2 = 13.28, p > .05, \chi^2 = 11.39, p > .05$, respectively) was attributable to differences among classrooms. In addition to low classroom-level variance, the number of classrooms in the present investigation was relatively small. Both of these factors led us to analyze the data using ordinary least squares regression rather than multi-level modeling.

The effects of directly measured behavioral regulation (HTT scores) on early math and vocabulary skills were analyzed, controlling for child age (centered), mother's education level (centered), and teacher-rated behavioral regulation. Participants with higher behavioral regulation on the HTT had significantly higher early math skills than their lower-scoring peers (see Table 4). Teacher-rated behavioral regulation, child age, and mother's education level were also significantly related to higher math skills (see Table 4). This regression model explained 23% of the variance in early math skills of Taiwanese preschoolers. The pattern of relations was similar for predictors of early vocabulary skills. Directly measured behavioral regulation (HTT), child age, and mother's education level were positive predictors of children's early vocabulary skills (see Table 4) when controlling for all other variables. In contrast to results for early math skills, teacher-rated behavioral regulation was not significantly related to early vocabulary skills. These results accounted for 18% of the variation in Taiwanese preschoolers' early vocabulary skills.

Relations between Behavioral Regulation on the Head-to-Toes Task and Teacher Ratings

To further investigate the psychometric properties of the HTT, we examined relations between the direct measure and teacher-rated measure of behavioral regulation. Descriptive analyses revealed that patterns of teacher-rated behavioral regulation scores were different from those found in HTT scores. For example, in contrast to a lack of significant gender differences found with the HTT when controlling for age $F(1, 13) = 2.70, p > .05$, girls were rated significantly higher by their teachers ($M = 3.94, SD = .63$) compared to boys ($M = 3.66, SD = .70$) on a scale from 1 to 5 $F(1, 13) = 5.05, p < .05$.

Similarly, although the HTT captured age differences, the teacher ratings did not differ for children who were 1 *SD* above and below the mean age of 3.82 years $F(1, 59) = 3.69$, $p > .05$, although there was a trend toward significance. In other words, the HTT was more sensitive to age differences compared to the teacher ratings, although the teacher ratings captured gender differences that were not present in HTT scores.

We also calculated the amount of variability in HTT scores and teacher ratings with the coefficient of variation by dividing the standard deviations by the means. The HTT had a larger coefficient of variation (87%) than the teaching ratings of behavioral regulation (18%). In other words, the direct measure of behavioral regulation revealed substantially greater variability than the teacher-rated measure. Further, intra-class correlations for the HTT and teacher-ratings indicated that there was substantially more variation at the classroom level for the teacher-rated measure than the HTT (ICC=.56 and .01, respectively). This suggests that there was much greater variability between classrooms for teacher ratings than for the HTT.

Finally, correlations and regression coefficients were examined for the HTT and teacher ratings of behavioral regulation. Scores on the two measures were not significantly correlated (see Table 3). Further, the HTT was not significantly related to teacher ratings of behavioral regulation after controlling for mother's education and child age ($\beta = .03$, $p > .05$). This model accounted for 4% of the variation in teacher ratings of Taiwanese preschoolers' behavioral regulation.

Discussion

The present study examined the psychometric properties of the HTT for Taiwanese preschoolers, including relations to early achievement and teacher-rated behavioral regulation. Results indicated that the HTT captured variability, and age differences, but not significant gender differences in Taiwanese preschoolers' behavioral regulation. Children's performance on the HTT was also significantly related to early math and vocabulary skills, but not to teacher ratings of behavioral regulation. These results suggest that the HTT offered predictive value beyond teacher ratings in the case of math achievement, and had a significant effect on vocabulary where the teacher ratings did not. Together these results indicate that the HTT may have potential as a screening tool in Taiwan to identify children that may need additional support.

Nature and Variability of Behavioral Regulation

Taiwanese children's scores on the HTT showed substantial variability within a sample of 3.5- to 4.5-year-olds. In addition, an increase in incorrect responses as the task progressed suggests that the children found it easier to pay attention, remember instructions, and inhibit their behaviors at the beginning of the task than during later items. This pattern supports the HTT as a measure of behavioral regulation which taps attention, working memory, and inhibitory control (McClelland, Cameron, Connor et al., 2007).

Previous research in the U.S. has found that the task captures the most variability in children's behavioral regulation between 4 and 5 years of age (Ponitz, McClelland, Jewkes et al., 2008). The current study involved slightly younger children, which could explain the higher amount of children who scored zero on the task. Specifically, about

one-fifth of the children earned a total of zero points on the ten-item task. This number of children earning zero points is similar to those found in a previous study in which 33% of four-year-old children in one U.S. location and 30% of four-year-old children in another U.S. location earned zero points on the HTT (Ponitz, McClelland, Jewkes et al., 2008). In that study the percent of children scoring at ceiling (6% and 3%) was also similar to the present study (5%).

One explanation for the group of children struggling with the task may be the HTT's requirement for a motor response rather than a verbal response. In previous research, preschoolers have had higher scores on behavioral regulation tasks that ask for a verbal response than those that ask for a motor response (Bell & Livesey, 1985; Livesey & Morgan, 1991). Zelazo and his colleagues (1995) refer to this phenomenon as *abulic dissociation*. It is possible that motor response tasks, such as the HTT, may be more challenging for children, and that children earning zero points in the present study may have scored higher on a verbal response behavioral regulation task. We were unable to directly address this issue in the present study, but future studies should include tasks that require verbal responses in addition to the HTT.

The patterns of HTT scores also reflected hypothesized relations regarding child age and gender. Specifically, the HTT was sensitive to age differences such that older children had higher scores than younger children. This finding is similar to findings from research conducted with older children in the U.S., Taiwan, China, and South Korea (Ponitz, McClelland, Jewkes et al., 2008; Son et al., 2008). Also, no significant gender differences were found on the HTT, which is similar to findings in China, Taiwan, and

South Korea with older children using an extended version of the HTT (Xuezhao Lan et al., 2008) but differs from findings in the U.S. (Ponitz, McClelland, Jewkes et al., 2008). It should be noted that in the present study, mean scores on the HTT were somewhat higher for girls than boys, but were not significantly higher. The discrepancy in gender differences between the U.S. and Asia could be due to a number of factors including differing cultural expectations for behavior (Tamis-LeMonda et al., 2002). For example, in one study, Taiwanese parents had higher expectations for their children to be persistent, assertive, and able to concentrate than European American parents (Jose et al., 2000). This finding suggests that Taiwanese parents may expect all children, regardless of gender, to be more regulated in their behavior than in the U.S. Taken together, the variability in HTT scores, distribution of scores as the task progressed, differences between older and younger children, and lack of gender differences support the utility of the measure for assessing the behavioral regulation of Taiwanese 3.5 to 4.5-year-olds and suggest that the HTT may have the potential to be used in the future as a screening tool to identify young Taiwanese children that need additional support.

Relations between Behavioral Regulation on the Head-to-Toes Task and Early Math and Vocabulary

Results of the present study provide evidence that behavioral regulation was related to academic achievement in Taiwan, and that a direct assessment of behavioral regulation can provide information about children's academic success. Specifically, in the present study, children with stronger behavioral regulation measured by the HTT had significantly higher early math and vocabulary skills, after controlling for background

variables and for teacher ratings of behavioral regulation. These findings suggest that behavioral regulation is important for the development of early math and vocabulary skills.

However, it is also possible that the HTT is a proxy for other constructs such as motor skills, language comprehension, or cognitive advancement. Although plausible, evidence from previous research as well as the present study suggests that these explanations are unlikely. For example, the HTT could be seen primarily as a measure of gross motor skills. However, previous research has found that motor skills are significantly but weakly related to academic skills (Son & Meisels, 2006) and in the present study, the HTT was strongly related to academic skills. Second, the HTT does require children to have a certain level of language comprehension in order to understand the task instructions. However, if the HTT were a measure of language comprehension, we would expect it to be more strongly correlated with vocabulary scores than was found in the present study ($r = .28$). Finally, although overlap likely exists between behavioral regulation and intelligence, it is unlikely that the HTT is a proxy for cognitive advancement. This is supported by previous studies finding that aspects of behavioral regulation uniquely predict academic achievement beyond the influence of intelligence (Blair, 2006; McClelland et al., 2006; McClelland et al., 2000). Moreover, Blair (2006) argues that fluid cognitive functioning, which includes aspects of behavioral regulation, is related to general intelligence, but is distinct and more amenable to experience. Finally, although some researchers have used scores on the PPVT vocabulary assessment as a proxy for general intelligence (Rowe, Jacobsen, & Van den Oord, 1999), in the present

study, relations were relatively weak between PPVT and HTT scores. Thus, although the HTT requires some degree of skills in multiple areas, it does not seem to be a primarily a measure of motor, language, or cognitive skills.

Relations between Behavioral Regulation on the Head-to-Toes Task and Teacher Ratings

Results of the present study found a number of differences when comparing the HTT direct measure to the teacher ratings of behavioral regulation. First, the HTT captured substantially greater variability in behavioral regulation than teacher ratings, which is important for understanding differences among children. Second, although no significant gender differences were found for Taiwanese preschoolers on the HTT, teachers rated girls as having significantly higher behavioral regulation than boys. In contrast, research in the U.S. has documented gender differences in both teacher ratings (Ready, LoGerfo, Burkam, & Lee, 2005) and direct measures of behavioral regulation (Ponitz, McClelland, Jewkes et al., 2008).

Additional discrepancies surfaced in relations between behavioral regulation and academic achievement. Specifically, although the HTT was significantly related to early math and vocabulary skills, the teacher ratings of behavioral regulation were significantly related to early math, but not vocabulary skills. This finding differs from previous research using teacher ratings in the U.S. that found significant links to early math, vocabulary and reading skills (McClelland et al., 2000). The consistent relation between the HTT and two types of academic achievement suggests that the HTT may be well-suited for identifying children's likelihood of doing well in school. Finally, children's performance on the HTT was not significantly related to teacher ratings of behavioral

regulation. This finding is similar to results from older children in China and Taiwan using an extended version of the HTT (Wanless, McClelland, Son et al., 2008), but contrasts with recent research in the U.S. and South Korea suggesting that children with higher scores on the HTT had higher teacher ratings of behavioral regulation (Ponitz, McClelland, Jewkes et al., 2008; Wanless, McClelland, Son et al., 2008).

The discrepancies in HTT and teacher-rated behavioral regulation in the present study may reflect a number of factors. To begin with, the direct measure and teacher-rated measure of behavioral regulation may be measuring different aspects of behavioral regulation. The direct measure requires children to pay attention to the instructions, remember the instructions throughout the task, and inhibit tendencies to respond in alternate ways. The teacher-rated measure asks the teachers to rate children's abilities on similar skills (paying attention, remembering instructions, inhibiting behavioral responses), but also may reflect other related skills. For example, one teacher-rated item, "Finds and organizes materials and works in an appropriate place, when activities are initiated" may reflect organizational skills. Other items such as, "Takes time to do his/her best on a task", and "Returns to unfinished tasks after interruption" may reflect persistence. Although these items tap elements of behavioral regulation, they may also capture other skills and explain why children's teacher-rated scores were not significantly related to HTT scores.

Another possible explanation is that teacher and classroom characteristics and cultural expectations may influence the way that teachers rate children. This possibility is supported by one study demonstrating that teacher's amount of teaching experience, self-

efficacy, teacher-child ratios in classrooms, and the locations of classrooms in school settings were related to teachers' ratings of preschoolers social behaviors (Mashburn et al., 2006). Other research on cultural differences has found that Taiwanese teacher ratings of children's behaviors may reflect cultural expectations, more than direct assessments, and more than for teachers in the U.S. (Jose et al., 2000). These teacher ratings may reflect the modesty and high academic expectations for children that are characteristic of Asian cultures (K. Lee, Xu, Fu, Cameron, & Chen, 2001; Stevenson et al., 1990). In sum, teachers in Taiwan could be rating children's behavioral regulation relative to cultural expectations, adjusting their ratings of children according to modesty and high academic standards.

Finally, the high amount of variability between teachers in different classrooms likely contributed to their lack of correlation with HTT scores. Variability between classrooms was particularly high for the teacher ratings of behavioral regulation, compared to the low variability between classrooms on all of the direct measures used in the present study (including the HTT, and the early math and vocabulary assessments). In other words, the teacher-rated measure showed substantial differences between classrooms and the direct measures found smaller differences. Together, these results suggest that the HTT may be a useful measure of behavioral regulation in Taiwan but further research comparing direct measures and teacher-rated measures is needed to determine whether significant gender differences in behavioral regulation and differential predictability for academic outcomes exist for Taiwanese preschoolers.

Practical Implications

The present study suggests that the HTT may be a useful measure for assessing Taiwanese children's behavioral regulation, and for predicting their early academic skills. The HTT is suited for classroom use because it is quick to administer, does not require special materials or computers, can be reliably given after brief training, and approximates tasks seen in early childhood learning settings. These task characteristics and the task's significant relations to early academic achievement suggest that the HTT has potential for further development as a screening tool that could be used to identify young Taiwanese children that may need additional support from teachers and parents to be successful in school.

