THE RELATION OF KINDERGARTEN ENTRY SKILLS TO EARLY LITERACY AND MATHEMATICS ACHIEVEMENT

by

PHILLIP SHAWN IRVIN

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2015

DISSERTATION APPROVAL PAGE

Student: Phillip Shawn Irvin

Title: The Relation of Kindergarten Entry Skills to Early Literacy and Mathematics Achievement

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Educational Methodology, Policy, and Leadership by:

Gerald Tindal	Chairperson
Charles R. Martinez	Core Member
Keith Zvoch	Core Member
Jane Squires	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2015.

DISSERTATION ABSTRACT

Phillip Shawn Irvin

Doctor of Philosophy

Department of Educational Methodology, Policy, and Leadership

June 2015

Title: The Relation of Kindergarten Entry Skills to Early Literacy and Mathematics Achievement

Learning-related behavioral and academic skills upon kindergarten entry, sometimes referred to as *kindergarten readiness*, is a construct of growing importance in education, having implications for early learning and eventual achievement. Much of the research on entry skills has been limited to initial status only with inferences drawn about preparedness for school. In this study, I examine the relation among kindergarten entry skills in literacy and mathematics as well as outcomes measured at the end of the kindergarten school year.

Two extant datasets were used—learning-related behavioral ratings and academic proficiency skills scores from a fall administration of a statewide kindergarten entry assessment and interim-formative assessment data collected for a subsample of students in the spring of the same academic year. The assessments were analyzed for their factor structures (using both exploratory and confirmatory factor analyses) as well as a hypothesized structural model. Factor analyses tested and confirmed the underlying structure and relations among items and measures included in the state entry assessment. Follow-up structural modeling confirmed the measurement model and concurrently estimated the effects of entry skills on emergent literacy and math skills measured in the spring, while accounting for student-level demographic characteristics.

Results indicated that the state's entry assessment measured three distinct skillsets: self-regulation and social-interpersonal learning-related behaviors, and academic proficiency. Importantly, kindergarten entry skills explained a large proportion of variance in spring emergent literacy achievement, beyond that of learning-related behavioral skills and student demographic characteristics. In contrast, these entry skills explained far less variance in spring math achievement. These findings are interpreted in the context of existing theory and recent empirical research.

CURRICULUM VITAE

NAME OF AUTHOR: Phillip Shawn Irvin

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
Antioch University McGregor, Yellow Springs, OH
Miami University, Oxford, OH

DEGREES AWARDED:

Doctor of Philosophy, Educational Methodology, Policy, and Leadership, 2015, University of Oregon
Master of Science, Educational Leadership, 2007, Antioch University McGregor
Master of Science, Geology and Geochemistry, 2002, Miami University
Bachelor of Science, Geology, 1997, Miami University

AREAS OF SPECIAL INTEREST:

Quantitative Research Methodology
Measurement

PROFESSIONAL EXPERIENCE:

Research Assistant and Graduate Teaching Fellow, Behavioral Research and Teaching, University of Oregon, Eugene, Oregon, 2011

Science Teacher, WPAFB, Dayton, OH, 2006-2010

PUBLICATIONS:

Anderson, D., Irvin, P. S., Alonzo, J., & Tindal, G. A. (2014). Gauging Item Alignment Through Online Systems While Controlling for Rater Effects. *Educational Measurement: Issues and Practice*. doi: 10.1111/emip.12038

Scalise, K., Timms, M., Moorjani, A., Clark, L., Holtermann, K., & Irvin, P. S. (2011). Student learning in science simulations: Design features that promote learning gains. *Journal of Research in Science Teaching*, 48(9), 1050-1078. doi: 10.1002/tea.20437

ACKNOWLEDGMENTS

I express my genuine appreciation to Dr. Gerald Tindal, who has far exceeded all expectations as my advisor and mentor. Working with Dr. Tindal since 2011, it is a true honor and privilege to associate my name with his legacy. I would also like to thank Dr. Julie Alonzo, who saw in me a promising researcher and a would-be asset to Behavioral Research and Teaching (BRT). Julie's friendship and mentorship has meant the world. I would like to thank my committee who helped me think critically about my research. I would also like to thank the Oregon Department of Education, and specifically Dr. Steve Slater, who provided Oregon Kindergarten Assessment data. As the state endeavors to implement and improve the entry assessment, I hope my work is way helpful. I would like to thank my friends and colleagues at BRT, whose quality and support constantly challenge me personally and professionally. I would also like to thank Brock Rowley, who while working on similar research was a lifeline I could bounce ideas off of and commiserate with late into the night. I would like to thank my friend Nicolas Harr, who over the years has become a steady source of encouragement. Lastly, I thank my family. My grandparents, who have long passed, are with me everyday. I work hard to live up to the version of me that they always saw and loved. My parents Ann and Phil and my brother Dave are constant and unwavering sources of support and inspiration. Their influence on my life, and willingness to always be present are integral to any personal or professional success I have or ever will achieve. Jasper and Indy were my side each and every time I worked late on coursework or my dissertation—dogs are the best. Finally, my wife Beth is my rock. Her love and patience (and even the nagging to finish) throughout this journey and our life together is no small miracle. Simply put, I love her forever.

For my grandparents, parents, brother and Beth who continue to teach me the value of learning and the importance of working with others—and my unborn daughter, with whom I'll strive to teach the same.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

With the importance of reading and math achievement to student success well documented and measured for accountability purposes starting in third grade, federal and state governments, along with corporate and non-profit interests, are investing anew in early learning programming (*The White House Summit on Early Childhood Education*, 2014). The investment by these groups includes millions toward universal preschool development (e.g., Preschool for all Initiative) and nationwide full-day kindergarten in an attempt to provide all children a jumpstart on learning and to identify and address early achievement and opportunity gaps (The White House & Office of the Press Secretary, 2013; U.S. Department of Education, 2014). States are responding by formally aligning early learning and K-3 systems to improve reading and math achievement, including developing comprehensive assessment and data systems that link preschool with the early primary years, and early primary years with later public schooling (*The White House Summit on Early Childhood Education*, 2014).

The development and improvement of kindergarten entry assessments—formal measurement of students' skills upon entering kindergarten, is a part of this investment and alignment process. In many cases, states receive support from the federal government through Race to the Top, Early Learning Challenge or enhanced assessment grants to develop such assessments (McGuinn, 2012; U.S. Department of Education, 2013). In 2010 just seven states had mandated statewide kindergarten entry assessments; this number grew to 25 in 2011, with 43 states at some point in the process of developing entry assessments in 2013 (Connors-Tadros, 2014). Over time, evidence of connection (concordant or experimental) between measured early learning, kindergarten entry, and early primary achievement outcomes may demonstrate evidence of policy and investment success—contingent on improvement in reading and mathematics achievement.

In the fall 2013, the Oregon Department of Education (ODE) conducted its first statewide field test of the new Oregon Kindergarten Assessment (OKA; Oregon Department of Education, 2013a), with four intended purposes cited: (a) to provide baseline behavioral and achievement skills data to stakeholders (i.e., educators, parents, policy makers), (b) to provide information to guide instructional decision-making, (c) to identify achievement gaps early among children in various demographic groups (e.g., geographical, cultural, gender, racial, and socioeconomic), and (d) to provide a single assessment tool for the state (Oregon Department of Education, 2013c).

The OKA battery provides the levels of academic skill and learning-related behaviors present in the first few weeks of kindergarten. Such early assessment may be informative as a gauge of entering skills, though the capacity of the OKA may also be restricted to practically inform decision-making in the intended areas and predict later achievement, in part, due to hypersensitivity and floor effects that may be observed when measuring such early skills (Catts, Petscher, Schatschneider, Bridges, & Mendoza, 2009). With the state's intended purposes in mind, the relation between and concordant validity of OKA entry skills, shown to be interrelated in an earlier analysis of pilot data (Tindal, Irvin, & Nese, Manuscript submitted for publication), need further investigated to establish valid and parsimonious factor structures and to examine their relation to later achievement in a broader sample. The former is important for accurately characterizing the state assessment and results, while the latter seems, albeit not explicitly stated, an intention behind the design and implementation of the OKA as an indicator of school preparedness and component of stakeholder (e.g., parents, policy makers, state board officials, district/school leaders) decision-making. In this study I investigate the reach of the OKA by examining the relation between students' learning-related behavior and academic achievement levels upon entry and their relation to important early/emergent literacy and mathematics indicators measured at the end of kindergarten. Examining the OKA battery and its predictive-concordant capacity may provide greater depth of understanding regarding the complex interplay of entry skills and their relation to students' achievement over the initial year of public schooling. Thus, in this study I examine a timely and sometimes contentious topic of early assessment of kindergarten school children within the context of a state initiative, framed by issues of measurement and validity.

Theoretical Framework: Learning Through Acquisition and Participation

In an analysis of 2012 OKA pilot data, Tindal et al. (Manuscript submitted for publication) modeled the relation between the learning-related behavioral and academic entry skills based on the theoretical views of Sfard (1998). In an attempt to unify potentially disparate theoretical views, Sfard argued that teaching and learning are fundamentally grounded in two distinct, though not inherently competing, conceptual metaphors: an *acquisition metaphor* (AM) and a *participation metaphor* (PM). At its core, the AM involves acquiring and developing frameworks (e.g., knowledge, concepts, meaning, sense), making them one's own through internal processes (e.g., reception, construction, internalization), and then using such frameworks across circumstances (e.g., transmission, translation, application). In short, the AM seems almost obvious when learning processes are considered across contexts, with phrases such as "knowledge acquisition" and "concept development" invoking perceptions of the "human mind as a container to be filled with certain materials and about the learner as becoming an owner of these materials" (Sfard, 1998, p. 5). Alternatively, the PM views the learner as an active participant in socialized learning contexts. Themes like "situatedness, contextuality, cultural embeddedness, and social mediation" (p. 7) frame the learner as being involved in activities as a part of a community of learners in which language and behavior are guided by certain norms and knowledge is co-built and utilized.

The two metaphors are distinguishable in many ways. The AM focuses on individual enrichment and development, while the PM focuses on growing bonds and building community. The AM situates the learner as being inward focused, while the PM positions the learner as looking and connecting outward. The AM is grounded in self-identification and possession, while the PM is based on group-identification and sharing. These metaphors connect what we know to what is discovered and created—working together to define learning in terms of our experiences as individuals and in surrounding groups—working separately *and* together.

Though theoreticians, researchers and educators may adhere to one or the other conceptual metaphor in practice and distinguish them as described, doing so is not necessary and may restrain learning experiences. As Sfard (1998) aptly stated, "the individual/social dichotomy does not imply a controversy as to the definition of learning, but rather rests on differing visions of the mechanism of learning" (p. 7). In the end, Sfard argued that learning and teaching are based on principles of *both* acquisition and participation.

Tindal et al. (Manuscript submitted for publication) argued that both metaphors are present and interrelated in the OKA, and thus, in the underlying construct of entering behavioral and academic skills of kindergartners addressed by the OKA battery. The researchers documented preliminary evidence supporting the underlying theoretical model of the OKA that empirically relates self-regulation and social-interpersonal participation behaviors with early academic acquisition skills. I seek to build on their argument and findings. For this study (and within the context of the OKA), kindergartners' entering academic skill in early literacy and numeracy, as well as their end-of-year early/emergent literacy and mathematics performance, represent the AM; kindergartners' entering learning-related behaviors indicate the PM. I extend the work of Tindal and colleagues by first analyzing and confirming the factor structure of the OKA using a statewide Oregon kindergarten sample. I then investigate the relation between entry academic and learning-related behavioral skills and important early/emergent literacy and mathematics skills measured at the end of the kindergarten school year using a portion of the statewide sample.

The Acquisition Metaphor

As is operationalized by the OKA, acquisition studies with kindergarten students typically focus on measuring achievement skills in the domains of early literacy (i.e., alphabetic and phonemic awareness) and mathematics (i.e., numeracy). Particular attention has been paid to developing measures that are technically adequate for identifying (screening) learners at risk of not meeting grade-level expectations and that are also sensitive to measuring both status (level) and change in student performance over time (growth) to aid in instructional decision-making (McConnell, McEvoy, & Priest, 2002). Researchers have also focused on ensuring such measures have valid and parsimonious factor structures in order to appropriately characterize assessments and associated results (Justice, Invernizzi, Geller, Sullivan, & Welsch, 2005).

Within the context of identifying risk and providing instructional information, one should consider hypersensitivity and floor effects when measuring these early developing skills. Hypersensitive measures are those with limited practical scope over the school year (e.g., letter naming fluency), and thus, might have limited utility in predicting higher-order skill development within and across grades (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Paris, 2005). Floor effects are observed when students exhibit very low levels of performance on a particular assessment. For example, students might score very low on a test of letter sounds fluency at the beginning of kindergarten because they do not yet have the skill to sound letters of the alphabet. As is the case with those that are hypersensitive, measures that exhibit floor effects may have limited predictive validity as early screening assessments (Catts et al., 2009). If as the state indicates results from the OKA are intended to guide decision-making and elucidate gaps between demographic groups, the relation of measured skills to one another and to other important skills needs investigated.

Measuring early literacy skills and their relations. In his seminal review of curriculum-based measurement (CBM), Tindal (2013) argued that three key events spurred research around early literacy achievement and growth. First, the National Reading Panel, comprised of expert researchers and educators from across the U.S., defined five essential components of reading: phonemic awareness, phonics, fluency, vocabulary and comprehension (National Institutes of Child Health and Human

Development, 2000). Second, *No Child Left Behind* ushered in the age of accountability testing that required students be proficient readers by the end third grade ("No Child Left Behind (NCLB) Act of 2001," 2002)—prompting measurement of early literacy skills as a means to predict the likelihood of later proficiency (U.S. Department of Education, 2008). Third, federal formula grants, funded initially in 2002 with continued funding though 2008 under the *Reading First* initiative, further focused attention on nationwide early literacy improvement (U.S. Department of Education, 2002). Following these events Fuchs, Fuchs, and Compton (2004) called for the field to "examine the tenability of reading tasks that address an earlier phase of reading" (p. 7); this is a call that has been heeded by researchers over time, and a focus reflected by the Obama Administration's current investment in early education initiatives across the country (U.S. Department of Education, 2013).

Researchers have long documented the development and importance of key alphabetic and phonological skills in terms of their interrelatedness and their proximal and distal effects on measures of status and growth in various reading-related skills (Speece, Ritchey, Cooper, Roth, & Schatschneider, 2004; Wagner, Torgesen, & Rashotte, 1994). Linklater, O'Connor, and Palardy (2009) found significant change on two kindergarten measures of early literacy skills: initial sound fluency (ISF; a measure of students' ability to sound out letters of the English alphabet) and phoneme segmentation fluency (PSF; a measure of students' ability to identify/sound phonemes in grade-level word lists). Using measures with documented reliability from the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS; Good & Kaminski, 2003), Linklater and colleagues found that the ability to sound letters and segment phonemes predicted unique variance in emergent reading skills at the end of kindergarten, specifically, nonsense word fluency (NWF), word reading fluency (WRF), and comprehension. Cummings, Kaminski, Good, and O'Neal (2011) assessed both pre-kindergarten and kindergarten students over three seasonal time points, and a subset of the sample multiple times in between the seasonal benchmarking, using alternate forms of another early literacy DIBELS measure, first sound fluency (FSF; a measure of early phonemic awareness skills). The researchers found the FSF measure reliable across all time points and sensitive to changes in early phonemic awareness skills for preschool and kindergarten students. Importantly, for the kindergarten portion of the sample, Cummings and colleagues found FSF to be moderately-highly correlated with PSF and a widely used criterion measure of early literacy skills, the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999).

Other research teams have also examined the interrelation of early and emergent literacy skills, also examining their importance to later reading skills in young children. Using growth curve analysis, Speece et al. (2004) controlled for background factors (i.e., family literacy and nonverbal IQ), and incorporated several skill indicators relevant to the transition from early to later reading to predict year-end third grade reading performance and the rate of growth for three measures of later reading skills: letter word identification, word attack and reading passage comprehension. The researchers measured general oral language, phonological awareness, emergent reading, and spelling, early in the kindergarten year and used them as predictors of intercept and growth. Amongst several findings that substantiate earlier concurrent and predictive validity studies of early literacy skills, perhaps the most significant was that kindergarten phonological awareness was a unique predictor of third grade word-level knowledge-skills, letter word identification, and word attack even in the presence of other important indicators. From a predictive validity viewpoint, it appears later word reading performance has been established for most students based on their ability to phonologically process letters and sounds.