Findings also indicate that supporting behavioral regulation is important for the academic success of children in high-achieving cultures such as Taiwan. Interventions that support behavioral regulation may be particularly valued because the Taiwanese culture places such a great emphasis on academic achievement. Aspects of interventions aimed at increasing behavioral regulation in the classroom could include providing emotional support and verbal feedback to the child when he or she is having difficulty regulating his or her behavior, and setting up opportunities for the child to engage in socio-dramatic roles that require levels of behavioral regulation that are challenging to the child (McClelland, Cameron, Wanless et al., 2007). In the classroom, recent research has shown that interventions to improve behavioral regulation can influence child outcomes. For example, Diamond and her colleagues found that the *Tools of the Mind* curriculum increased preschool children's attention, memory, and inhibitory control. This curriculum included socio-dramatic play, encouraging children to use private speech, and

teaching the children to draw pictures to increase their abilities to remember and to pay attention (Diamond, Barnett, Thomas, & Munro, 2007). In another example, preschoolers with low behavioral regulation in the fall, who participated twice each week for 8 weeks in games that required them to integrate attention, memory, and inhibitory control, had stronger behavioral regulation in the spring of the preschool year, compared to children who did not receive the games (Tominey & McClelland, 2008). Together, these results suggest that interventions aimed at improving behavioral regulation may have real benefits for strengthening these skills prior to formal schooling. Assessing behavioral regulation in reliable and valid ways in countries such as Taiwan is the first step towards identifying and intervening with children who may need help with behavioral regulation and early achievement.

Limitations

Findings from the present study lay the foundation for use of the HTT in Taiwan, but a number of limitations should be noted. First, the participants in this study all attended public preschools and lived in an urban area in Taiwan which may have limited the diversity of the sample. In future work, it would be important to include children from all types of preschools and geographic areas in Taiwan, in order to get a more accurate picture of the presence and variability of behavioral regulation in Taiwanese preschoolers. Second, the present analyses only include data from one time point. To fully understand links between behavioral regulation and academic achievement and to examine growth in behavioral regulation, data from more than one time point are needed.

Third, only one teacher-rated measure and one direct measure of behavioral regulation were used in the present study. For the teacher-rated measure, it is possible that some of the items may have tapped behavioral regulation as well as other related skills. Using additional teacher-rated measures of behavioral regulation and providing teacher training to increase the reliability of the teacher-rated measure may clarify the relation between HTT scores and teacher-rated scores. Future research should also include additional direct measures to shed light on the overlap of domains that may be captured by the HTT. For example, additional direct measures that require verbal responses may clarify the role that gross motor skills play in influencing HTT scores. Also, direct measures that did not require children's language skills would directly test the possibility that the HTT may be capturing language comprehension. Finally, controlling for general intelligence would also allow researchers to disentangle the relation between HTT scores and cognitive advancement. In sum, findings from the present study lay the groundwork for future use of the HTT with young children in Taiwan. Future research that included additional measures of behavioral regulation would shed light on the most practical and valid ways to measure these skills in early learning settings.

Conclusion

Findings from the present study extend previous research on behavioral regulation in the U.S. and suggest that the HTT may be a useful measure of behavioral regulation, predictor of early academic success, and have potential as a screening tool in Taiwan. Taiwan's strong cultural emphasis on academic achievement makes it particularly

important for Taiwanese teachers to be able to identify children that need additional academic support. Results from this study suggest that the HTT is a reliable, easy-to-use measure of behavioral regulation for Taiwanese preschoolers. Children's behavioral regulation scores on the HTT demonstrated variability in Taiwanese 3 and 4 year olds, captured age differences, and were significantly related to early math and vocabulary achievement. Taiwanese parents, teachers, researchers, and policy makers can use these findings to better support children's early school success.

Table 1.1

Descriptive Statistics for Predictor and Outcome Variables (N=152)

	<i>M</i>	<i>SD</i>	<i>Range</i>
Direct Measure of Behavioral Regulation (HTT; Total Score)	7.88	6.86	0 – 20.00
Teacher-rated Behavioral Regulation (Mean Score)	3.68	.70	2.00 – 5.00
Child Age (in Years)	4.09	.29	3.58 – 4.58
Mother's Education Level (Less than High School Degree = 1, High School Degree =2, College Degree =3, Graduate Degree =4)	2.47	.62	1.00 – 4.00
Vocabulary	29.80	12.97	4.00 – 72.00
Math	12.34	8.10	0 – 39.00

Table 1.2

Head-to-Toes Task Scores (N=152)

	Percent of Children Attaining Each Score	Age of Children in Years Attaining Each Score, <i>M (SD)</i>	Age Range of Children in Years Attaining Each Score
0 (Floor)	21	4.00 (.28)	3.58 – 4.50
1-9	40	4.07 (.28)	3.58 – 4.50
10-19	34	4.16 (.29)	3.58 – 4.58
20 (Ceiling)	5	4.25 (.33)	3.67 – 4.50

Note. HTT is the Head-to-Toes Task.

Table 1.3

Correlations between Background, Predictor, and Outcome Variables (N=152)

Variables	1	2	3	4	5	6	7
1. Direct Measure of Behavioral Regulation (HTT)	-						
2. Teacher-rated Behavioral Regulation	.07	-					
3. Child Age (in Years)	.20*	.20*	-				
4. Gender (Girl =0, Boy=1)	-.15 [†]	-.18*	-.03	-			
5. Mother's Education	.05	-.01	.01	.08	-		
6. Vocabulary	.28***	.02	.20*	-.10	.28***	-	
7. Math	.31***	.27***	.30***	-.05	.20*	.39***	-

Note. HTT is the Head-to-Toes Task. [†] $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 1.4

Effects of HTT on early math and vocabulary skills (N=152)

<i>Math Skills</i>	ΔR^{2a}	<i>Coefficient</i>	<i>SE</i>	β	<i>t-value</i>
Intercept	---	.51	3.46	---	.15
Centered Child Age (months)	.09***	.51	.17	.22	2.89**
Centered Mother's Education	.03*	2.09	.89	.17	2.35*
Teacher-rated Behavioral Regulation	.05**	2.53	.88	.21	2.87**
Direct Measure of Behavioral Regulation (HTT)	.06**	.29	.09	.24	3.30**
<i>Vocabulary Skills</i>	ΔR^{2a}	<i>Coefficient</i>	<i>SE</i>	β	<i>t-value</i>
Intercept	---	31.06	5.72	---	5.43***
Centered Child Age (months)	.04*	.64	.29	.17	2.20*
Centered Mother's Education	.08***	5.37	1.47	.27	3.65***
Teacher-rated Behavioral Regulation	.00	-1.22	1.46	-.06	-0.84
Direct Measure of Behavioral Regulation (HTT)	.05**	.45	.14	.24	3.11**

Note. HTT is the Head-to-Toes Task. * $p < .05$; ** $p < .01$; *** $p < .001$. ^a To calculate ΔR^2 , independent variables were added individually, in the order they are listed in the table.

MEASURING BEHAVIORAL REGULATION IN FOUR SOCIETIES

Shannon B. Wanless^aMegan M. McClelland^aClaire C. Ponitz^bSeung-Hee Son^cXuezhao Lan^dAlan C. Acock^aFrederick J. Morrison^dJo-Lin Chen^eFu-Mei Chen^eKangyi Lee^fMiyoung Sung^gSu Li^h

^a Oregon State University, Department of Human Development and Family Sciences. ^b University of Virginia, Center for the Advanced Study of Teaching & Learning (CASTL). ^c Purdue University, Department of Child and Family Studies. ^d University of Michigan, Departments of Education and Psychology. ^e Fu Jen Catholic University, Department of Child and Family Studies. ^f Seoul National University, Department of Child and Family Studies. ^g Seokyeong University, Department of Child Studies. ^h Chinese Academy of Sciences. 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, and the Oregon State University Department of Human Development and Family Sciences.

Abstract

High academic achievement in societies such as Taiwan, South Korea, and China may be better understood by examining the role of behavioral regulation (the integration of attention, working memory, and inhibitory control) in early academic skills. Few reliable and valid measures of behavioral regulation, however, have been developed and evaluated outside of the United States (U.S.). In the present study, the psychometric properties of a direct measure of behavioral regulation, the Head-Toes-Knees-Shoulders Task (HTKS), were examined with 3 to 6 year old children in the U.S., Taiwan, South Korea, and China. Specifically, we investigated the nature and variability of HTKS scores, and examined relations between the HTKS and early math and vocabulary skills, and teacher-rated behavioral regulation skills. Results indicated that the HTKS captured substantial variability in all four societies. HTKS scores were also significantly and positively related to early math, vocabulary, and early literacy skills after controlling for important child and demographic variables including teacher-rated behavioral regulation. Finally, HTKS scores were significantly related to teacher ratings of behavioral regulation in the U.S. and South Korea but not in Taiwan and China. These findings suggest that the HTKS is a useful measure of behavioral regulation for young children in these four societies, and that behavioral regulation seems to be important for early academic success in the U.S. and Asian countries. Results and discussion note the possibility that in some societies, the HTKS and teacher-rated measures may tap different aspects of behavioral regulation.

The United States (U.S.) repeatedly scores below average on international academic assessments, while some societies including Taiwan, South Korea, and China, consistently score at or near the top (Baldi et al., 2007; Mullis, Martin, Gonzalez, & Chrostowski, 2004). 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 et al., 1997; Stevenson et al., 1993). These studies, however, have not measured behavioral regulation (the integration of attention, working memory, and inhibitory control), which is a significant predictor of academic achievement in the U.S. throughout preschool and elementary school (Blair & Razza, 2007; McClelland et al., 2006; McClelland, Cameron, Connor et al., 2007). It is possible that behavioral regulation has not been measured in previous studies because reliable and valid measures do not exist outside the U.S. In the present study, the psychometric properties of a direct measure of behavioral regulation, the Head-Toes-Knees-Shoulders Task (HTKS), were examined with children in the U.S., Taiwan, South Korea, and China, including relations between the HTKS and early achievement and teacher-rated behavioral regulation.

We investigated the use of the HTKS in academically-oriented societies such as Taiwan, South Korea, and China based on previous research documenting the importance of early behavioral regulation for academic achievement (Ponitz, McClelland, Matthews et al., 2008; Wanless, McClelland, Acock, Chen, & Chen, 2008). Children in these Asian societies are intently focused on educational success and their daily schedules typically include extracurricular academic classes and long hours of studying to improve their

exams scores and class rankings (Bao, 2004; Dwyer, 2004; Yi & Wu, 2004). Reliable and valid assessments for young children, however, are limited and early childhood experts in Asia have called for the development of psychometrically sound and culturally appropriate tools to help identify children that may need additional help to be academically successful (Kim et al., 2003; Tsai et al., 2006). Identifying a measure of behavioral regulation that predicts academic achievement and could be developed into a screening tool 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, Cameron, Connor et al., 2007; McClelland, Cameron, Wanless et al., 2007). 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 moving ahead in line to get water more quickly. Attention, working memory, and inhibitory control individually and collectively contribute to behavioral regulation and to the school success of young children.

Attention is the ability to focus on a task while ignoring distractions (Rothbart & Posner, 2005; Rueda et al., 2005). Research in the U.S. suggests that young children need to attend to activities amidst distractions in order to succeed academically in vocabulary acquisition and particularly in math (Dixon Jr. & Salley, 2007; Duncan et al., 2007;

Howse, Lange et al., 2003). Working memory is the ability to remember and apply information while encountering and processing new stimuli (Gathercole & Pickering, 2000). This skill has been related to math and language skills (Adams et al., 1999; Espy et al., 2004) and is particularly important during the transition to formal schooling (Senn et al., 2004). Finally, inhibitory control is the ability to stop one response in favor of a more adaptive behavior (Dowsett & Livesey, 2000). Preschoolers' are often expected to use inhibitory control to follow classroom rules and this skill has been linked to math and early literacy outcomes (Blair & Razza, 2007; van der Schoot et al., 2004).

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 frequently expected to orchestrate these skills for learning. This may be particularly true in Asia where classroom characteristics such as high teacher child ratios and frequent teacher lectures place high demands on behavioral regulation (Hsieh, 2004; Kim et al., 2003; Pang & Richey, 2007). However, reliable and valid measures of behavioral regulation are not available in academically-oriented societies such as Taiwan, South Korea, and China. In the present study, we extend research on a measure of behavioral regulation called the HTKS, and examine its psychometric properties in the U.S. and these Asian societies.

Behavioral Regulation and Early Academic Achievement

Behavioral regulation has repeatedly been related to early academic success in preschool and early elementary school concurrently and over time (Blair & Razza, 2007; Howse, Calkins et al., 2003; McClelland, Cameron, Connor et al., 2007; McClelland et

al., 2000; Ponitz, McClelland, Matthews et al., 2008). In the U.S., one study found that gains in behavioral regulation over the prekindergarten year significantly predicted improvement in early literacy, math, and vocabulary skills (McClelland, Cameron, Connor et al., 2007). In Taiwan, four-year-olds' behavioral regulation also significantly predicted early math and vocabulary skills (Wanless, McClelland, Acock et al., 2008). Moreover, in some studies, the relation between behavioral regulation and early academic achievement has been particularly pronounced for early math skills (Blair & Razza, 2007; Duncan et al., 2007; Ponitz, McClelland, Matthews et al., 2008).

In U.S. research using the HTKS to measure behavioral regulation, higher kindergarten behavioral regulation in the fall was significantly related to higher math, early literacy, and vocabulary scores in the spring (Ponitz, McClelland, Matthews et al., 2008). For U.S. first graders, similar relations were present for HTKS behavioral regulation and reading comprehension and vocabulary skills (Connor et al., 2008). 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, only includes one study from one Asian society which used a simple version of the HTKS called the Head-to-Toes Task (HTT) developed for younger children. Further, previous research has not systematically compared the use of the HTKS across societies by using participants of similar age ranges and including similar outcomes. In the present study, we address these issues by examining links between HTKS behavioral regulation and early academic skills in the U.S, Taiwan, South Korea, and China.

Considering Cultural Context

A growing body of evidence suggests that Asian preschoolers may have particularly strong behavioral regulation skills including attention, working memory, and inhibitory control (Oh & Lewis, 2008; Sabbagh et al., 2006). For example, researchers found that preschoolers in China had significantly stronger skills than their U.S. counterparts, including inhibitory control measured by the Day/Night task, the Grass/Snow task, the Bear/Dragon task, and the Whisper task, as well as working memory measured by the KRISP task (Sabbagh et al., 2006). In fact, these Chinese preschoolers were scoring on par with U.S. children that were six months older. Preschoolers in South Korea also had stronger inhibitory control skills measured by the Day/Night task and by Luria's Hand Game than their Western counterparts (Oh & Lewis, 2008). Further, this trend was evident in China and South Korea when examining social behaviors that are related to behavioral regulation, such as cooperation and compliance (Farver, Kim, & Lee, 1995; Orlick, Zhou, & Partington, 1990). In addition, research by Ahadi and colleagues (1993) found that Chinese children had less variability in attention, inhibitory control, and impulsivity than their counterparts in the U.S. Taken together, these findings suggest that beyond academic skills, Asian children may also have stronger behavioral regulation skills than Western children, although no research has used a direct measure that assesses overall behavioral regulation.

Measuring Behavioral Regulation

The majority of previous research on behavioral regulation in the U.S. and abroad has used measures that do not directly assess the integration of attention, working

memory, and inhibitory control. Instead, research has traditionally measured behavioral regulation with teacher ratings or as an aggregated score comprised of individual attention, working memory, and inhibitory control tasks (Bronson et al., 1995; Howse, Calkins et al., 2003). Although these methods have been useful for understanding teachers' perceptions of children's behavioral regulation or for researching individual components of behavioral regulation, they have important limitations. First, teacher ratings often differ from scores on other assessments including child and parent ratings, and direct assessments of executive functioning (Kunter & Baumert, 2006; Loo & Rapport, 1998; Mahone et al., 2002; Mahone & Hoffman, 2005; Wall & Paradise, 1981). Teacher ratings in Asia have also been found to be different than direct measured child behaviors (Jose et al., 2000; Wanless, McClelland, Acock et al., 2008). Further, teacher-rated assessments often rely on likert scales which have been found to be limited for cross-cultural comparisons because culturally-based teacher expectations for children's behaviors may influence ratings in a way that makes them difficult to compare across cultures with different values (Heine et al., 2002). As a result, direct assessments may offer a unique perspective on child behaviors that may not be captured by teacher ratings.

Second, 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 that relations among the different tasks measuring attention, working memory, and inhibitory control are often weak (Archibald & Kerns, 1999; Espy & Bull, 2005; Oh & Lewis, 2008). We suggest that children's ability to orchestrate these three skills may be

qualitatively different from their ability to employ each skill individually, and may be a key predictor of academic success.

The HTKS Task is a direct measure designed to capture the integration of attention, working memory, and inhibitory control. In the HTKS, children are asked children to do the opposite of four paired commands: “touch your head” and “touch your toes”; “touch your knees” and “touch your shoulders.” One pair of commands is used in the first part of the task, and the second pair is incorporated in the second part of the task. This 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 setting, which reflects the way that children must use these skills in classroom settings. It is also quick to administer, does not require special materials, and can be given reliably in classroom settings. These characteristics demonstrate the potential of the HTKS viable as a screening tool to help identify children that made need additional academic support.

Goals of the Present Study

Research suggests that behavioral regulation is important for early academic success in the U.S., but less research has examined these links in high-achieving Asian societies such as Taiwan, South Korea, and China. Measures that reliably and validly measure behavioral regulation across societies are needed to investigate this skill and its relations to academic achievement. In the present study, the psychometric properties of 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, and (2) relations to early math, vocabulary, and early literacy skills, and to teacher-rated behavioral regulation.

We hypothesized that the HTKS would capture substantial variability in behavioral regulation in all four societies, as has been found with previous research in the U.S. (Ponitz, McClelland, Matthews et al., 2008). Further, based on previous studies using the HTKS in the U.S. and the HTT (a simplified version of the HTKS) in Taiwan, we anticipated that scores on the HTKS behavioral regulation task would be significantly related to early math, vocabulary, and early literacy skills in all four societies. It was less clear whether scores on the HTKS would relate to teacher-rated behavioral regulation because results from previous research have been somewhat mixed. In the U.S., teacher ratings were significantly related to the HTKS, but research using a simpler form of the HTKS (the HTT) in Taiwan, did not find significant relations between direct and teacher-rated behavioral regulation (Ponitz, McClelland, Matthews et al., 2008; Wanless, McClelland, Acock et al., 2008). Similarly, in previous research in Taiwan, teachers had rated children differently than research assistants' observations of children's behavior (Jose et al., 2000). Therefore, it seems plausible that HTKS and teacher-rated behavioral regulation would be related in the present study, but that the relation would be weaker in the Asian societies than in the U.S.

Method

Participants

Data for the present study were collected in four different societies: the U.S., Taiwan, South Korea, and China. Parents and teachers of all children volunteered to

participate in this study. 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, and each sample's age range overlapped 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 - 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 because no significant differences in HTKS scores were found between the two sites (see McClelland, Cameron, Connor et al., 2007 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 U.S. sample, 19% ($n = 58$) of children came from minority ethnic groups (non-Asian and non-Caucasian).

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), 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 a subgroup of children had mothers with a high school degree or less ($n = 80$). The majority of children had parents who were born in Taiwan (77% of mothers, 100% of father), with the remaining mothers originally from China (4%), Vietnam (4%), Indonesia (1%), or the Philippines (1%).

South Korea. The 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 childcare centers in Seoul (the capital city of South Korea) and Kyonggi province. Less than half of these children were girls (40%, $N = 91$), and mothers had an average education level between a high school and college degree. Eighty-eight of the mothers in this sample had a high school degree or less. 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 skill 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). In all four samples, teachers completed a questionnaire rating behavioral regulation for each child in their

classrooms in the fall of the school year, except in the U.S. where it was completed in the spring of the school year. In addition, the Head-Toes-Knees-Shoulders Task (HTKS), and academic tasks (math, early literacy, and vocabulary) were given to children in the fall of the school year. The school year in South Korea, however, begins in the spring and ends in the winter, so the fall assessment did not represent the beginning of the school year for these children. All academic assessments were included 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.

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, 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, this information was collected from school records.

Direct Measure of Behavioral Regulation

The HTKS was used in the present study to directly measure behavioral regulation. Children were given the HTKS in their native language. Each of the 20 items may be scored with 0 for an incorrect response (touching head when asked to touch head), 1 for a self-correct (initially responding incorrectly, but correcting themselves), or a 2 for a correct response (touching toes when asked to touch head). Total scores on the HTKS range from 0 to 40 points. There are two forms of the HTKS which are equivalent except for the order of the initial items. Form 1 starts with head-toes items and Form 2 starts with knees-shoulders items. In the present study, there were no significant differences between forms in any the U.S., Taiwan, or China when controlling for age, (p 's > .05), which is consistent with previous research in the U.S. and Taiwan (Ponitz, McClelland, Matthews et al., 2008; Wanless, McClelland, Acock et al., 2008). In South Korea, only one version of the task (Head-Toes) 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. the HTKS ($\alpha = .98$; Connor et al., 2008), and the Head-to-Toes Task, a simple version of the HTKS (weighted $\kappa = .90$; McClelland, 2008) have shown strong inter-rater reliability in previous research. In the present study, there were no significant differences between examiners in children's HTKS scores after controlling for child age in Taiwan, $F(28, 155) = 1.05, p > .05$. In South Korea, the HTKS was rated by two research assistants for a sub-sample of the participants ($n = 72$) and the raters' scores were significantly correlated ($.71, p < .001$). This information was

not available for the other samples. In, previous research on the Head-to-Toes Task in Taiwan has demonstrated strong inter-rater reliability (weighted $\kappa = .85$, $p < .001$; Wanless, McClelland, Acock et al., 2008).

Teacher-rated measure of Behavioral Regulation

In all four samples, the Child Behavior Rating Scale (CBRS) was used to measure teacher-rated behavioral regulation (Bronson et al., 1995). Items on the CBRS ask teachers to rate the children's behaviors when using materials, interacting with peers, and completing tasks, using a scale of 1 (never) to 5 (usually/always). To determine whether a behavioral regulation factor was present in each of the four societies, CBRS scores were analyzed using principle axis factor analysis with a promax rotation. In each of the four samples, the first dimension that emerged was the same 10-item behavioral regulation factor that has been found in other research in the U.S. ($\alpha = .94 - .95$; Matthews, Ponitz, & Morrison, 2008; Ponitz, McClelland, Matthews et al., 2008), and in Taiwan ($\alpha = .94$; Wanless, McClelland, Acock et al., 2008). In the present study, the CBRS behavioral regulation factor in each society had strong reliability (U.S.: $\alpha = .94$; Taiwan: $\alpha = .94$; South Korea: $\alpha = .94$; China: $\alpha = .95$). 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 behavioral regulation factor ranged from 1 to 5 with higher scores demonstrating higher levels of behavioral regulation.

Academic Achievement

United States. Math, literacy, and vocabulary subtests from the Woodcock Johnson Psycho-Educational Battery-III Tests of Achievement (Woodcock & Mather, 2000) or the Spanish version, the Bateria Woodcock-Muñoz-R (Woodcock & Muñoz-Sandoval, 1996) were used for children in the U.S. The Applied Problems subtest assessed early math 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 words, and the Picture Vocabulary subtest uses pictures to assess expressive vocabulary. W-scores were used in order 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 U.S., inter-rater reliability of these subtests is greater than .85 (Woodcock & Mather, 2000).

Taiwan. Early math and vocabulary skills were assessed in Taiwan with measures that had previously been translated into Mandarin 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 counting objects on a page, determine greater than and less than, and to correctly identify numbers (Ginsburg & Baroody, 1990). In previous research in Taiwan, 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. The PPVT-R has had a split-half reliability of .90 - .97 in Taiwanese samples (Lu & Liu, 1998).

South Korea. Early math and early vocabulary were measured with subtests of the Korean-Wechsler Preschool and Primary Scale of Intelligence (K-WPPSI; Park, Kwak, & Park, 1989). The math subtest measured relative magnitude, counting, and calculation. In previous research, the Korean math 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. 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 Hangeul Word Reading in which children are asked to pronounce two-syllable Korean words and non-words (Choi & Yi, 2007). The internal consistency of 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 math 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 math composite score. Counting and calculation scores in the present study were significantly correlated ($r = .44, p < .001$). The task had previously been translated into Mandarin Chinese. In previous research counting and calculation tests had a reliability of .84 and .87, respectively, and were correlated with teacher reports and cognitive tasks (X. Lan, personal communication, July 24, 2008). For the vocabulary task, all traditional characters were translated into Mandarin characters, and the children were asked to read the characters aloud.

Results

The present study investigated the psychometric properties of a direct measure of behavioral regulation, the HTKS, in the U.S., and high-achieving societies including Taiwan, South Korea, and China. Specifically, we examined (1) the nature and variability of behavioral regulation measured with the HTKS in each society, and (2) relations to early math, vocabulary, and early literacy skills, and to teacher-rated behavioral regulation.

Missing data and multiple imputation. In the present study, 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 math, early literacy, and vocabulary scores. 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 analyses or as auxiliary variables, and any remaining missingness is at random (Schafer & Graham, 2002). Although there is no way to definitely test this assumption, auxiliary variables not included in the analyses, but that might explain missingness, were chosen for each sample and included in the imputation model. These mechanism variables included the amount of preschool experience, number of hours in

preschool each week, family income, or whether the child spoke Spanish or was an ethnic minority. For variables with more than 5% missing data, logistic regressions were calculated between possible auxiliary variables and dummy variables that indicated whether variables were missing or not. In the U.S. and Taiwan, the missingness of variables that had more than 5% missing data (mother's education and teacher-rated behavioral regulation) was not significantly predicted by any of the auxiliary variables. In South Korea, variables missing more than 5% of the data (math, early literacy, vocabulary, and mother's education) were significantly predicted by amount of preschool experience, so preschool experience was included as an auxiliary variable in the multiple imputation model. In China, no variables had more than 5% missing data. Although this is not definitive, these analyses suggest that the missing data in the present study may be missing at random (Acock, 2005; Meng, 1995; Rubin, 1996).

In order to obtain unbiased parameter estimates, multiple imputation (using Stata) was used to create 10 imputed datasets (Acock, 2005). Descriptive statistics for each society using the original and imputed data are presented in Table 1. Results presented in all analyses were based on the imputed data. Descriptive statistics, correlations, and regression coefficients were based on information from all 10 imputed datasets. In addition, correlation significance levels were adjusted to account for the inflated sample size (see Table 2). Standardized betas were calculated by averaging the betas for regressions on each of the 10 imputed datasets. Finally, R-squares were calculated by averaging the natural log of each R-squared from regressions on each of the 10 imputed datasets, and exponentiating this mean (Acock, 2007). These calculations were done

because R-squares have a fixed upper limit of 1.0 and can not be assumed to be normally distributed across the 10 imputed datasets.

Nature and Variability of Behavioral Regulation measured by the HTKS

Our first research question examined the nature and variability in HTKS scores in four samples of children. Variability in HTKS scores in each society was found with China showing the least and Taiwan showing the most 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 on the task (40 points). China ($N = 8$, 6.7%) and the U.S. ($N = 7$, 2.3%), had more children, but still a relatively small number, reach 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 distribution was the most skewed, with many high scoring children, and Taiwan was the only society with a positive skew, with the majority of children in Taiwan had low scores. Taiwanese children, however, were younger and U.S. children were older on average than their counterparts in other samples.