Ritchey and Speece (2006) used a variety of measures (e.g., DIBELS, CTOPP, Woodcock Reading Mastery Test-Revised) to examine the complex and interrelated nature of early alphabetic and phonemic fluency skills (i.e., letter names fluency (LNF), LSF, and PSF), including their capacity to predict reading skills later in kindergarten. Two key findings related to early literacy skill development emerged from the study. First, students exhibited significant growth on all three alphabetic and phonemic fluency measures—these skills are thus present early and develop over kindergarten. Second, LSF served as a connective mechanism between early (i.e., LNF and PSF) and higher-order reading skills (i.e., word reading and spelling). Ritchey and Speece argued that researchers should focus attention on these early skills, in particular LSF, as means of supporting emergent reading skill development. Soon after, Ritchey (2008) followed up her work with Speece, bolstering their findings on the importance of LSF as an early literacy skill critical to reading development in early elementary school. LSF was measured over the latter half of kindergarten and found to significantly predict word reading and oral reading fluency (ORF) at the end of the kindergarten year.

Based on the studies highlighted here, researchers have used a variety of technically adequate measures, finding that early alphabetic (e.g., the ability to name and sound letters) and phonemic (e.g., phoneme awareness and segmenting) skills are interrelated and important to predicting emergent reading skills (e.g., word reading, vocabulary, spelling, reading fluency and comprehension). Such early and transitional literacy skills are correlated to each other and predictive of later (higher order) reading skills, and thus, are often included in kindergarten entry assessments, including the OKA (Oregon Department of Education, 2013c).

Measuring early math skills and their relations. Though comparatively less research has been published on early math skills than on early/emergent literacy and reading (Gersten et al., 2012), researchers studying math skills in young students identify aspects of counting, cardinality, numeracy, geometry, and early operations as important in early learning contexts and to later math skills development (see Foegen, Jiban, & Deno, 2007; Gersten et al., 2012). Clements, Sarama, and Lieu (2008) developed and validated the Research-based Early Maths Assessment (REMA) using a broad range of empirical findings around important early math core ideas and skills and associated learning trajectory research (see, for example, Clements, Wilson, & Sarama, 2004), and later developed and validated a shorter form more conductive to classroom use (Weiland et al., 2012). The researchers documented unidimensional measurement of a single latent trait comprised of developmental progressions of math skills across five main content areas: (a) verbal number counting, object counting, number recognition and subitising (i.e., without counting, quickly recognizing the number of objects in a small group), (b) number comparison (i.e., number sequencing, numeral recognition, number composition and decomposition, and adding and subtracting); (c) geometry, (i.e., shape identification, shape composition and decomposition, comparison and congruence, construction of shapes, and transformations); (d) measurement; and (e) patterns. Together, the researchers argued, these skills comprise the basis of early math skill development in young (pre- and kindergarten) children.

Earlier, VanDerHeyden et al. (2004) developed math measures including counting objects, selecting numbers, naming numbers, identifying shapes, counting, and visual discrimination. Based on the performance of 60 four-year-old preschool students, VanDerHeyden and colleagues found that the majority of measures were reliable across alternate forms and that they correlated moderately with the Test of Early Mathematics Ability (TEMA-2; Ginsburg & Baroody, 2005) and the Brigance Screen (Brigance, 1985), a global measure of early academic skills. In a follow-up that sampled preschool and kindergarten students over two consecutive years, VanDerHeyden, Broussard, and Cooley (2006) used their preschool early mathematics measures along with newly developed kindergarten tasks comprised of selecting numbers, counting objects, counting and visual discrimination to document evidence of screening accuracy in instructional and intervention contexts. Mathematics performance on the preschool measures correlated moderately to strongly with performance in kindergarten indicating that emergent counting and numeracy skills develop in early school contexts and are related to one another over time.

Seethaler and Fuchs (2011) examined test-retest reliability and concurrent and predictive validity of kindergarten students' initial and final performances over 14 weeks on alternate test forms of a fluency measure that assessed counting, addition and subtraction. Reliability was strong, ranging from .80 to .87, and students' initial and final math scores correlated moderately (.61 and .69) with the TEMA-3 (Ginsburg & Baroody, 2005) administered at the end of kindergarten. Additionally, Seethaler and Fuchs evaluated the instrument as a screener of academic risk, using students' initial performance as a predictor of later math difficulty. They found high sensitivity (90%), but low specificity (64%), including 24 false positives out of 87 students, and suggested a two-stage screening process to improve classification accuracy and save instructional time and resources. These findings appear to support earlier work by Gersten, Jordan, and Flojo (2005) who found that difficulties in number sense (early numeracy) were predictive of difficulties in math for kindergarteners, though not without false positives as well. VanDerHeyden (2011) placed this problem of sensitivity and specificity into practical context. She argued that while screeners may predict whether a student is at risk for math (or reading) learning difficulties within response to intervention (RTI) contexts, such predictions are problematic because "at-risk" typically is without precise definition. Thus, cut-scores used to determine who is and is not at-risk should demonstrate adequate consequential validity in the context in which they are operationalized (Gersten, Keating, & Irvin, 1995).

Lembke and Foegen (2009b) took a forward-looking approach from kindergarten and assessed over 300 kindergarten and first-grade students to evaluate the technical adequacy of four different early numeracy indicators including number identification, quantity discrimination, quantity array, and missing number. The researchers documented strong alternate form (.80 to .90) and test-retest (.80 to .88) reliability. Additionally, Lembke and Foegen found that scores from the beginning of the year for both kindergartners and first graders were significantly predictive of scores at the end of the year on the TEMA-3 (.34 to .68) and teacher ratings of math abilities (.49 to .70), with number identification and missing number being the most highly predictive.

As appears the case with early literacy skills research, researchers of early math have focused on identifying and measuring sets of skills fundamental to early math skills development. A variety of technically adequate measures gauge status, screen for risk, and monitor change over time. In large part, researchers have focused on skills involving number sense (i.e., early numeracy and connected simple operations), "a child's fluidity and flexibility with numbers, the sense of what numbers mean, and an ability to perform mental mathematics and to look at the world and make comparisons" (Gersten & Chard, 1999, pp. 19-20), which appears to broadly cover many of the early math tasks highlighted in the research synthesized here, and is also a key part of the OKA (Oregon Department of Education, 2013c).

The Participation Metaphor and Its Relation to Acquisition

In addition to students acquiring academic skills, Sfard (1998) argued that active interactive participation in the classroom (and school) community is also an integral mechanism in learning. Young students' self-regulatory and social learning-related behaviors in the classroom community, specifically developing relations to self, peers and adults, are key factors when adapting to school and learning processes, and complexly connected with academic skill acquisition (Ladd, Birch, & Buhs, 1999). Kindergarten represents the first time many students have participated in formal education/school settings. According to the U.S. Commerce Department, in 2013 fewer than half of 3- and 4-year-old children were enrolled in some form of early childhood (i.e., preschool) programming, split about evenly between publicly- and privately-sponsored programs (U.S. Census Bureau, 2013). For students who have participated in early childhood programs or Head Start, the experience of classroom learning may not be foreign, but the public kindergarten school/classroom environment represents a different rule- and norm-governed system that needs negotiated. Thus, like academic acquisition skills, relations to self, peers and adults in educational settings, today often specified as self-regulation and social-interpersonal/emotional behaviors (though other terms have and are used, e.g., work-related skills), have been studied theoretically and from measurement perspectives, including in various studies analyzing kindergarten entrance and early achievement.

In educational contexts where students interact with the self, peers and adults, a social cognitive perspective may be useful in better understanding self-regulatory behaviors and their interplay with external (influences. In a seminal work that defined the mechanisms behind self-regulation, Albert Bandura argued that almost all behavior is purposeful, and therefore, self-regulated in that individuals envision desired (internal and external) outcomes and devise behaviors to reach such outcomes (Bandura, 1991). Bandura stated that "self-regulation is a multifaceted phenomenon operating through a number of subsidiary cognitive processes including self-monitoring, standard setting, evaluative judgment, self-appraisal, and affective self-reaction" (Bandura, 1991, p. 282). For example, a student desiring to learn the letters of the alphabet might practice writing each letter using an exercise workbook at home. Behavior is thus internally regulated, though it is also an extrapsychic affair, because while it mediates internal influences it is also affected by external factors within social settings (Bandura, 1986). Extending the simple example above, the same student might also routinely practice writing each letter at home and in school because the teacher has instructed him/her to do so and because doing so draws praise from teachers and parents. In this example, the student's behavior is internally regulated while at the same time being externally influenced.