To compare average HTKS scores across societies, we calculated HTKS means adjusted for age using coefficients from a multiple regression model with centered child age and an indicator variable for each society (except the reference society) as the predictors. After controlling for child age, China had a significantly higher HTKS mean

than the other three samples, the U.S. and South Korea did not have significantly different means, and Taiwan had a significantly lower HTKS mean than the other three samples (see Table 3). HTKS scores were also examined by age group (see Table 4). Three-year-olds were not included in this comparison because there were few children in this age group outside of South Korea. For four-year-olds, two distributions of scores emerged. First, the U.S. and China had similar score distributions for this age group, with most of the children scoring between 20 and 39 points (the task ranged from 0-40). In Taiwan and South Korea, however, four-year-olds' scores were more evenly distributed across the range of the task. For five-year-olds, the majority earned between 20 and 39 points, with some children in South Korea earning 1-19 points, and some children in the U.S. earning 1-19 or zero points. There were not enough five-year-olds in Taiwan to compare to the other samples. The large majority of six-year-olds in the U.S. and South Korea (there were not enough in Taiwan or China for comparisons) earned between 20-39 points on the HTKS.

HTKS relations to Math, Vocabulary, and Early Literacy

To investigate our second research question, we used correlations and regressions within each sample to determine whether HTKS scores related to early academic outcomes. In all samples that measured math (U.S., Taiwan, South Korea, China) and vocabulary (U.S., Taiwan, South Korea), higher HTKS scores were significantly related to higher math skills, and higher HTKS scores were significantly related to higher vocabulary scores (see Table 2). In addition, HTKS scores were strongly related to early

literacy scores in South Korea and moderately related in the U.S., but were weakly related to early literacy in China. Early literacy was not assessed in Taiwan.

Although children in all four samples were nested in preschool classrooms, there were too few classrooms in each sample to allow for multi-level modeling. Due to the small number of classrooms we analyzed the data within each sample using ordinary least squares regression. Thus, our final regression models examined the relations between directly measured behavioral regulation (HTKS) on early math, vocabulary, and early literacy skills for each sample controlling for demographic variables and teacher-rated behavioral regulation (see Table 5). In all four societies, HTKS scores were significantly related to early math skills and the regression models explained between 28% (Taiwan) and 57% (South Korea) of the variance in early math skills. In the three samples that assessed early literacy (U.S., China, and South Korea), HTKS scores also significantly predicted early literacy skills and these models explained between 11% (China) and 39% (South Korea) of the variance in early literacy. Finally, HTKS scores significantly predicted vocabulary skills in the U.S. and Taiwan, but not in South Korea. This model in the U.S. accounted for 24% of the variance in vocabulary skills, and 17% of the variance in Taiwan (vocabulary was not assessed in China). In the U.S., HTKS scores were a larger predictor of math and early literacy skills than all other child demographic predictors, but in Asia, child age consistently had a modestly larger effect on math and early literacy than HTKS scores.

Relations between HTKS direct measure scores and Teacher-rated Behavioral Regulation

To better understand the psychometric properties of the HTKS in the four societies, we compared HTKS scores to teacher-rated behavioral regulation scores. In the U.S. and in South Korea, higher scores on the HTKS were significantly related to higher teacher-rated scores (U.S. = .29, South Korea = .23; see Table 2). In Taiwan and China, however, relations between HTKS and teacher-rated scores were weak and not significant (Taiwan = .09, China = .12).

When comparing age-adjusted mean HTKS and age-adjusted mean teacher-rated behavioral regulation scores across samples, differences emerged between the direct measure and the teacher ratings. In contrast to HTKS mean scores, where China had the highest scores and Taiwan had the lowest scores, teacher ratings were highest in Taiwan and the U.S. Further, mean Taiwanese and U.S. scores in these samples were not significantly different from one another (see Table 3). Teacher ratings in China and South Korea, however, were relatively low and were not significantly different from each other. We also compared the amount of variability in HTKS scores and teacher ratings by calculating the coefficient of variation (standard deviation divided by the mean). In all four societies, the HTKS captured greater variability than the teacher ratings of behavioral regulation (see Table 3), with the largest difference in HTKS and CBRS coefficients of variation in Taiwan (67% difference) and the smallest difference in China (8% difference).

Discussion

The present study examined the psychometric properties of the HTKS Task in four societies, and relations between HTKS scores, early achievement, and teacher-rated

behavioral regulation in each sample. Overall, there were many similarities in HTKS psychometric properties across all four groups, but some differences also emerged. The HTKS captured variability in behavioral regulation in each society, showing the most variability in Taiwan and the least in China. Scores on the HTKS were also significantly related to early math, vocabulary, and early literacy, with the exception of vocabulary in South Korea, beyond the teacher-rated measure, and more strongly than the teacher-rated measure. Finally, HTKS scores were differentially related to teacher-rated behavioral regulation with a significant relation present in the U.S. and South Korea but not in Taiwan or China.

Nature and Variability of Behavioral Regulation

The HTKS proved to be a useful task for directly measuring behavioral regulation in all four samples, but some differences emerged. The HTKS captured individual variation in behavioral regulation in all four samples; the greatest amount of variability was found in Taiwan, and the least variability was found in China. The relatively low variability in HTKS scores in China may be due to the fact that overall, children earned higher scores and some children approached ceiling levels on the task (although only 8 children reached ceiling levels). This finding is congruent with previous research that Chinese children have shown less variability than U.S. children in many related areas including inhibitory control, impulsivity, and attention (Ahadi et al., 1993). Further, teaching practices in China tend to place less emphasis on individuality and more emphasis on whole group instruction in which, “all children are expected to do the same thing at the same time” (Pang & Richey, 2007, p.3). This practice may lead to lower

variability in Chinese children's scores than in other societies. The relatively high variability in Taiwan may relate to the influence of Western values emphasizing independence, such as is seen by the use of the National Association for the Education of Young Children's developmentally appropriate teaching practices in Taiwan (McMullen et al., 2005). Variability in Taiwan's HTKS scores, however, exceeded that of the U.S. sample which has a particularly strong emphasis on independence. This may be due to the relatively young age of the Taiwanese sample and the rapid rate of behavioral regulation growth during younger years (Ponitz, McClelland, Jewkes et al., 2008).

In addition to high variability, the Taiwanese sample had the highest number of children scoring zero points. South Korean children, although fewer than in Taiwan, also had a relatively high number of children with zero points on the task. One possible reason for this finding is that children in Taiwan and South Korea were younger compared to children in the U.S. and China. Age is not solely responsible for a score of zero points, however, as is demonstrated by the Chinese sample which had a mean age and age range similar to South Korea but did not have many children with zero points. Previous research suggests that the gross motor demands of the HTKS may have been more challenging than verbal demands on behavioral regulation, and this characteristic of the task may account in part for the relatively large number of zeros in South Korea and Taiwan (Bell & Livesey, 1985; Livesey & Morgan, 1991). It is also possible that the children who scored zero points did not understand the instructions for the HTKS. The task began with practice and training items during which the children were reminded of the instructions to do the opposite, up to three times. After three reminders, the research

assistant continued to administer the task whether the child gave correct responses or not. Therefore, it is possible that the children who earned zero points on the task did not understand the instructions, but were tested anyway. However, in previous research using the HTT, the simpler version of the HTKS, the majority of children who earned zero points on the testing section of the task, earned at least one point in the practice section (Ponitz, McClelland, Jewkes et al., 2008).

HTKS scores adjusted for age indicated that China had significantly higher behavioral regulation scores than the other three samples. The U.S. and South Korea had similar HTKS scores, and Taiwan had significantly lower scores than the other three samples. China's high scores reflect similar findings from previous work using multiple executive functioning tasks (Sabbagh et al., 2006). South Korea's similarity to the U.S., however is somewhat different than previous findings that South Korean preschoolers' executive functioning was stronger than their Western counterparts. This study, however, was comparing South Korean children to British children and not to U.S. children (Oh & Lewis, 2008). Both South Korea and Taiwan's HTKS scores may have been particularly low due to the relatively large number of children who earned zero points on the task.

Overall, the results of the present study suggest that the HTKS captured variability in preschoolers' behavioral regulation in all four samples. Cultural differences, however, were found related to the degree of variability in HTKS scores and to child age. These differences may indicate that the task would be most useful for slightly different age groups in each society. In the U.S., the small number of 4-year-olds in the present sample makes it difficult to determine the usefulness of the task for this age group.

Findings from the present study support the use of the HTKS with U.S. 5-year-olds and suggest that further research should consider its use with a larger sample of 4-year-olds. The high percentage of U.S. 6-year-olds who had high scores on the task suggests that the task may not be useful for differentiating older U.S. children. In Taiwan, the task may be most useful with children that are four-years-old or older due to the relatively high number of children that earned zero points in this age group. The Taiwanese sample in the present study was limited to mostly 4-year-olds, so the usefulness of the HTKS with older Taiwanese children can not be determined. Due to the fairly even distribution of scores for 4-year-olds, however, it seems likely that the task may be useful for 5-year-olds as well. Also, in South Korea, a relatively large number of 4-year-olds earned zero points on the HTKS, suggest that the task may not be useful for younger children. Similarly, the high number of 6-year-olds who earned high scores in South Korea suggests that the task may not provide differentiating information for South Korean children older than 6 years. Finally, in China, where four-year-olds generally scored in the upper range of the task, the HTKS may also be useful for three-year-olds. The high percentage of Chinese 5-year-olds that had high scores on the task indicates that the task may not be useful children older than 5 years. Samples with broader age ranges need to be assessed in future research to further clarify the younger and older age limits of the task in each society. Taken together, these findings suggest that the HTKS may be useful for measuring behavioral regulation across multiple societies, but that the age ranges in which the task is most informative may be culturally specific.

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

In the present study, HTKS scores predicted early math in all four samples, early literacy in all three samples that it was measured (U.S., South Korea, and China), and early vocabulary in each sample that it was measured (U.S. and Taiwan), with the exception of South Korea. This relation was present after controlling for teacher ratings of behavioral regulation and other important background variables. The predictive utility of the HTKS is particularly noteworthy given that these effects were found after taking teacher ratings into account. Further HTKS effects on academic domains were consistently stronger than the effects of the teacher ratings. Findings suggest that behavioral regulation is important for academic success in all four societies, and that the HTKS may be useful for identifying children that may struggle in early math, vocabulary, and early literacy. Specifically, consistent and significant relations between HTKS scores and academic achievement point to the possibility of developing the HTKS into a screening tool to identify children needing additional academic support. Since early childhood experts in Taiwan and South Korea have expressed the need for reliable and valid assessments for young children, an HTKS screening tool may be particularly useful in Asian societies (Kim et al., 2003; Tsai et al., 2006).

Of all of the academic domains tested in each sample, vocabulary in South Korea was the only domain not significantly related to HTKS behavioral regulation scores. In contrast, HTKS scores in South Korea explained a greater amount of variance in math and early literacy skills than in any other society. These results may need to be

interpreted with caution, however, due to the relatively large amount of missing data in South Korea's academic outcomes (33.9% of early math, vocabulary, and early literacy). Further, the lack of a significant relation between HTKS scores and vocabulary in South Korea may reflect the vocabulary measures that were used in each sample. Although the vocabulary measures in the U.S., Taiwan, and South Korea (vocabulary was not assessed in China) all required children to either recognize or produce the names of pictures, the South Korean measure also asked children to define words. In other words, attention, working memory, and inhibitory control, skills that underlie behavioral regulation, may relate to children's ability to name words, but not be as important for defining words.

In the U.S. and China, HTKS scores were more strongly related to early math than vocabulary or early literacy. In Taiwan and South Korea, relations between HTKS scores and early math, vocabulary, and early literacy were more similar among outcomes, apart from not finding a significant relation between HTKS scores and vocabulary in South Korea. The strong relation between HTKS scores and early math fits other similar findings in the U.S. (Blair & Razza, 2007; Duncan et al., 2007). Previous researchers finding similar relations between HTKS scores and math skills in the U.S. suggested that behavioral regulation may be particularly important for math skills because relatively little time is spent on math instruction compared to early literacy instruction in schools (Ponitz, McClelland, Matthews et al., 2008). In other words, child factors such as behavioral regulation may be particularly important when children are exposed to fewer opportunities to learn math skills. Findings from the present study, together with findings from previous research, suggest that the skills children need to be successful on the

HTKS (attention, working memory, and inhibitory control) are also needed to be successful on early math assessments, as well as vocabulary and early literacy.

Overall, the findings from the present study indicate the importance of HTKS behavioral regulation for early academic success in the U.S., Taiwan, South Korea, and China, beyond teacher-ratings. Although this relation has previously been found in the U.S., findings from the present study suggest that the HTKS may be useful cross-culturally for understanding children's early school success. Specifically, results support the potential of the HTKS as a screening tool to help early childhood professionals in the U.S., Taiwan, South Korea, and China identify children that may need additional help in order to succeed in school. This may be also useful in the U.S. where many children are struggling with academic skills compared to their counterparts in other countries (Baldi et al., 2007; Mullis et al., 2004). Finally, teachers and parents in Taiwan, South Korea, and China may also find the HTKS particularly valuable because of the strong emphasis that these cultures place on academic success and the need for reliable and valid screening tools for young children (Bao, 2004; Dwyer, 2004; Kim et al., 2003; Tsai et al., 2006; Yi & Wu, 2004).

Relations between Behavioral Regulation on the HTKS and Teacher Ratings

Results of this study found culturally-specific evidence of relations between the HTKS and teacher ratings of behavioral regulation. In the U.S. and South Korea, there were significant positive relations between both types of behavioral regulation assessments. In Taiwan and China, however, these measures were not significantly related. In previous research, HTKS scores were significantly related to teacher ratings in

the U.S., but not related in Taiwan (Ponitz, McClelland, Matthews et al., 2008; Wanless, McClelland, Acock et al., 2008).

Teacher-ratings of behavioral regulation also differed from HTKS scores in the amount of variability they captured. Overall, the HTKS captured more variability than the teacher-ratings of behavioral regulation in all samples. Differences between the amount of variability captured by the HTKS and teacher-ratings were greatest in Taiwan and smallest in China. In other words, in Taiwan, the teacher-ratings explained a relatively small amount of the variability in children's behavioral regulation compared to the HTKS direct measure. In China, the HTKS captured less variability than it did in other samples although the task still detected more variability in this sample than the teacher-ratings of behavioral regulation. The finding that the HTKS consistently captured more variability than the teacher-rated measure in every sample suggests that it may be more useful for differentiating children's levels of behavioral regulation skills.

Finally, teacher-ratings also differed from HTKS behavioral regulation scores in their predictability of academic outcomes. In all four samples and for all academic outcomes assessed, the HTKS was more consistently and more strongly predictive of academic skills than teacher-rated behavioral regulation. Children's scores on this task, therefore, may provide teachers and parents with richer information about their children's academic achievement than they would gain from teacher reports. There are two reasons that may explain the differences between these two types of assessments. First, the HTKS may measure a different aspect of regulation than teacher ratings: an aspect which is more relevant for academic success. An accumulation of experiences with the children in

the classroom may affect teachers' ratings, even if those experiences do not solely reflect behavioral regulation. For example, children may have difficulty with behavioral regulation when they feel frustrated or angry, but these situations incorporate aspects of emotion regulation that are not part of the HTKS. Teacher-ratings, in this situation, would reflect behavioral regulation as well as other related components, instead of only assessing behavioral regulation. Second, teacher-ratings may reflect child, classroom, teacher, or cultural characteristics that may not affect HTKS scores (Mashburn et al., 2006). One cross-cultural study of likert scales, for example, suggested that cultural ideas shaped teachers expectations for the child characteristics that represented each score on a likert scale, and these expectations led to the *reference-group effect*. This meant that the lack of a common reference group made cross-cultural comparisons of likert scale scores problematic (Heine et al., 2002). Taken together, results comparing HTKS scores and teacher-ratings suggest that in all four samples, the HTKS provides a unique window on children's academic achievement, beyond that of teacher ratings.

Overall, the usefulness of the teacher-rated measure for cross-cultural research remains unclear. Although this measure has been used in multiple samples in the U.S. and its properties seem to be similar in South Korea, more information is needed in order to understand the use of the CBRS in Taiwan and China. Preliminary results in Taiwan and China, however, suggest that its practical utility for predicting early academic skills may be limited in these two societies. It is suggested that future research use this measure with caution in Taiwan and China until its psychometric properties have been further

clarified. Moreover, researchers should use multiple teacher-rated measures in Taiwan and China in order to shed further light on their results from additional perspectives.

Practical Implications

The need to understand factors that predict academic success remains critical as the U.S. lags behind other high-achieving countries like Taiwan, South Korea, and China (Baldi et al., 2007; Mullis et al., 2004). Findings from the present study indicate that the HTKS direct measure of behavioral regulation demonstrated strong psychometric properties: it captured variability in all four samples and was significantly predictive of academic achievement, beyond the effect of teacher-rated behavioral regulation and important background variables. Further, the HTKS is particularly useful for assessment in classrooms because it is easy-to-administer, can be given in less than ten minutes, and requires no special materials or a computer to use. The task also reflects demands that are seen in the classroom, making it particularly ecologically valid and attractive for researching behaviors that are relevant to classroom contexts.

These findings suggest that the HTKS may be developed into a screening tool for identifying children that may need additional help to succeed in school in these societies. It is important to recognize, however, that screening tools sometimes have the unintended consequence of labeling children as “low-performing” or “at-risk”, when in fact they may develop typically. These labels can help children receive additional services that will help them improve their behavioral regulation and thus school readiness, but they can also hold them back from participating in challenging classes or lead to teachers and parents treating them differently based on a perceived limitation in their skills. In societies that

have strong academic orientations, such as South Korea, Taiwan, and China, the negative consequences of earning a low score on a screening tool that is meant to predict academic achievement, may be even greater than in the U.S. Parents in these societies may be particularly concerned if their child earns a low HTKS score, and may place additional social and emotional pressure on the child to achieve academically. Although this concern should be seriously considered, screening tools are still valuable in academically-oriented societies because of their ability to identify children that legitimately need additional or different types of academic support. Before the HTKS can be used as a screening tool in any society, careful consideration should be paid to labeling children and useful ways to present scores to parents and teachers.

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 et al., 2000). Results from the present study support this assertion and suggest that in all four societies, including three with particularly high-academic outcomes, supporting behavioral regulation development in preschool may improve children's school readiness. As such, attention to promoting the development of behavioral regulation in early childhood settings across societies is warranted. Behavioral regulation interventions used in the U.S. have shown initial signs of improving children's skills. Specifically, 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 (Diamond et al., 2007). In addition, preschool games designed to help

children practice attention, working memory, and inhibitory control skills played over an eight week period were shown to improve children's HTKS behavioral regulation scores (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 behavioral regulation games, especially for children who have low HTKS scores.

Limitations

The present study revealed a number of findings about the use of the HTKS for measuring behavioral regulation across four societies. There are a few limitations, however, that should be noted. First, all of the participants in this study from Asian societies lived in urban areas which could have limited the diversity of the sample. 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, age groups and ranges were not consistently represented. This limitation made it difficult to compare behavioral regulation distribution, and floor and ceiling effects across samples. Age-adjusted means, however, did allow for some cross-cultural comparison to be made. In future research, age ranges should also 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 math, 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, data used in the present study was only from one time point. Future research should use longitudinal data in order to examine growth in behavioral regulation and relations to achievement over time. Fifth, using more than one direct measure of behavioral regulation and more than one teacher-rated measure would provide additional information about the differences in these two types of assessment. Specifically, the use of more than one direct measure may help clarify whether children earning zero points on the HTKS did not understand the instructions, despite multiple reminders. Finally, although all samples used the HTKS and the CBRS teacher-rated scale, the academic measures differed. For the present study, academic measures were chosen based on their reliability and validity within each society and not for cross-culturally consistency. This may have improved the accuracy of measurement within each society, but made comparisons more difficult. However, it is noteworthy that such consistency was found across samples and measures in relations between HTKS scores and achievement outcomes. Future research should consider using consistent measures across all samples to better address this issue.

Conclusion

This preliminary study examined the psychometric properties of a direct measure of behavioral regulation, the HTKS, in four societies. Findings suggest that the HTKS was reliable, captured variability, and predicted academic achievement in the U.S., Taiwan, South Korea, and China. Cultural differences emerged in the specific age ranges

that were best suited for the HTKS, and in relations between HTKS directly measured behavioral regulation and teacher-rated behavioral regulation. Overall, results suggest that the HTKS is a reliable and valid measure that can be used in the U.S., Taiwan, South Korea, and China to further examine behavior regulation in preschoolers. These findings extend previous research on behavioral regulation in the U.S. and provide the foundation for research on behavioral regulation in high-achieving societies including Taiwan, South Korea, and China to help ensure that all children are successful in school.

Table 2.1

Descriptive statistics by society for original data and imputed data, M (SD), and percent missing in original data

	United States <i>N</i> = 310		Taiwan <i>N</i> = 158		South Korea <i>N</i> = 227		China <i>N</i> = 119	
	Original	Imputed	Original	Imputed	Original	Imputed	Original	Imputed
Child Age	5.5 (.3) 0.6%	5.5 (.3)	4.6 (.3) 0.6%	4.6 (.3)	5.1 (.9) 3.1%	5.1 (.9)	5.0 (.6) 0.0%	5.0 (.6)
Gender	0.5 (.5) 0.0%	0.5 (.5)	0.5 (.5) 0.0%	0.5 (.5)	0.6 (.5) 0.0%	0.6 (.5)	0.5 (.5) 0.0%	0.5 (.5)
Mother's Education	3.3 (.7) 19.4%	3.3 (.8)	2.5 (.5) 13.3%	2.5 (.6)	2.6 (1.1) 28.6%	2.7 (1.1)	--- ---	---
HTKS	26.3 (11.1) 0.0%	26.3 (11.0)	15.8 (13.2) 0.6%	15.8 (13.1)	24.0 (13.0) 0.0%	24.0 (12.9)	31.8 (8.8) 0.0%	31.8 (8.8)

Table 2.1 *continued*

CBRS	4.0 (0.7)	4.0 (0.7)	3.9 (0.6)	3.9 (0.6)	3.8 (0.7)	3.8 (0.7)	3.8 (0.8)	3.8 (0.8)
	35.5%		13.9%		0.4%		1.7%	
Math	434.0 (15.3)	434.0 (15.3)	18.1 (8.7)	18.1 (8.6)	12.8 (3.9)	12.9 (4.01)	6.0 (2.8)	6.0 (2.8)
	0.6%		1.3%		33.9%		0.8%	
Vocab	477.4 (13.5)	477.5 (13.5)	38.4 (15.8)	38.4 (15.7)	13.7 (6.1)	14.0 (6.2)	---	---
	0.6%		0.6%		33.9%			
Early	375.3 (30.2)	375.4 (30.2)	---	---	28.8 (30.5)	28.3 (31.8)	23.6 (20.3)	23.6 (20.2)
Literacy	0.6%				33.9%		0.0%	

Note. Mother's education is coded as (1 = 0-8 years, 2 = >8-≤12 years, 3 = >12-≤16 years, 4 = >16 years)

Table 2.2

Correlation Matrix

	HTKS	Vocabulary	Math	Early Literacy
US (<i>N</i> = 310)				
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 (<i>N</i> = 158)				
Child Age	.20*	.27***	.37***	---
Gender	-.06	-.13	-.17*	---
Parent Education	.18*	.19*	.10	---
CBRS	.09	.12	.33***	---
HTKS	---	.30***	.34***	---
South Korea (<i>N</i> = 227)				
Child Age	.53***	.51***	.68***	.55***
Gender	-.03	.03	.06	-.11 [†]
Parent Education	.22***	.20**	.26***	.11 [†]
CBRS	.23***	.13*	.22***	.26***

Table 2.2 *continued*

HTKS	---	.36***	.59***	.50***
China (<i>N</i> = 119)				
Child Age	.24**	---	.42***	.23**
Gender	-.09	---	-.19*	.01
CBRS	.12	---	.15	.14
HTKS	---	---	.40***	.24**

[†] $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

Table 2.3

HTKS direct measure and CBRS teacher-rated measure descriptives

		United States	Taiwan	South Korea	China
<i>HTKS</i>	Mean	23.82 _a	19.57	24.43 _a	32.34
	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
	<i>CBRS</i>	Mean	3.27 _b	3.29 _b	3.08 _c
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

Note. HTKS is the Head-Toes-Knees-Shoulders Task. CBRS is the Child Behavior Rating Scale. Means are adjusted for child age. Adjusted means with the same subscript are not significantly different from each other.

Table 2.4

HTKS Score Distribution

Percent of Children in Each Age Group, Scoring Within Each HTKS Range, by Society										
	U.S.			Taiwan	South Korea				China	
<i>n</i>	4yrs	5yrs	6yrs	4 yrs	3yrs	4yrs	5yrs	6yrs	4yrs	5yrs
	34	257	19	145	29	78	79	41	44	69
0	---	4	---	24	---	15	---	---	---	---
Floor										
1-19	---	21	---	35	39	29	14	---	---	---
20-39	71	72	95	41	29	55	84	92	79	90
40	---	---	---	---	---	---	---	---	---	---
Ceiling										

Note. Cells are defined if there are more than 10 children in the category.