In an early study that measured learning-related behaviors in classroom settings, Cooper and Farran (1988) developed a teacher-rated scale with two independent subscales termed "work-related skills" (e.g., listening, following directions, remaining on task) and "interpersonal skills" (e.g., sharing, playing cooperatively, relating positively with peers) to frame groups of behaviors deemed critical to kindergarten success. The researchers collected observational ratings for 650 kindergarten students and found that students classified as being maladjusted to the classroom were associated with lower than typical work-related skills, while interpersonal skills appeared unrelated. Cooper and Farran argued that views of kindergarten entry (readiness) should be expanded beyond academic knowledge and social interactions to include indicators and measures of work-related skills.

Ladd et al. (1999) also looked at relative behavioral and environmental risk factors prior to and just after kindergarten entry. Among several propositions tested, Ladd and colleagues framed participatory behaviors as being predicted by behavioral and environmental factors within the first three months of kindergarten, and as an antecedent to achievement. The researchers measured kindergartners' participation using the Cooperative and Independent Participation subscales of the Teacher Rating Scale of School Adjustment (Birch & Ladd, 1997), which evaluated the extent to which students accepted authority and worked well with others and displayed self-directed learning behaviors. They measured early achievement using the Visual and Quantitative composites of the Metropolitan (School) Readiness Test (Nurss & McGauvran, 1986). Ladd and colleagues found that, on average, higher levels of participatory behaviors exhibited early on in kindergarten predicted higher levels of early literacy and numeracy/operations achievement midway through the kindergarten school year. In line with Sfard's (1998) theorizing, Ladd et al. (1999) argued that such early learning-related behaviors (self-regulatory and social in nature) demonstrated "an adaptive response to the culture of kindergarten, and over time…higher levels of learning and achievement" (p. 1386). Ladd and colleagues' line of reasoning was bolstered by and, in part, based upon the findings of Finn (1993). Finn found that such participatory behaviors fostered powerful learning and reading and mathematics skill development, beyond that attributable to demographic characteristics in a nationwide sample of transitioning middle school students.

McClelland, Morrison, and Holmes (2000) extended the work of Cooper and Farran (1988) using their behavioral rating scale to examine work-related skills as a predictor of early literacy (i.e., letter and word recognition, letter naming, receptive/picture vocabulary, passage reading) and mathematics (e.g., number recognition, addition, multiplication) upon kindergarten entry and at the end of second grade. McClelland and her colleagues found that entry work-related skills predicted modest though unique variance in all achievement outcomes beyond other predictors of early achievement (i.e., IQ, entrance age, amount of preschool experience, parental education level, ethnicity, and home literacy environment) at *both* near and distal time points. Children with lower work-related skills ratings scored significantly lower on achievement measures at the beginning of kindergarten and at the end of second grade (controlling for earlier achievement). The key finding here was that work-related skills (representing the PM) continued to significantly predict early literacy, reading and math skills over the first three years of students' public schooling.

McClelland and Morrison (2003) later teamed again to explore whether "learning-related social skills" were present in preschoolers over two time points one year apart. McClelland and Morrison used the Social Skills subscale of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) and the Mastery Behaviors subscale of the Child Behavior Rating Scale (CBRS; Bronson, Goodson, Layzer, & Love, 1990) to measure skills in the domains of independence, responsibility, self-regulation, and social cooperation. Although learning-related behavioral skills did not change within the preschoolers, the researchers observed significant variation between students at both time points. Perhaps the most important of McClelland and Morrison's findings was the presence of lower- and higher-levels of learning-related behaviors in pre-kindergarten students—similar to the findings of McClelland et al. (2000) with kindergarten and second grade students.

In another extension of their previous work the McClelland team investigated if learning-related skills (measured just after kindergarten entry) predicted the level and growth in reading and math achievement over elementary school (McClelland, Acock, & Morrison, 2006). The researchers once again used the Cooper-Farran Behavioral Rating Scales (Cooper & Farran, 1988) to measure students' learning-related behaviors, broadly framing such skills as self-regulation and social competency (e.g., self-control, staying on task, organizing work materials, working independently, listening, following directions, and participating appropriately in student groups). McClelland and colleagues found that kindergartners' learning-related behavioral skills significantly predicted the level of reading and math skills between kindergarten and sixth grade while controlling for background variables. Additionally, learning-related skills significantly predicted growth of early reading and math skills between kindergarten and second grade, though not between third and sixth grade. Lower learning-related skills were associated with lower reading and math scores and lower growth, commensurate with the team's previous studies (i.e., McClelland & Morrison, 2003; McClelland et al., 2000). Together the studies by the McClelland team suggest that participatory behaviors as described by Sfard (1998) are measurable in early school contexts, that absent intervention they persist across prekindergarten and into the early primary years, and that they positively relate to the level and growth in academic skills both proximal and distal to kindergarten entry.

Summary and Study Context

Founded on Sfard's (1998) theoretical perspective, findings from the previous empirical research on acquisition- and participation-related skills suggest a number of important inferences. First, students' early learning-related skills appear largely characterized by self-regulation and social behavioral skills, while early literacy is characterized by alphabetic and phonemic awareness skills and by the end of kindergarten early/emergent reading skills (e.g., vocabulary, word reading), and early math by numeracy- and early arithmetic-related skills. Second, like the early literacy and mathematics achievement skills, entry learning-related skills are measurable prior to and after kindergarten entry. Third, learning-related skills are consistently and positively related to early literacy and math achievement status and growth over the short and longer-term in the early primary years. Fourth, lower literacy and math scores appear consistently related to lower learning-related behavioral skills while controlling for prior achievement and key demographic factors. Broadly then, it would appear that teacher ratings of student participatory (learning-related) behaviors may be an important assessment tool, in concert with measures of early literacy and math achievement, upon kindergarten entry as a means to identify students at risk of poor academic outcomes at the end of kindergarten and beyond. Consequently, it is with justification from both theory and empirical research findings on the relation of early learning-related behaviors and academic skills that I present the following dissertation study framed by a statewide kindergarten entry assessment initiative.

Research Questions

The purpose of this dissertation study is to identify and investigate the relation between student learning-related behavioral and academic achievement skills upon entry into kindergarten, and the relation of these entry skills to the level of early/emergent literacy and mathematics achievement in the spring of the same school year. Specifically, I address the following two research questions:

1. What are the underlying dimensions (latent factors) and interrelations of the learning-related behavioral and academic skill components of the OKA? I hypothesize that the underlying factor structure of the OKA replicates that which is formally reported by the ODE and supports the preliminary findings of Tindal et al. (Manuscript submitted for publication)—two learning-related behavior factors (Self-regulation and Social-interpersonal) and a single academic proficiency skill factor comprised of the three achievement measures (LNF, LSF, and Numbers and Operations).

2. What is the relation of kindergarten students' entering learning-related behaviors and academic skill to the level of early/emergent literacy and mathematics achievement measured in the spring of the same school year when controlling for student demographic

characteristics? I hypothesize that on average, the greater students' entering level of academic skill and self-regulation behaviors, the greater their achievement on all spring achievement measures will be. Further, I hypothesize that the positive effect of entry academic achievement substantively exceeds that of the effects of either entering learning-related behaviors. I base this hypothesis on the fact that prior achievement almost ubiquitously predicts proximal-later achievement in early literacy and math (e.g., Lembke & Foegen, 2009a; Linklater et al., 2009; Speece et al., 2004; VanDerHeyden et al., 2006), and on the curious negative relation between entry academic skills and social-interpersonal behaviors estimated in analysis of 2012-2013 OKA pilot data (Tindal et al., Manuscript submitted for publication). Finally, I hypothesize that prior achievement (represented by kindergarten students' entering academic skill composite in the OKA) renders the influence of some demographic factors statistically non-significant or practically unsubstantial, with negative effects of Economic Disadvantage, Disability Status and Limited English Proficiency being likely exceptions (U.S. Department of Education, 2015).

CHAPTER II

METHODS

Two extant datasets are used with data collected during the 2013-2014 academic year. The Oregon Department of Education (ODE) provided the first dataset and represents students' learning-related behavioral ratings and academic skills open entering kindergarten as measured by the initial statewide field-test of Oregon's Kindergarten Assessment (OKA). The second dataset comes from the easyCBM interim-formative assessment database (Alonzo, Tindal, Ulmer, & Glasgow, 2006) and represents early literacy and mathematics achievement for a portion of the statewide sample measured in the spring of the same kindergarten school year. As extant data, the sample is one of convenience rather than design so causal inferences are not appropriate. I structure the methods by presenting a description of the sample (demographic characteristics) and data preparation, including, measures and statistical analyses).

Sample and Data Preparation

The OKA was administered to a cohort of approximately 43,000 kindergarten students in September-October 2013 and easyCBM interim early literacy and mathematics benchmark assessments administered to a smaller portion of the kindergarten cohort in spring 2013. Extant data were cleaned and merged using SPSS version 22 for Macintosh prior to statistical analyses (SPSS Inc., 2010). Only students with a valid total score on one or more measure of the OKA were included in the state dataset and only students with a valid spring score on *any of the selected benchmark measures* were included in the easyCBM dataset. For all measures (see next section), negative scores and those scores that fell outside of the acceptable interpretation range of the respective assessments were deleted and coded as missing.

Initially, the OKA dataset included 43,072 students. A count variable was created that totaled the number of OKA assessment measures (out of a possible four possible) that were included in state reporting for student group averages (e.g., by demographic, district, or school group). A total of 842 students had not taken any of the four OKA test measures and were recorded as missing.

Of the remaining students in the OKA dataset, 997 (2.4%) had results from one measure, 560 cases (1.3%) had results from two measures, 1,710 (4.0%) had results for three measures, and 38,963 cases (92.3%) had all four OKA measures. For all cases that were not included in state reporting of group averages, item and total scores of zero were deleted and coded as missing. If an OKA measure was included for state reporting, zeroes for both individual items and total test scores were retained as long as the student was flagged as having attempted the given test segment; otherwise, they were deleted and coded as missing.

For the Numbers and Operations (Early Math) measures, 582 cases were missing data for all 16 items and did not have a total score reported. Of the remaining cases, 4,403 students were flagged as having taken the Spanish language equivalent version of the Numbers and Operations test. Additionally, another 312 students had a total score of zero, but were flagged as having not attempted the measure. These data were deleted and coded as missing. Finally, individual items were summed and matched to the total score as a check for accuracy. Across all student cases in the dataset, 40,588 were deemed valid for the Numbers and Operations test segment of the OKA. For the LNF measure, 494 students did not have a total score reported and were coded as missing for analyses. Across all student cases in the dataset, 40,676 students were deemed valid for the LNF test segment. For the LSF test segment, 864 students did not have a total score reported and were coded as missing. A total of, 40,306 students were deemed valid for the LSF test segment of the OKA. In the end, the full analytic sample for the OKA extant dataset included test results from 694 schools in 189 districts in Oregon.

For the Approaches to Learning measure of the OKA, 806 cases were missing data for all 15 items comprising the behavioral rating measure, and 769 item ratings (across all student cases) were out-of-range and therefore deleted and coded as missing. Individual items were summed and matched to the total score as a check for accuracy. Across all students in the dataset, 40,364 cases were deemed valid for the Approaches to Learning test segment.

A similar data cleaning process was used for the easyCBM extant dataset, whereby the sample was restricted to only Oregon kindergarten students with a valid score on one or more of the spring interim benchmark assessments, with instances of repeated district identification numbers rectified or deleted. In summary, and prior to merging with the OKA dataset, 9,526 Oregon kindergarten students had a valid score for the easyCBM LSF interim benchmark assessment; 9,564 students had a valid score for the PSF benchmark; 9,534 students had a valid score for the WRF benchmark; and 5,185 students had a valid test score for the spring math benchmark. Separate datasets for each interim benchmark measure were merged into a single extant dataset using the students' easyCBM identification number. In the end, the analytic sample for the easyCBM extant dataset included spring interim benchmark assessment results from 159 schools in 49 districts in Oregon.

To merge the two extant datasets, three unique identifiers were created. The first identifier combined students' 4-digit district identification number with the district-given student identification number assigned in fall 2013 (the beginning of the school year for this study) in both the OKA and easyCBM datasets. Because many students switched districts during the school year, a second unique identifier used the 4-digit district identification number with the district-given student identification number recorded for spring 2014 (the end of the school year in this study). The final unique identifier was the SSID, which was included for every student in the OKA dataset and was used as the easyCBM identification number for many students. After merging the two extant datasets, 7,199, 7,275, 7,236, and 4,246 Oregon kindergarten students with a valid interim benchmark score for the LSF, PSF, WRF, and NCTM Math, respectively, were matched/merged with the OKA dataset and included in the analytic sample. For the easyCBM extant dataset, roughly 76% of Oregon kindergarten students had a valid score on the emergent literacy benchmark assessments and 82% of students had a valid score on the NCTM Math benchmark assessment.

Table 1 displays the demographic counts and percentages for the (statewide) full analytic sample in this study, each random subsample used in factor analyses, and for the easyCBM-matched subsample used in structural modeling measured across all measures. Demographics counts and percentages are taken from the OKA dataset and are complete for all student cases. As shown in Table 1, demographic makeup is comparable across all (sub)sample populations for sex, Nonwhite, disability status, economic disadvantage status, and limited English proficiency status. The largest difference between samples is that there are roughly 6% more White students in the easyCBM-matched subsample, and 6% less Hispanic students as compared to the full analytic sample and each 50% random subsample, which may limit generalizability of inferences.

Measures

Measure development and technical adequacy is described in the following section. As noted in more detail, the two measures are highly related with the OKA being a subset of the easyCBM interim-formative assessment system. The OKA early literacy measures used alternate (progress monitoring) forms of the benchmark easyCBM and the math used a subset (progress measure) of the easyCBM NCTM Math benchmark.

Oregon Kindergarten Assessment. The OKA is an individually administered assessment battery consisting of measures in three domains: early literacy, early numeracy, and learning-related behaviors/interactions. Teachers administer three achievement measures, rate their students on observed behavioral frequencies, and upload scores to a secure website. The early literacy and numeracy measures included in the OKA are single grade-level progress-monitoring test forms from the easyCBM interim-formative assessment system. Included are two measures of alphabetic early literacy (LNF and LSF) and an early numeracy measure (Numbers and Operations). In addition, the rating scale, called Approaches to Learning, requires teacher judgments about students' behavior in the classroom. All technical adequacy information for these achievement measures is presented under the section on easyCBM (following a description of the Approaches to Learning measure).

Table 1

Demographics for Statewide Full Analytic Sample, Random Subsamples, and easyCBM-matched Subsample

Demographic Characteristic	Full Analytic		EFA50		CFA50		easyCBM	
	n	%	*n*	%	*n*	%	*n**	%
All Students	41,170	100.00	20,585	100.00	20,585	100.00	9,164	100.00
Sex								
Female	20,074	48.76	9,978	48.47	10,096	49.05	4,524	49.37
Male	21,906	51.24	10,607	51.53	10,489	50.95	4,640	50.63
Race/Ethnicity								
Asian	1,410	3.42	684	3.32	726	3.53	392	4.28
Black	977	2.37	506	2.46	471	2.29	188	2.05
Hispanic	9,790	23.78	4,867	23.64	4,923	23.92	1,564	17.07
American Indian/Alaskan Native	553	1.34	287	1.39	266	1.29	112	1.22
Multi-Ethnic	2,310	5.61	1,149	5.58	1,161	5.64	594	6.48
Pacific Islander	316	0.77	157	0.76	159	0.77	47	0.51
White	25,814	62.70	12,935	62.84	12,879	62.56	6,267	68.39
Disability Status								
Non-disability	37,276	90.54	18,641	90.57	18,635	90.53	8,341	91.02
Disability	3,894	9.46	1,944	9.44	1,950	9.47	823	8.98
Economic Status								
Not Economically Disadvantaged	19,251	46.76	9,644	46.85	9,607	46.67	4,252	46.40
Economically Disadvantaged	21,919	53.24	10,941	53.15	10,978	53.33	4,912	53.60
English Proficiency Status								
Not Limited English Proficient	33,601	81.62	16,854	81.88	16,747	81.36	8,055	87.90
Limited English Proficient	7,569	18.38	3,731	18.12	3,838	18.64	1,109	12.10

Note. Demographic breakdown by full analytic sample, the two 50% random subsamples, and the matched easyCBM subsample using both count and percentages relative to the associated (sub)sample. *casewise deletion.

The OKA was piloted in September-October 2012 with a representative sample of 1,228 kindergarten students from 16 schools in 13 districts (Oregon Department of Education, 2013b). Over 2012-2013, the ODE in collaboration with the Oregon Early Learning refined OKA content and administration procedures. Training and support materials were developed with particular attention to English Language Learners and those enrolled in special education. A series of informational and training webinars were conducted throughout the state, and a website with training/support materials was built (http://www.ode.state.or.us/search/page/?=3908). Required trainings took place and the OKA was formally field-tested in September-October 2013 with kindergarten students from across Oregon, making up the original database for this study. Descriptive statistics for all OKA measures for the full analytic sample are shown in Table 2, with additional descriptive statistics for all subsamples displayed in Table 3. Descriptive statistics for the OKA by demographic group are shown in Tables B.1 and B.2 in Appendix B.