Table 2.5

Coefficients (standard errors) and betas from regressions of early math, vocabulary, and early literacy skills on behavioral regulation by society

	<i>United States</i>			<i>Taiwan</i>		<i>South Korea</i>			<i>China</i>	
	<i>Math</i>	<i>Vocab</i>	<i>Early Literacy</i>	<i>Math</i>	<i>Vocab</i>	<i>Math</i>	<i>Vocab</i>	<i>Early Literacy</i>	<i>Math</i>	<i>Early Literacy</i>
R^2	.3	.2	.2	.3	.2	.6	.3	.4	.3	.1
Intercept	417.3*** (2.8)	462.1*** (2.6)	354.0*** (5.9)	21.2*** (2.0)	44.7*** (3.9)	11.8*** (.6)	13.8*** (1.5)	19.6* (7.7)	3.3*** (.9)	8.8 (7.3)
HTKS	.6*** (.1)	.3*** (.1)	.6*** (.2)	.2** (.1)	.3** (.1)	.1** (.0)	.0 (.1)	.6* (.2)	.1*** (.0)	.4* (.2)
	$\beta = .4$	$\beta = .2$	$\beta = .2$	$\beta = .2$	$\beta = .2$	$\beta = .3$	$\beta = .1$	$\beta = .2$	$\beta = .3$	$\beta = .2$

Table 2.5 *continued*

Child	4.4 [†]	7.3***	12.5*	8.1***	12.1**	2.5***	3.5***	15.3***	1.5***	6.8*
Age	(2.4)	(2.3)	(5.1)	(2.2)	(4.2)	(.3)	(.6)	(3.5)	(.4)	(3.1)
(years)	$\beta = .1$	$\beta = .2$	$\beta = .1$	$\beta = .3$	$\beta = .2$	$\beta = .5$	$\beta = .5$	$\beta = .4$	$\beta = .3$	$\beta = .2$
Gender	4.2**	1.2	3.3	-1.6	-3.8	-.7	.1	-6.1	-.4	4.3
(boys=1, girls=0)	(1.6)	(1.4)	(3.3)	(1.3)	(2.5)	(.4)	(.8)	(5.1)	(.5)	(3.8)
	$\beta = .1$	$\beta = .1$	$\beta = .1$	$\beta = -.1$	$\beta = -.1$	$\beta = -.1$	$\beta = .0$	$\beta = -.1$	$\beta = -.1$	$\beta = .1$
Parent's	4.4**	5.8***	7.2**	.8	4.1 [†]	.7**	1.0*	1.2	---	---
Educ.	(1.2)	(1.0)	(2.3)	(1.1)	(2.1)	(.2)	(.4)	(2.0)		
	$\beta = .2$	$\beta = .3$	$\beta = .2$	$\beta = .1$	$\beta = .2$	$\beta = .2$	$\beta = .2$	$\beta = .0$		

Table 2.5 *continued*

CBRS	3.3*	-.1	5.8*	3.1**	.4	.4	.5	5.5 [†]	.3	3.8
	(1.4)	(1.2)	(2.7)	(1.2)	(2.1)	(.5)	(.9)	(3.0)	(.3)	(2.5)
	$\beta = .2$	$\beta = -.0$	$\beta = .1$	$\beta = .2$	$\beta = .0$	$\beta = .1$	$\beta = .1$	$\beta = .1$	$\beta = .1$	$\beta = .2$
Site	-4.6*	2.7	-5.6	---	---	---	---	---	---	---
(MI=1,	(1.9)	(1.8)	(4.3)							
OR=0)	$\beta = .0$	$\beta = .1$	$\beta = -.1$							

Note. Parent Education is scored as 1 = 0-8 years, 2 = 9-12 years, 3 = 13-16 years, 4 = greater than 16 years. CBRS is the teacher-rated Child Behavior Rating Scale. HTKS is the Head-Toes-Knees-Shoulders Task. Child Age, Parent Education, and CBRS scores were centered. In the U.S. sample, site was also controlled because data were collected in Michigan (MI) and Oregon (OR). [†] < .10; * $p < .05$; ** $p < .01$; *** $p < .001$.

CONCLUSION

Early childhood experts in the U.S. and abroad are calling for increased focus on behavioral regulation as a means of increasing children's school readiness (Kim et al., 2003; Zero to Three, n.d.). Reliable and valid measures of behavioral regulation, however, are not available internationally. The present studies lay the foundation for this research in high-achieving Asian societies including Taiwan, South Korea, and China by investigating the psychometric properties of the Head-to-Toes Task (HTT) and the Head-Toes-Knees-Shoulders Task (HTKS) that measure behavioral regulation in young children. The HTT and HTKS measure behavioral regulation as an integrated, contextually-relevant construct, are practical for use in early childhood classrooms, and have face validity according to early childhood experts from the societies studied.

The first study investigated the reliability and validity of the HTT for 3.5- to 4.5-year-olds in Taiwan. The second study used a more complicated version of the HTT, the HTKS, with 3- to 6-year-olds in the U.S., Taiwan, South Korea, and China. Each study examined the psychometric properties of a direct measure of behavioral regulation by addressing two main research questions within each sample. Specifically, the nature and variability of behavioral regulation, and relations between directly measured behavioral regulation and academic skills, and teacher-rated behavioral regulation were analyzed. In both Study 1 and Study 2, the tasks captured variability in children's behavioral regulation and predicted academic skills independent of teacher ratings. These findings suggest that the HTT and the HTKS may have potential for future development as

screening tools to identify children that may need additional help to be successful in school.

Overview of Findings

Findings from Study 1 indicate that the HTT was reliably administered in Taiwan, captured variability in children's behavioral regulation and significant differences by child age, and significantly predicted early math and vocabulary skills beyond the effect of teacher ratings of behavioral regulation. There were, however, no significant differences in HTT scores by child gender and no significant relation between HTT scores and teacher ratings of behavioral regulation. Possible reasons for this lack of gender differences and lack of a relation with teacher ratings are discussed below. Taken together, these findings indicate that the HTT may be a reliable and valid measure of behavioral regulation for young Taiwanese children.

Findings from Study 2 reflected a similar pattern as Study 1: the HTKS was reliably administered in all four samples, captured significant differences by child age, and significantly predicted early math, vocabulary, and early literacy controlling for teacher ratings of behavioral regulation in all samples that measured these academic skills, except for vocabulary in South Korea. In particular, the HTKS predicted academic outcomes beyond the effect of teacher ratings and more strongly than teacher ratings. Differences between societies emerged in the ages that the task was most appropriate for and relations to teacher ratings of behavioral regulation. Specifically, the HTKS was most appropriate for 5- to 6-year-olds in the U.S., 4-year-olds in Taiwan, 4- to 6-year-olds in South Korea, and 4- to 5-year-olds in China. Future research involving broader age

ranges of children, however, is needed to determine whether the task can be useful for children beyond these age ranges. Finally, HTKS scores were significantly related to teacher-ratings in the U.S. and South Korea, but not in Taiwan or China. In sum, the HTKS was useful for assessing behavioral regulation across all four samples, but some unique properties of the task emerged in each sample.

Variability in HTT and HTKS Behavioral Regulation

In both Study 1 and Study 2, the majority of the children scored between floor and ceiling levels of the tasks and scores represented the full range of the HTT and HTKS. Taiwanese children assessed with the HTT in Study 1 showed an increase in incorrect answers as the task progressed, suggesting that it became more difficult to pay attention, remember the instructions, and inhibit their behaviors over the course of the task. Age differences in HTT and HTKS scores were present in all samples in Study 1 and Study 2 with older children performing better on the task than younger children. Finally, although previous research in the U.S. has found significant gender differences in HTT scores, with girls performing better than boys, these were not found in the young Taiwanese sample in Study 1 (Ponitz, McClelland, Jewkes et al., 2008). Gender differences, although not assessed for the HTKS in Study 2, were examined for the Study 2 samples in a related study. Findings were consistent with Study 1 and previous U.S. research indicating that the U.S. sample in Study 2 showed significant gender differences, but these differences were not found in the Taiwanese, South Korean, or Chinese samples (Xuezhao Lan et al., 2008).

Distributions of task scores were within accepted standards for normality, but were somewhat negatively skewed in the U.S. and China, positively skewed in Taiwan (in both Study 1 and Study 2) and bimodal in South Korea (Kline, 2005). It is important to note that although these different distributions may indicate cultural specificity, they cannot be directly compared because of the different ages in each sample. In Study 1, for example, the Taiwanese sample was relatively young (3.5 to 4.5 years old) which may have been the reason for the greater number of children earning zero points in this sample. In Study 2, there was still a relatively large number of zeros in the Taiwanese sample, but these children were younger than those in the other three samples. Similarly, the U.S. sample in Study 2 may have been negatively skewed because of the relatively older age of the children.

In Study 1, the percentage of Taiwanese children earning floor (0 points, 21%) and ceiling (20 points, 5%) scores was similar to those found in previous studies in the U.S. using the HTT with children of similar ages (Ponitz, McClelland, Jewkes et al., 2008). It is possible that the task's requirement of a motor response, which has previously been found to be more difficult for children than verbal responses, may be one reason that some children scored zero points (Bell & Livesey, 1985; Livesey & Morgan, 1991). It is also possible that the children did not understand the task instructions despite receiving up to three reminders before the testing section began. In Study 2, the differences in ages in each sample limit comparisons of floor and ceiling effects that can be made across societies. The high number of zero scores in Taiwan (24%) and South Korea (15%) suggests that the HTKS may not be useful for Taiwanese and South Korean children who

are younger than 4 years old. Further, no sample in Study 2 had 10 or more children of any age group reaching ceiling levels (40 points) suggesting that the HTKS did not seem to be too simple for the children tested because very few children in any of the four samples earned the highest number of points on the task.

Although cross-cultural comparisons could not be made of task score distributions, comparisons were made of age-adjusted mean scores. In Study 2, China had the highest HTKS scores which were significantly higher than the other three samples and had the smallest variance. This low variance has been found in previous research on Chinese attention, inhibitory control, and impulsivity, and may relate to cultural beliefs and behaviors that emphasize the importance of not distinguishing oneself from the group, or developing independence (Ahadi et al., 1993; Pang & Richey, 2007). The U.S. and South Korea had similar means that were not significantly different, although the South Korean sample had greater variance than the U.S. sample. Taiwan had the lowest HTKS scores which were significantly lower than the other three samples, and they also had the greatest variance. This large amount of variance in Taiwan may reflect the influence of Western cultural values, including teaching practices endorsed by the National Association for the Education of Young Children that emphasize children's independence (McMullen et al., 2005). It is surprising, however, that the Taiwanese sample would demonstrate greater variance than the U.S. sample which heavily emphasizes individualism. It is possible that the relatively young age of the Taiwanese sample compared to the U.S. sample was related to the large amount of variance in Taiwanese HTKS scores. The Taiwanese sample also may have exhibited greater variance because

younger children may be developing behavioral regulation at a more rapid rate which could have differentiated between children who were only months apart in age (Ponitz, McClelland, Jewkes et al., 2008).

Overall, findings from Study 1 and Study 2 indicate that the HTT and HTKS tasks were both useful for detecting variability in behavioral regulation. There were differences in task distributions, variances, and mean scores adjusted for age, suggesting that behavioral regulation may develop differently in each society. These differences, however, may be due in part to the differences in ages in each sample, even though there was overlap in child age across all four samples. Future research should include a variety of behavioral regulation tasks, such as tasks with verbal responses, and larger samples with broader and more consistent age ranges in order to further investigate the validity of the HTT and HTKS and the most useful ages for using the tasks in each society.

Relations between HTT and HTKS Behavioral Regulation and Early Math, Vocabulary, and Early Literacy Achievement

Results from Study 1 and Study 2 indicate that the HTT and HTKS were significantly related to early academic achievement and that a direct assessment of behavioral regulation provided useful information about children's early academic success beyond teacher ratings of behavioral regulation and more strongly than teacher ratings of behavioral regulation. Specifically, findings from Study 1 indicated that HTT behavioral regulation scores were significantly and positively related to early math and vocabulary skills for Taiwanese preschoolers (aged 3.5 to 4.5), when controlling for teacher-rated behavioral regulation and background variables. In Study 2, HTKS

behavioral regulation was significantly and positively related to early math (U.S., Taiwan, South Korea, and China), vocabulary (U.S. and Taiwan), and early literacy skills (U.S., South Korea, and China) in the samples that these domains were measured, except for vocabulary in South Korea. These findings suggest that the HTT and HTKS may show potential as screening tools that would help teachers identify children that need extra academic support. Such tools would be particularly useful in societies such as Taiwan, South Korea, and China where academic success is highly valued but the availability to screening tools for young children is limited (Bao, 2004; Dwyer, 2004; Kim et al., 2003; Sharma, 1997; Stevenson et al., 1990; Tsai et al., 2006; Yi & Wu, 2004).

In Study 2, additional consistency was found in the particularly strong relation between HTKS scores and math in all four samples, and this pattern was most pronounced in the U.S. and China. The finding is congruent with other similar findings in the U.S. which may be related to the increased importance of child factors such as behavioral regulation in light of the relatively limited amount of time that is spent on math instruction compared to early literacy instruction in U.S. schools (Blair & Razza, 2007; Duncan et al., 2007; Ponitz, McClelland, Matthews et al., 2008). One exception to the consistency across Study 1 and Study 2 relations between HTT and HTKS scores and academic skills was seen in the South Korean sample. The lack of relation between South Korean HTKS scores and vocabulary skills may reflect differences in the measurement of vocabulary in this sample. Specifically, the South Korean vocabulary assessment was the only measure across the four samples that asked children to produce a definition of a word. This increased demand of the assessment may draw upon an aspect of vocabulary

skills that is not related to behavioral regulation. Moreover, the high degree of missing data in the South Korean academic outcomes suggests that findings from this sample need to be interpreted with caution, although original and imputed descriptives were highly similar.

With one exception, the overall consistency found across Study 1 and Study 2 suggests that the link between HTT and HTKS behavioral regulation and academic achievement is not culturally specific. In other words, the skills that are needed to be successful on these tasks, including the integration of attention, working memory, and inhibitory control, are also needed for academic success in early math, vocabulary, and early literacy in multiple societies. Despite cultural differences between the U.S. and these Asian societies, including differences in educational systems and practices, behavioral regulation scores on the tasks consistently differentiated children's academic achievement (Hsieh, 2004; Lin & Tsai, 1996; Sang-Hun, 2008; "What the numbers say," 1998). Further, in the U.S., previous research has shown that this relation is present over time. Specifically, in one U.S. study, gains in behavioral regulation over the prekindergarten year significantly predicted improvement in literacy, vocabulary, and math skills (McClelland, Cameron, Connor et al., 2007). Future research investigating the development of these skills over time in Taiwan, South Korea, and China may clarify relations between growth in behavioral regulation and future academic success.