Approaches to Learning. This measure from the OKA uses a portion of the Child Behavior Rating Scale (CBRS; Bronson et al., 1990) and is based on the Bronson Social and Task Skill Profile (Bronson, 1994). The measure focuses on the frequency of learning-related behavioral strategies students use in typical classroom situations. It is comprised of 15 items and uses a five-point scale. Teachers rate students on the frequency with which they observe such behaviors (1 = *never*, 2 = *rarely*, 3 = *sometimes*, 4 = *frequently/usually*, and 5 = *always*). Table 4 lists the 15 item stems comprising the Mastery Behaviors scale of the CBRS, along with means and standard deviations for the full analytic sample and two random subsamples (EFA50, CFA50). Item abbreviations used throughout this study, including in figures, are bolded.

Table 2

Descriptive Statistics for 2013-14 OKA Total Scores (Full Analytic Sample)

OKA	*n*	Miss	Min	Max	*M*	*SD*	Skew	Kurtosis
LNF	40,676	494	0	100	18.49	16.71	0.74 (0.01)	-0.09 (0.02)
LSF	40,306	864	0	100	6.72	9.71	1.79 (0.01)	3.12 (0.02)
Math*	40,588	582	0	16	8.02	3.17	0.24 (0.01)	-0.38 (0.02)
SR**	40,364	806	10	50	35.35	8.52	-0.38 (0.01)	-0.18 (0.02)
Social**	40,364	806	0	25	19.51	4.37	-0.67 (0.01)	0.12 (0.02)
AL total**	40,364	806	14	75	54.85	12.14	-0.45 (0.01)	-0.09 (0.02)

Note. Total n = 41,170 casewise. Reported n vary based on cleaning procedures, and pairwise deletion, with the number of missing values (Miss) displayed for each measure relative to total casewise sample, where: LNF = Letter Names Fluency, LSF = Letter Sounds Fluency, Math = Numbers and Operations (academic skill measures), and SR = Self-regulation and Social(-interpersonal) = sub-measures of the Approaches to Learning (AL) behavioral rating measure.

*Of the total, the state flagged 4,403 (10.7%) students for the Spanish language version of the Numbers and Operations math assessment.

**SR and Social descriptive statistics are based on sub-scores totaled for items 1-10 and 11-15, respectively, from the Approaches to Learning segment of the OKA—these results based on exploratory and confirmatory factor analyses. AL total statistics represent total score (items 1-15) from the Approaches to Learning segment.

Table 3

Descriptive Statistics for 2013-2014 OKA for Full Analytic Sample, Random Subsamples, and easyCBM-matched Subsample

OKA Segment	Full Analytic			EFA50			CFA50			easyCBM		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
LNF	40,676	18.49	16.71	20,351	18.47	16.77	20,325	18.52	16.65	9,114	19.74	16.55
LSF	40,306	6.72	9.71	20,153	6.70	9.73	20,153	6.74	9.68	9,102	6.94	9.82
Math	40,588	8.02	3.17	20,301	8.03	3.18	20,287	8.01	3.16	9,072	8.13	8.13
SR*	40,364	35.35	8.52	20,190	35.31	8.55	20,174	35.38	8.48	9,098	35.58	8.47
Social*	40,364	19.51	4.37	20,190	19.49	4.38	20,174	19.52	4.36	9,098	19.50	4.40
AL total*	40,364	54.85	12.14	20,190	54.81	12.19	20,174	54.90	12.09	9,098	55.08	12.16

Note. Full analytic n = 41,170 and EFA50/CFA50 n = 20,585 casewise, with missing data ≤ 2.1% for all OKA measures. easyCBM sample matched sample n = 9,164 casewise, with missing data ≤ 1.0% for all easyCBM measures. Means and spread for the four OKA battery measures are comparable across the four (sub)samples.

*SR and Social descriptive statistics are based on sub-scores totaled for items 1-10 and 11-15, respectively, from the *Approaches to Learning* behavioral rating segment of the OKA—these based on exploratory and confirmatory factor analyses. AL total statistics represent total score (items 1-15) from the *Approaches to Learning* segment.

Table 4

Abbreviations and Descriptive Statistics for 2013-2014 Approaches to Learning (Child Behavior Rating Scale; CBRS) of OKA (Full Analytic Sample, EFA50, CFA50)

	Full		EFA50		CFA50	
CBRS item stem	*M*	*SD*	*M*	*SD*	*M*	*SD*
1. Observes rules and **follows** directions without requiring repeated reminders	3.60	0.99	3.60	0.99	3.60	0.99
2. **Completes** learning tasks involving two or more steps (e.g. cutting and pasting) in organized way.	3.66	0.99	3.65	0.99	3.66	0.99
3. Completes tasks **successfully**.	3.71	0.92	3.71	0.92	3.71	0.91
4. **Attempts** new challenging tasks.	3.65	0.98	3.65	0.98	3.66	0.97
5. **Concentrates** when working on a task; is not easily distracted by surrounding activities.	3.35	1.04	3.35	1.04	3.35	1.03
6. **Responds** to instructions and then begins an appropriate task without being reminded.	3.56	1.03	3.56	1.03	3.57	1.02
7. Takes **time** to do his/her best on a task.	3.67	0.96	3.67	0.96	3.67	0.95
8. **Finds** and organizes materials and works in an appropriate place when activities are initiated.	3.69	0.93	3.69	0.93	3.69	0.92
9. Sees own **errors** in a task and corrects them.	3.00	1.01	3.00	1.01	3.01	1.01
10. **Returns** to unfinished tasks after interruption.	3.49	0.97	3.48	0.97	3.49	0.96
11. Willing to **share** toys or other things with other children when playing; does not fight or argue with playmates in disputes over property.	3.90	0.89	3.90	0.89	3.91	0.89
12. **Cooperative** with playmates when participating in a group play activity; willing to give and take in the group, to listen to or help others.	3.89	0.92	3.89	0.92	3.89	0.92
13. Takes **turns** in a game situation with toys, materials, and other things without begin told to do so.	3.91	0.92	3.90	0.92	3.91	0.91
14. **Complies** with adult directives, giving little or no verbal or physical resistance, even with tasks.	3.91	0.99	3.91	0.99	3.91	0.99
15. Does not **fuss** when he/she has to wait briefly to get attention from teacher or other adult; child may be asked once to wait by the teacher or adult.	3.92	1.02	3.92	1.02	3.92	1.01

Note. Total n = 40,364. The Self-regulation (SR) skills latent factor is comprised of items 1-10 (above dividing line), while the Social-interpersonal (S) skills latent factor is comprised of items 11-15 (below dividing line). Bolded words represent the item abbreviations used in this study.

Reliability and validity evidence. Tindal et al. (Manuscript submitted for publication) documented very strong internal consistency in their analysis of OKA pilot data from 2012-2013. In an older study, Abt Associates (1988) documented strong internal consistency (α = .96), and moderate test-retest reliability from fall to spring (r = .67). Bronson, Tivnan, and Seppanen (1995) found moderate relation to the Preschool Inventory (r = .34), a measure of early cognitive achievement. Later, McClelland and Morrison (2003) reported high internal consistency (r = .95) in a study of preschoolers (ages 3-5).

easyCBM early and emergent literacy and math. This study uses a series of early/emergent literacy and math measures from the easyCBM interim-formative assessment system. The LNF and LSF early literacy measures used in the OKA are alternate progress forms of the seasonal benchmarks in the easyCBM system. The early numeracy measure (Numbers and Operations), the third academic measure of the OKA, is subset of the NCTM Math benchmark. The remaining measures, LSF, PSF, WRF and NCTM Math, are spring benchmark measures. All easyCBM measures, comprising both the OKA and dependent endogenous outcomes in structural modeling are described in what follows.

The LSF, PSF, WRF and NCTM Math benchmark measures were administered to Oregon kindergartners (a portion of the full analytic sample) in the spring of the 2013-2014 academic year. For each early/emergent literacy measure, raw scores are calculated based on the number of letter sounds, phonemes and words correctly spoken, with students' self-corrections counting as correct responses. For NCTM Math, raw scores are calculated based on the number of math problems that students answer correctly out of a possible 45 test items.