Relations between Behavioral Regulation on the HTT and HTKS and Teacher Ratings

In both studies, a complex pattern of results was found in relations between directly measured and teacher-rated behavioral regulation (CBRS). For example, the HTT

and HTKS consistently captured more variation in children's behavioral regulation than the teacher-rated measure in the U.S., in both samples from Taiwan, in South Korea, and in China. However, in Study 2, the difference in amount of variation in HTKS and teacher-rated scores differed by society and was most pronounced in Taiwan and least pronounced in China. In other words, in Taiwan the tasks captured substantially more variability in children's scores than teacher-ratings, but in China, the two types of assessments captured a more similar amount of variability. This finding suggests that the HTT and HTKS were consistently more sensitive to differences in children's behavioral regulation compared to the teacher-rated measure in all samples, and particularly in Taiwan. As a result, these tasks may prove to be more useful than teacher-ratings for identifying differences among children indicate that they are struggling with behavioral regulation.

Both studies also indicated that the HTT and HTKS were more consistently predictive of all academic outcomes than teacher-rated measures of behavioral regulation. Specifically, in Study 1, the HTT significantly predicted early math and vocabulary scores in Taiwan, after controlling for teacher-rated behavioral regulation and background variables. Teacher ratings, however, significantly predicted early math, but not vocabulary skills. In Study 2, the HTKS significantly predicted all academic outcomes in all samples that they were measured except for early vocabulary in South Korea, after controlling for teacher-rated behavioral regulation and background variables. In other words, the HTKS predicted early math in all four samples, vocabulary in the U.S. and Taiwan, and early literacy in the U.S., South Korea, and China. Teacher-rated

behavioral regulation, however, did not significantly predict early math in South Korea or China, vocabulary in the U.S., Taiwan, or South Korea, or early literacy in China. These differences suggest that the HTT and HTKS may be a more useful measure than teacher-rated measures such as the CBRS for predicting children's school success. It is important to note, however, that there may be measurement issues related to the teacher-rated assessment that led to limited relations to academic achievement.

More specifically, differences in directly measured and teacher-rated behavioral regulation may be related to three issues. First, the two types of assessments may actually assess different aspects of behavioral regulation. The Head-to-Toes Task and HTKS ask children to simultaneously use attention, working memory, and inhibitory control to produce certain behaviors. The teacher-rated measure lists example behaviors that include elements of behavioral regulation, but may also tap additional skills such as organization or perseverance. Teacher's ratings on these items may reflect children's behavioral regulation as well as other similar skills which could contribute to the lack of significant relation between the measures in Taiwan (in both studies) and in China. Second, teachers may rate children differently based on teacher, classroom, cultural, and child characteristics (Mashburn et al., 2006). For example, previous research has found differences in Taiwanese teacher-rated scores of social skills and direct assessments which may reflect cultural values such as modesty that influence ratings (Jose et al., 2000; K. Lee et al., 2001). Third, previous research suggests that likert scales such as the CBRS may be problematic for cross-cultural research due to the *reference-group effect* (Heine et al., 2002). This effect refers to teachers rating children in comparison to other children

and to the norms and values of their particular society. In other words, without training on the use of the likert scale, it is possible that a score of 4 in one society, for example, may not be equivalent to a score of 4 in another society. Although the teacher-rated measure used in the present study has been used in many samples in the U.S., and its properties seem to be similar in South Korea, further research is needed to understand the usefulness of this assessment in Taiwan and China. In particular, the limited predictive utility of the CBRS in Taiwan and China suggests that this measure may need to be used with caution in these societies until further research can clarify whether this measure is appropriate in these societies. Future studies are needed that incorporate additional direct and teacher-rated measures of behavioral regulation to clarify the differences between these types of assessments.

Practical Implications

The focus of this dissertation was to examine the psychometric properties of two related measures of behavioral regulation; a less complicated version designed for younger children (the HTT) and a more complicated version (the HTKS) across multiple societies. Taken together, findings from these studies support the initial validity of both tasks in Taiwan and of the HTKS in the U.S., South Korea, and China. Further, results indicate that these tasks, which have face validity as measures of behavioral regulation across samples, significantly predict academic achievement beyond teacher-ratings of behavioral regulation and other important variables. These findings are important for early childhood professionals in each society because these tasks could help identify children that at-risk for school difficulty. These direct measures are particularly attractive

to early childhood professionals and researchers because they are quick to administer, require no special materials or computer, and approximate demands often seen in classroom settings. Findings from the present research suggests the need for further research into the development of the HTT and HTKS for use as screening tools so that teachers could quickly and effectively identify children that may need additional support to be successful in school.

These findings suggest that the HTKS may be developed into a screening tool for identifying children that may need additional help to succeed in school in these societies. It is important to recognize, however, that screening tools sometimes have the unintended consequence of labeling children as “low-performing” or “at-risk”, when in fact they may develop typically. These labels can help children receive additional services that will help them improve their behavioral regulation and thus school readiness, but they can also hold them back from participating in challenging classes or lead to teachers and parents treating them differently based on a perceived limitation in their skills. In societies that have strong academic orientations, such as South Korea, Taiwan, and China, the negative consequences of earning a low score on a screening tool that is meant to predict academic achievement, may be even greater than in the U.S. Parents in these societies may be particularly concerned if their child earns a low HTKS score, and may place additional social and emotional pressure on the child to achieve academically. Although this concern should be seriously considered, screening tools are still valuable in academically-oriented societies because of their ability to identify children that legitimately need additional or different types of academic support. Before the HTKS can be used as a

screening tool in any society, careful consideration should be paid to labeling children and useful ways to present scores to parents and teachers.

Results from the present study can be used for further cross-cultural research on behavioral regulation. Previously, the vast majority of research on behavioral regulation had been conducted in the U.S. which may have been due to the lack of reliable and valid behavioral regulation measures for other societies. Cross-cultural research may be particularly important for understanding the mechanisms through which children in academically-oriented societies such as Taiwan, South Korea, and China attain high academic success. Results from the present studies suggest that behavioral regulation may play a role in the high academic achievement of these societies.

Finally, the consistent links found between behavioral regulation and academic achievement in the U.S., Taiwan, South Korea, and China suggest that early childhood interventions aimed at strengthening behavioral regulation may be useful for supporting academic development in the U.S. and in these Asian countries. Some examples of interventions in the U.S. include *Tools of the Mind* (a curriculum using sociodramatic play and other learning strategies) or games aimed at improving behavioral regulation, which have been shown to be effective for young children (Diamond et al., 2007; Tominey & McClelland, 2008). These interventions, however, have not yet been tested outside of the U.S. The results of this research may be particularly important for early childhood professionals and parents in Taiwan, South Korea, and China who highly value education and often supplement children's school experiences with additional classes to help children achieve high academic success (Bao, 2004; Sang-Hun, 2008; Yi & Wu,

2004). Parents and teachers in these Asian societies may be especially interested in the usefulness of behavioral regulation interventions, given the findings of the present studies.

Limitations

Although findings from the present research contribute to the understanding of cross-cultural measurement of behavioral regulation, there were a number of limitations. First, data collection in the overall longitudinal study in Taiwan was conducted at two time points: the end of the preschool year, and the beginning of the following prekindergarten year. The timing of these collection points was scheduled due to practical constraints, and future research may schedule them at the beginning and end of a school year in order to be able to analyze the effects of being in a classroom for one academic year. In terms of the sample of children, all of the Taiwanese, South Korean, and Chinese children were from urban locations which may have limited the generalizability of the results within each society. In addition, the present studies were limited in their ability to investigate the variability of behavioral regulation skills because only one direct behavioral regulation task was used. Future studies employing multiple direct measures that differ in the types of responses required (e.g. motor, verbal) would further clarify whether the nature of behavioral regulation skills in each society differed based on the type of task used. Understanding cross-cultural similarities and differences in behavioral regulation would also be clearer if children of the same ages were assessed in each society. Although in Study 2 there was overlap in the ages of the children across samples, not all age groups were equally represented in each sample, and were therefore difficult to compare. However, it is noteworthy that such consistent relations were found across

child ages and societies. Finally, although the present studies began to address the ages that the HTKS was best suited for in each society, assessing a broader age range of children in each society would help to clarify the developmentally appropriate age ranges for the task.

Another limitation was that although the results from the present studies provided information about the link between behavioral regulation and academic achievement, both studies were cross-sectional. Future studies analyzing data from more than one time point would address the relation between the development of behavioral regulation and growth in academic skills. In addition, some of the variables in some of the samples (including teacher-rated behavioral regulation in the U.S. and academic outcomes in South Korea) had a relatively large amount of missing data, and results using these variables had to be interpreted with caution. Finally, future research should address the inconsistency in relations between the HTT and HTKS and the teacher-rated CBRS measure across societies. The present studies provided preliminary information about this relation in four societies, but future studies using more than one teacher-rated measure in each society would garner further insight into the differences between these types of assessments. Specifically, teacher-rated measures that do not include likert scales may prove to be more reliable and thus more similar to direct measures of behavioral regulation. Notwithstanding these limitations, findings from the present study can be used to provide a starting point for future longitudinal research with broader samples and more behavioral regulation measures.

Future Directions

Findings from the present dissertation suggest the need for multiple lines of future research. First, the presence of mediators/moderators that influence the relations between behavioral regulation and academic achievement are not fully understood. For example, it is possible that this relation is partially a function of the extent to which children's behavioral regulation influences the amount of direct instruction they receive from a teacher which would then influence their academic outcomes. Further, it is possible that the relation between behavioral regulation and academic achievement is different for children depending on their gender, age, socioeconomic status, or culture. Future research considering these intervening variables would help to clarify how parents and teachers can help support children's academic success. Second, in terms of further investigation into the use of the HTT and HTKS, their significant relation to early academic achievement suggests that these tasks may show potential for being developed as screening tools. These tools would be particularly useful in Taiwan, South Korea, and China where there is a strong cultural emphasis on academic achievement, but few assessments are available to identify children that may need additional academic support (Tsai et al., 2006). A screening tool would also be valuable in the U.S. where children are struggling academically compared to their counterparts in other societies, and evidence suggests that strong behavioral regulation may be one key to school success. If the HTT and HTKS are developed into screening tools, it will also be necessary to determine whether teachers administering the tasks are as reliable as the data collectors in the present dissertation studies.

Third, these results suggest that interventions that support children's behavioral regulation development may be useful. A few such interventions are beginning to be piloted in U.S. classrooms, but more research is needed to determine what elements are most important for improving behavioral regulation skills, and whether these interventions may be useful in Taiwan, South Korea, and China (Diamond et al., 2007; Tominey & McClelland, 2008). One way to determine what culturally-appropriate elements may be needed for interventions in each society is to examine the relations between teaching and parenting behaviors and child behavioral regulation in each society. This line of research is the third goal of the longitudinal study from which this dissertation research is based. Although this goal was beyond the scope of the present dissertation, it is an important next step for understanding how to help children develop behavioral regulation skills. It would be particularly interesting to investigate this question in high-achieving cultures such as Taiwan, South Korea, and China because specific teaching and parenting behaviors may influence early academic success via child behavioral regulation. Since interventions are most successful when they include elements in the home and the school, researching the role of parents and teachers in supporting behavioral regulation may clarify how adults in both settings can work together to encourage children's academic success (Cox, 2005).

Finally, expanding the use of the HTT and HTKS outside of the U.S. offers new opportunities for cross-cultural comparisons and as well as understanding of child development within Taiwan, South Korea, and China. These societies, however, are all within Asia and limit our understanding of behavioral regulation to this region. Future

research that examines the use of these tasks in additional high-achieving societies, such as Finland, would further clarify the role of behavioral regulation in school success (Baldi et al., 2007).

Conclusion

The majority of previous research on behavioral regulation has either been conducted in the U.S. or used measures that may not include the integration of attention, working memory, and inhibitory control. The present studies extend this work by examining a direct measure of behavioral regulation that requires children to integrate executive function skills in four different societies. Specifically, these studies lay the foundation for future cross-cultural research in behavioral regulation by providing preliminary evidence of the reliability, validity, and practical utility of the HTT and HTKS in a number of countries. Findings suggest that these tasks were useful for young children in the U.S., Taiwan, South Korea, and China and scores significantly related to children's academic success in the four societies. These findings have implications for the use of the task in future research and as possible screening tools in early childhood classrooms. Moreover, the present studies provide evidence that supporting early behavioral regulation in all four societies may help children be more successful in school.

BIBLIOGRAPHY

- Acock, A. C. (2005). Designing a research study (pp. 2-22). Corvallis, OR: Oregon State University.
- Acock, A. C. (2007). 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(5/6), 364-373.
- Ahadi, S. A., Rothbart, M. K., & Ye, R. (1993). Children's temperament in the US and China: similarities and differences. *European Journal of Personality*, 7, 359-377.
- Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, 5(2), 115-129.
- Baldi, S., Jin, Y., Skemer, M., Green, P. J., & Herget, D. (2007). *Highlights from the PISA 2006: Performance of U.S. 15-year-old students in sciences and mathematics literacy in an international context*. National Center for Education Statistics.

Bao, L. (2004). China. In J. J. Slater (Ed.), *Teen life in Asia* (pp. 19-33). Westport, CT: Greenwood Press.

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., et al. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*(3), 299-313.

Bell, J. A., & Livesey, D. J. (1985). Cue significance and response regulation in 3- to 6-year-old children's learning of multiple choice discrimination tasks. *Developmental Psychobiology, 18*, 229-245.

Berk, L. E. (2008). *Infants and children: Prenatal through middle childhood* (6th ed.). Boston: Pearson.

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. (2006). 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, 29*(2), 109-125.

- 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*(2), 647-663.
- Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28*(2), 561-571.
- Brock, L. L., Rimm-Kaufman, S. E., & Nathanson, L. (2008). The contributions of 'Hot' and 'Cool' executive function to learning-related behaviors, engagement, and academic achievement in kindergarten. *Manuscript in review*.
- 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.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595-616.

- Choi, N.-Y., & Yi, S.-H. (2007). The effects of alphabet knowledge on Korean kindergarteners' reading of Hangeul words. *Journal of Korean Home Management Association, 25*(3), 151-168.
- Chow, B., McBride-Chang, C., Cheung, H., & Chow, C. (2008). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. *Developmental Psychology, 44*(1), 233-244.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q. M., Glasney, S., & Morrison, F. J. (2008). Individualized instruction and children's literacy, behavioral and social self-regulation in first grade classrooms. *Manuscript in review*.
- Cox, D. D. (2005). Evidence-based interventions using home-school collaboration. *School Psychology Quarterly, 20*(4), 473-497.
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*, 1387-1388.

- Dixon Jr., W. E., & Salley, B. J. (2007). "Shhh! We're tryin' to concentrate": Attention and environmental distracters in novel word learning. *The Journal of Genetic Psychology, 167*(4), 393-414.
- Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "Executive skills" training. *Developmental Psychobiology, 36*(2), 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*(6), 1428-1446.
- Dwyer, E. (2004). South Korea. In J. J. Slater (Ed.), *Teen life in Asia* (pp. 205-222). Westport, CT: Greenwood Press.
- Espy, K. A., & Bull, R. (2005). Inhibitory processes in young children and individual variation in short-term memory. *Developmental Neuropsychology, 28*(2), 669-688.
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Stern, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology, 26*(1), 465-486.

- Fahie, C. M., & Symons, D. K. (2003). Executive functioning and theory of mind in children clinically referred for attention and behavior problems. *Journal of Applied Developmental Psychology, 24*(1), 51-74.
- Farver, J. M., Kim, Y. K., & Lee, Y. (1995). Cultural difference in Korean- and Anglo-American preschoolers' social interaction and play behaviors. *Child Development 66*, 1088-1099.
- Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology, 70*, 177-194.
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1-16.
- Ginsburg, H. P., & Baroody, A. J. (1990). *Test of early mathematics ability* (2nd ed.). Austin, TX: Pro-Ed.
- Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective likert scales? The reference-group effect. *Journal of Personality and Social Psychology 82*(6), 903-918.

- 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*(1), 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*(2), 151-174.
- Hsieh, M.-F. (2004). 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 5*(3), 309-329.
- Hsu, H.-H. (1998). *The concept of early learning mathematics seminar collection*. Paper presented at the Life, games, mathematics in young children, Taipei, Taiwan.
- Hsu, H.-H. (2000). *Child mathematical aptitude test TEMA-2*. Taipei, Taiwan: Psychological Publishing Company Ltd.

- Hughes, D., Seidman, E., & Williams, N. (1993). Cultural phenomena and their research enterprise: Toward a culturally anchored methodology. *American Journal of Community Psychology, 21*(6), 687-703.
- Huntsinger, C. S., Jose, P. E., Liaw, F.-R., & Ching, W.-D. (1997). Cultural differences in early mathematics learning: A comparison of Euro-American, Chinese-American, and Taiwan-Chinese families. *International Journal of Behavioral Development, 21*(2), 371-388.
- Jose, P. E., Huntsinger, C. S., Huntsinger, P. R., & Liaw, F.-R. (2000). Parental values and practices relevant to young children's social development in Taiwan and the United States. *Journal of Cross-Cultural Psychology, 31*(6), 677-702.
- Kim, J., Lee, Y.-s., Suen, H., & Lee, G. S. (2003). A Delphi study of young children's readiness in Korea: Challenges and implications for early childhood schooling. *Educational Research and Evaluation, 9*(4), 345-355.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

- Kotchick, B. A., & Forehand, R. (2002). Putting parenting in perspective: A discussion of the contextual factors that shape parenting practices. *Journal of Child and Family Studies, 11*(3), 255-269.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231-251.
- Lan, X., Wanless, S. B., McClelland, M. M., Son, S.-H., Ponitz, C. C., Morrison, F. J., et al. (2008). Gender differences in behavioral regulation in four cultures. *Manuscript in preparation.*
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.
- Lee, I.-C. (2004). A comparison of the play performance of boys with autism and that of boys without disabilities in Taiwan. *Dissertation Abstracts International: Section B: The Sciences and Engineering 64*(12-B), 6048.
- Lee, K., Xu, F., Fu, G., Cameron, C. A., & Chen, S. (2001). Taiwan and Mainland Chinese and Canadian children's categorization and evaluation of lie- and truth-

telling: A modesty effect. *British Journal of Developmental Psychology*, *19*, 525-542.

Lin, Y.-W., & Tsai, M.-L. (1996). Culture and the kindergarten curriculum in Taiwan. *Early Childhood Development and Care*, *123*, 157-165.

Livesey, D. J., & Morgan, G. A. (1991). The development of response inhibition in 4- and 5-year-old children. *Australian Journal of Psychology*, *43*(3), 133-137.

Loo, S. K., & Rapport, M. D. (1998). Ethnic variations in children's problem behaviors: A cross-sectional, developmental study of Hawaii school children. *Journal of Child Psychology and Psychiatry*, *39*(4), 567-575.

Lu, L., & Kao, S.-F. (2002). Traditional and modern characteristics across the generations: Similarities and discrepancies. *Journal of Social Psychology*, *142*(1), 45-59.

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

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., et al. (2002). Validity of the behavior rating inventory of executive function

in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, 17(7), 643-663.

Mahone, E. M., & Hoffman, J. C. (2005). Behavior ratings of executive function among preschoolers with ADHD [Abstract]. *International Neuropsychological Society*, 11(S1), 129.

Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekindergarteners; relationships and behaviors. *Journal of Psychoeducational Assessment*, 24(4), 367-380.

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2008). Early gender differences in self-regulation: A sixth-sense for academic achievement? *Manuscript in preparation*.

McClelland, M. M. (2002). The emergence of work-related skills in preschool children: Loyola University Chicago.

McClelland, M. M. (2008). Inter-rater 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. B., & Murray, A. (2007). Executive function, self-regulation, and social-emotional competence: Links to school readiness. In O. N. Saracho & B. Spodek (Eds.), *Contemporary Perspectives on Research in Social Learning in Early Childhood Education*.

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*(3), 307-329.

McClelland, M. M., & Piccinin, A. M. (2008). Relations between preschool attention and sociability and later achievement and social outcomes. *Manuscript in preparation*.

McClelland, M. M., & Wanless, S. B. (2008). Touching your toes in four cultures: Developing a new measure of behavioral regulation for young children,

International Society for the Study of Behavioural Development. Wurzburg, Germany.

McMullen, M. B., Elicker, J., Wang, J., Erdiller, Z., Lee, S.-M., Lin, C.-H., et al. (2005).

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*, 20(4), 451-464.

Meng, X. L. (1995). Multiple imputation inferences with uncongenial sources of input.

Statistical Sciences, 10, 538-573.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003*

International Mathematics Report. Chestnut Hill, MA: TIMSS & PIRLS

International Study Center.

Ogbu, J. U. (1981). Origins of human competence: A culture-ecological perspective.

Child Development 52, 413-429.

Oh, S., & Lewis, C. (2008). Korean preschoolers' advanced inhibitory control and its

relation to other executive skills and mental state understanding. *Child*

Development 79(1), 80-99.

Organization for Economic Cooperation & Development. (2001). *Starting strong: Early childhood education and care*. Paris: OECD.

Orlick, T., Zhou, Q. Y., & Partington, J. (1990). Co-operation and conflict within Chinese and Canadian kindergarten settings. *Canadian kindergarten settings, Canadian Journal of Behavioral Science*(22), 20-25.

Pagani, L. S., Vitaro, F., Tremblay, R. E., McDuff, P., Japel, C., & Larose, S. (2008). When predictors fail: The case of unexpected pathways toward high school dropout. *Journal of Social Issues*, 64(1), 175-194.

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

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

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black-white test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The black-white test score gap* (pp. 229-272). Washington, D.C.: Brookings Institution Press.

- Pickering, S. J., & Gathercole, S. E. (2004). Distinctive working memory profiles in children with special educational needs. *Educational Psychology, 24*(3), 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.
- Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2008). A structured observation of behavioral regulation and its relation with kindergarten outcomes. *Manuscript in review.*
- Ratner, C., & Hui, L. (2003). Theoretical and methodological problems in cross-cultural psychology. *Journal of the Theory of Social Behavior, 33*(1), 67-94.
- 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*(1), 21-38.

- 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*(4), 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*(2), 147-166.
- 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. A. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 338-357). Malden, MA: Blackwell Publishing.
- Rowe, D., Jacobsen, K., & Van den Oord, E. (1999). Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. *Child Development 70*(5), 1151-1162.

- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*, 473-489.
- Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573-594.
- Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L., J., & Lee, K. (2006). The development of executive functioning and theory of mind. *Psychological Sciences, 17*(1), 74-81.
- Sang-Hun, C. (2008, August 13). A taste of failure fuels an appetite for success at South Korea's cram schools. *New York Times*, p. A6.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.
- Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy, 2*(4), 419-436.

- Senn, T. E., Espy, K. A., & Kaufmann, P. M. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology, 26*(1), 445-464.
- Sharma, Y. (1997, 1997, February 16). Tough fight for places among tiny elite. *Times Educational Supplement*.
- Sheng, L., McGregor, K. K., & Marian, V. (2006). Lexical-Semantic organization in bilingual children: Evidence from a repeated word association task. *Journal of speech, language, and hearing association, 49*, 572-587.
- Shweder, R. A., Goodnow, J., Hatano, G., LeVine, R. A., Markus, H., & Miller, P. (1998). The cultural psychology of development: One mind, many mentalities. In R. M. Lerner (Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 865-937). New York: Wiley.
- 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*(2), 173-187.

- Son, S.-H., & Meisels, S. J. (2006). The relationship of young children's motor skills to later reading and math achievement. *Merrill-Palmer Quarterly*, 52(4), 755-778.
- Son, S.-H., Ponitz, C. C., Lan, X., Wanless, S. B., McClelland, M. M., Morrison, F. J., et al. (2008). The development of behavioral regulation in four cultures. *Manuscript in preparation*.
- Sorensen, C. W. (1994). Success and education in South Korea. *Comparative Education Review*, 38(1), 10-35.
- Stevenson, H. W., Chen, C., & Lee, S.-Y. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science*, 259(5091), 53-58.
- Stevenson, H. W., Lee, S., Chen, C., Stigler, J. W., Hsu, C. C., & Kitamura, S. (1990). Contexts of achievement: A study of American, Chinese, and Japanese children. *Monographs of the Society for Research in Child Development*, 55(1-2, Serial No. 221)).
- Sue, S., & Okazaki, S. (1990). Asian-American education achievements: A phenomenon in search of an explanation. *American Psychologist* 45, 913-920.

- Tamis-LeMonda, C. S., Wang, S., Koutsouvanou, E., & Albright, M. (2002).
Childrearing values in Greece, Taiwan, and the United States. *Parenting: Science
and Practice, 2*(3), 185-208.
- Tobin, J. (2007). An ethnographic perspective on quality in early childhood education. In
J. Zhu (Ed.), *International perspectives on early childhood education* (pp. 131-
143). Shanghai: East China Normal University Press.
- Tominey, S., & McClelland, M. M. (2008). Red light, purple light: Initial findings from
an intervention to improve self-regulation over the pre-kindergarten year.
Manuscript in preparation.
- Tsai, H.-L. A., McClelland, M. M., Pratt, C., & Squires, J. (2006). Adaptation of the 36-
Month Ages and Stages Questionnaire in Taiwan: Results from a preliminary
study. *Journal of Early Intervention, 28*(3), 213-225.
- van der Schoot, M., Licht, R., Horsley, T. M., Aarts, L. T., van Koert, B., & Sergeant, J.
A. (2004). Inhibitory control during sentence reading in dyslexic children. *Child
Neuropsychology, 10*(3), 173-188.

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

Wall, S. M., & Paradise, L. V. (1981). A comparison of parent and teacher reports of selected adaptive behaviors of children. *Journal of School Psychology, 19*(1), 73-77.

Wanless, S. B., McClelland, M. M., Acock, A. C., Chen, F.-M., & Chen, J.-L. (2008). Behavioral regulation and early academic achievement in Taiwan. *Manuscript in preparation*.

Wanless, S. B., McClelland, M. M., Son, S.-H., Lan, X., Ponitz, C. C., Morrison, F. J., et al. (2008). Measuring behavioral regulation in four cultures. *Manuscript in preparation*.

What the numbers say. (1998). *Curriculum Review, 37*(6), 2.

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

Woodcock, R. W., & Muñoz-Sandoval, A. F. (1996). *Batería Woodcock-Muñoz-R:*

Pruebas de aprovechamiento. Itasca, IL: Riverside.

Wu, C.-C. (2007). Reading beliefs and strategies of Taiwanese mothers with preschoolers in relation of the children's emergent literacy. *Dissertation Abstracts International*

Section A: Humanities and Social Sciences, 68(5-A), 1808.

Yi, C.-C., & Wu, C.-I. (2004). Taiwan. In J. J. Slater (Ed.), *Teen life in Asia* (pp. 223-241). Westport, CT: Greenwood Press.

Zelazo, P. D., Reznick, J. S., & Piñon, D. E. (1995). Response control and the execution of verbal rules. *Developmental Psychology* 31(3), 508-517.

Zero to Three. (n.d.). *Zero to Three: Response to prekindergarteners left behind*.

Retrieved March 24, 2006, from http://www.zerotothree.org/left_behind.html

Zhu, J., & Zhang, J. (2008). Contemporary trends and developments in early childhood education in China. *Early Years*, 28(2), 173-182.