One unique characteristic of easyCBM early literacy and mathematics measures (including the LNF measure) as compared to other interim-formative assessment systems is that items were scaled with a Rasch model (Alonzo & Tindal, 2007a, 2007b, 2009d). Items were compiled into test forms to maximize form comparability *within* each grade; thus, average difficulty is considered approximately equivalent across all test forms for a given grade-level measure. Observed changes on these measures is attributable to changes in students' letter sounding, phoneme segmentation, word reading, and math skills, rather than to variance in difficulty across test forms. Results from Rasch analyses help ensure test forms have adequate difficulty range to sufficiently screen students of varying skill level into risk categories, along with an adequate number of items at the lower tail of the distribution in order to detect small changes (growth) in performance of students whose early/emergent literacy and math skills are monitored over time. The difficulty and fit of each item in the early literacy and NCTM Math benchmark test forms are reported in a series of technical reports (Alonzo & Tindal, 2007a, 2007b, 2009d). Descriptive statistics for easyCBM early/emergent literacy (LSF, PSF, and WRF) and NCTM Math spring benchmarks for the easyCBM-matched subsample are displayed in Table 5. Descriptive statistics for the spring benchmarks by demographic group are shown in Table B.3 in Appendix B.

Table 5

Descriptive Statistics for 2013-2014 easyCBM Spring Benchmark Measures

Measure	*n*	Min	Max	*M*	*SD*	Skew	Kurtosis
LSF	7,199	0	107	33.57	14.86	0.27 (0.03)	0.59 (0.06)
PSF	7,275	0	70	40.15	15.52	-0.47 (0.03)	0.00 (0.06)
WRF	7,236	0	60	14.86	14.35	1.76 (0.03)	2.64 (0.06)
Math	4,246	0	45	36.36	6.44	-1.56 (0.04)	3.77 (0.08)

Note. Total n = 9,164 cases. Reported n vary based on cleaning procedures described in Methods section, and pairwise deletion, relative to total easyCBM-matched subsample, where: LSF = Letter Sounds Fluency, PSF = Phoneme Segmenting Fluency, and Math = NCTM Math.

Letter names fluency. The LNF measure assesses students' ability to name the letters of the English alphabet aloud (Alonzo & Tindal, 2007a). LNF is individually administered with students shown lower case and capitalized letters organized in rows on a one-page test form and instructed to name as many as they can in 60 seconds. The maximum score is 100 letters correctly named per minute.

Reliability evidence. Alonzo and Tindal (2007a) found alternate test form reliability ranged from .82 to .89, and test-retest reliability ranged from .79 to .82 across three test forms. Alonzo and Tindal (2009b) documented alternate test form reliability to range from .87 to .90 across several test forms, providing evidence that students' letter naming fluency scores were quite stable regardless of the test form administered.

Validity evidence. Lai, Alonzo, and Tindal (2013) found a strong relation between the easyCBM LNF benchmarks and the spring Stanford-10 Word Reading assessment (r = .75). Wray, Lai, Saez, Alonzo, and Tindal (2014) the easyCBM LNF seasonal benchmarks with the Dynamic Indicators of Basic Early Literacy Skills

(DIBELS) LNF measure using a large sample of kindergarten and first grade students in the Pacific Northwest. Overall, nonparametric correlations were high, $\rho = .86$ for kindergarten and $\rho = .80$ for first grade. Wray et al. (2014) also found that LNF scores loaded strongly onto the latent reading trait (factor) across seasonal models, suggesting that reading ability is a strongly predicted by performance on the LNF measure.

Letter sounds fluency. The LSF measure assesses students' skill in orally sounding letters of the English alphabet (Alonzo & Tindal, 2007a; Lai et al., 2010). The measure is individually administered with students shown a series of lower case and capitalized letters organized in rows on a one-page test form, and instructed to sound out as many letters as they can in 60 seconds. The maximum score is 110 letter sounds correctly named per minute.

Reliability evidence. Alonzo and Tindal (2009b) found strong alternate test form reliability of .88 to .92 for several kindergarten LSF test forms. Test-retest reliability was investigated in two separate studies by Alonzo and Tindal (2009b) and Wray et al. (2014). Alonzo and Tindal documented moderate test-retest correlations ranging from .64 to .68, whereas Wray and colleagues, using different test forms, found strong test-retest correlations of .77 to .87.

Validity evidence. The relation of the LSF measure was moderate with both the DIBELS Initial Sound Fluency measure ($\rho = .55$) and the DIBELS Nonsense Word Fluency measure ($\rho = .58$; Lai et al., 2013), and strong with the Stanford-10 Word Reading ($r = .71$; Wray et al., 2014). Lai et al. (2010) showed that the kindergarten LSF benchmark scores loaded strongly onto the latent reading factor across seasonal models, with fair to good model fit. Lai and colleagues also examined the utility of LSF growth across seasonal benchmark testing in predicting Stanford-10 Reading scores. The researchers found moderate to large effects of .51 for the lower two quartiles to .41 for the upper quartile for the kindergarten sample.

Phoneme segmenting fluency. PSF is an early literacy test of phonemic awareness that assesses students' skill in identifying phonemes within a series of grade-level words (Alonzo & Tindal, 2007a). Assessors individually administer the measure, saying aloud each word to the student, with students verbally segmenting as many words into phonemes as they can in 60 seconds. The maximum score is 70 phonemes correctly named per minute.

Reliability evidence. Anderson, Park, Lai, Alonzo, and Tindal (2012) documented that alternate test form reliability ranged from .81 to .90 for the kindergarten measure. Alonzo and Tindal (2009a) reported correlation test-retest coefficients of .45 to .47, indicating a modest relation. Anderson and colleagues found that test-retest correlations ranged from .32 to .81, with a median value of .57, indicating a low to strong relation. Sample size was quite small for these studies (n = 19 to 42).

Validity evidence. The concurrent relation between PSF benchmark measure and the DIBELS PSF measure was strong (ρ = .85; Lai et al., 2013), and moderate with the Stanford-10 Word Reading (ρ = .41; Wray et al., 2014). PSF benchmark scores loaded moderately onto the latent reading factor across seasonal models (Lai et al., 2010). Lai and colleagues also examined PSF growth across the three seasonal testing occasions and found it generally significant, positively related, and a modest predictor of spring Stanford-10 achievement as compared to raw PSF score. Sáez, Irvin, Alonzo, and Tindal

(2012) documented weak alignment between the PSF kindergarten measure and the Common Core State Standards (CCSS) Foundational Skills.

Word reading fluency. WRF is a word reading fluency measure that assesses students' skill in reading single grade-level words aloud. The measure is individually administered with students shown a series of English language words organized in rows on a one-page test form and instructed to read as many words as they can in 60 seconds. The maximum score in kindergarten is 60 words correctly read per minute.

Reliability evidence. Wray et al. (2014) evaluated alternate form reliability for nine kindergarten WRF test forms across five time points from winter to spring 2012-2013 and found moderately strong to very strong correlations between the measures, ranging from .74 to .94, with the strongest relations found between measures administered at the same time point. Alonzo and Tindal (2009b) investigated alternate form and test-retest reliability of three Grade 1 WRF test forms, and found correlations ranged from .95 to .96. Anderson et al. (2012) investigated alternate form reliability using four additional Grade 1 test forms and found correlations ranged from .89 to .97. Alonzo and Tindal (2009b) reported that test-retest correlations ranged from .94 to .95,, while Anderson et al. (2012) found test-retest correlations ranged from .87 to .95, for their respective studies..

Validity evidence. Wray et al. (2014) found that WRF was a significant predictor of spring performance on the Stanford-10, with the unique variance explained ranging from 13% to 21% across the four time points examined. In the same study, WRF was regressed upon by easyCBM LNF, LSF and PSF. The variance explained by the model across time ranged from 49% to 56%. LNF (14% to 41% unique variance explained) and

LSF (1% to 2% unique variance explained) were significant predictors of spring WRF at all time points, while PSF was not a significant predictor at any time point.

NCTM Math. The easyCBM benchmark math measure, written to align with NCTM Focal Point Standards, includes three seasonal test forms designated for benchmark screening and 30 designated for progress monitoring, 10 test forms assessing each of the three Focal Point standard skill domains at each grade level (Alonzo, Lai, & Tindal, 2009a, 2009b, 2009c; Alonzo & Tindal, 2009c, 2009d; Lai, Alonzo, & Tindal, 2009a, 2009b, 2009c, 2009d). All math test forms are designed to be group administered using computers, with paper-pencil versions available. Kindergarten NCTM Math benchmarks each have 45 total multiple-choice items, with the number of items divided as equally as possible across the test to address the three NCTM Focal Point standards skill domains in Measurement, Geometry, and Numbers and Operations—the latter making up the math measure included in the OKA battery.

In addition to developing alternate grade-level test forms of approximate equivalent difficulty, test development aimed to maximize accessibility using principles of Universal Design for Assessment (e.g., precisely defined construct targets, consideration of all potential test-takers, non-biased items) as outlined by Thompson, Johnstone, and Thurlow (2002) and Johnstone, Altman, and Thurlow (2006), as well as guidelines to writing quality multiple-choice items given by Haladyna (2004) and Downing (2006a, 2006b). In accordance with these principles, a Spanish language versions of all test forms are also available (Alonzo & Tindal, 2009d).

Reliability evidence. For the kindergarten NCTM Math measures, internal consistency ranged from .83 to .87, and split-half reliability ranged from .80 to .82 across the three seasonal benchmark screeners (Anderson et al., 2010).

Validity evidence. Anderson et al. (2010) documented evidence of acceptable item functioning and unidimensionality (across the Focal Point Standards). Wright and Linacre (1994) suggested mean square outfit for less high stakes decision-making should range from .70 to 1.30, criteria largely met by the kindergarten NCTM Math measures.

Anderson et al. (2010) also compared the NCTM Math benchmarks to the mathematics portion of the comprehensive TerraNova 3 test battery (CTB McGraw-Hill, 2010) and found approximately 53% of the total variance in the TerraNova 3 accounted for by the three seasonal benchmarks. The spring benchmark had the highest regression coefficient, uniquely explaining 12.4% of the variance. A spring model, run to examine the concurrent validity evidence of easyCBM NCTM Math, was also significant accounting for approximately 52% of the variance in TerraNova 3.

Nese et al. (2010) found that across grades, Focal Points, and test forms, the ratings of easyCBM math items aligned to NCTM Focal Points standards were generally strong. Irvin, Park, Alonzo, and Tindal (2012) investigated alignment with the CCSS and found reasonable ratings, though gaps in alignment were also observed. Generally, Irvin and colleagues found that benchmark items appeared more strongly aligned with on-grade CCSS compared to prior-grade standards, with alignment much stronger at the CCSS domain level as compared to the individual standard level.

Variables

Predictor variables used in the structural modeling were based on results of the exploratory and confirmatory factor analyses of the OKA in the current study, as well as previous analysis of 2012-2013 OKA pilot data (Tindal et al., Manuscript submitted for publication) that preliminarily confirmed the states' theoretical kindergarten entry model (Oregon Department of Education, 2014). Three separate continuous latent factor variables are included as predictors of spring achievement. The first factor is a latent academic achievement (*Academic Skill Proficiency*) indicating the level of academic skill upon entrance into kindergarten. It is comprised of LNF, LSF and Numbers and Operations of the OKA. Two additional latent factors are comprised of items from the *Approaches to Learning* (CBRS) measure of the OKA, representing two distinct though related behavioral constructs (*Self-regulation* and *Social-interpersonal*). Scores on the easyCBM spring LSF, PSF, WRF achievement benchmarks are loaded onto a single latent endogenous outcome factor termed *Emergent Literacy*, and *NCTM Math* is included as a separate continuous outcome variable (see Figure 1).

Six student-level demographic covariates were included as covariate predictors of spring achievement in structural modeling: (a) Sex [0 = male, 1 = female]; (b) Nonwhite-Hispanic [0 = Not Nonwhite-Hispanic, 1 = Nonwhite-Hispanic]; (c) Nonwhite-Non-Hispanic [0 = Not Nonwhite-Non-Hispanic, 1 = Nonwhite-Non-Hispanic]; (d) Disability Status [0 = Non-disability, 1 = Disability]; (e) Economic Disadvantage [0 = Not Economically Disadvantaged, 1 = Economically Disadvantaged]; and (f) Limited English Proficiency [0 = Not Limited English Proficient, 1 = Limited English Proficient].

Analysis

A combination of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were used in documenting the factor structure underlying and learning-related and academic achievement skill relations within the OKA (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Floyd & Widaman, 1995; Preacher & MacCallum, 2003). The subsequent structural equation model (SEM) was specified to document the relation of kindergarten entry skills' to early/emergent literacy and math achievement measured in the spring of kindergarten controlling for student demographic characteristics. The EFA was conducted on a random subsample (50% of the full analytic sample). An independent replication of the factor structure using CFA was conducted using the remaining 50% of the full analytic sample. Results of these factor analyses, consistent with previous results from the analysis of 2012-2013 OKA pilot data (Tindal, Irvin, & Nese, April 2013; Tindal et al., Manuscript submitted for publication) informed the measurement model portions of the structural modeling. EFA, CFA and SEM analyses are described in what follows.

Exploratory factor analysis. Prior to conducting the EFA data were analyzed descriptively using SPSS v.22 for Macintosh (IBM Corp., 2012). EFA was used to account for both the common and unique variance associated with the extracted common factors (Fabrigar et al., 1999). The EFA was conducted using Mplus version 7.3 (Muthén & Muthén, 1998-2012) with maximum likelihood and robust standard error . This method, termed MLR in Mplus, uses the unreduced correlation matrix to estimate parameters (Fabrigar et al., 1999). MLR, a full information likelihood estimator, is capable of handling both continuous and categorical data and is a preferred method when missing are present (≤ 2.1% aross all OKA segments; Arbuckle, 1996; Muthén & Muthén, 2015b). Additionally, MLR does not assume multivariate normality, an important consideration given the positive (OKA academic achievement measures) and negative (OKA learning-related behavior ratings) skew observed in OKA data (see Table 2). To allow factors to correlate, a reasonable assumption given the interrelatedness of the academic and learning-related behavioral skills in theory (akin to acquisition and participatory skills as described by Sfard, 1998) and in previous research (e.g., Ladd et al., 1999; McClelland et al., 2006), geomin rotation was used (Preacher & MacCallum, 2003), the default oblique rotation method in Mplus (Muthén & Muthén, 1998-2012).

Based on analysis of 2012-2013 OKA pilot data (Tindal et al., Manuscript submitted for publication), I specified and compared two factor solutions. The first was a two-factor solution, comprised of a single OKA academic proficiency skills latent factor (made up of LNF, LSF and Numbers and Operations) and a single learning-related behavioral factor comprised of all 15 items from the CBRS. The second was a three-factor solution, comprised of the same academic proficiency skills latent factor as in the two-factor solution, though with two behavioral latent factors comprised of items from the CBRS.

It was reasoned that the behavioral rating items comprising the CBRS would either load to a single latent factor (i.e., with an underlying construct of "learning-related behaviors") or most likely that the items would separate into two distinct factors (i.e., underlying constructs of "self-regulation" and "social-interpersonal" learning-related behaviors) as reported by ODE and observed in EFA of 2012-2013 OKA pilot data (Tindal et al., Manuscript submitted for publication). The distinct EFA solutions were specified a priori and compared based on reasonableness and interpretability of the item/measure factor loadings, with the most parsimonious factor structure sought, where each factor explains as much variance as possible in the items and measures comprising the OKA (Kaplan, 2009). The two factor solutions were also compared using the chi-square difference test and the Akaike and Bayesian information criteria (AIC and sample-size-adjusted BIC, respectively) available when using MLR estimation with categorical data. For AIC and BIC information criteria, lower values indicate a better fitting model to the observed data, taking into account the simplicity of the model as useful parameters are included in the model (Akaike, 1973; Schwartz, 1978).

Confirmatory factor analysis. Given the preliminary evidence of the factor structure and interrelation documented, I used CFA to confirm the appropriateness of the number of factors, the pattern of indicator loadings, and the correlation between latent factors documented in the EFA of OKA test measures (Boomsma, 2000; Brown & Moore, 2012; Jackson, Gillaspy Jr., & Purc-Stephenson, 2009). I first specified three unidimensional models to independently investigate the structure of the extracted factors. A follow-up CFA concurrently estimated the factor structures and interrelation between the latent factors of the OKA measures. A final CFA tested the interrelation of easyCBM spring benchmark measures to determine whether or not the outcomes measures should be included in structural modeling as a single latent factor (i.e., with an underlying construct of "spring achievement") or as two separate latent factors (i.e., underlying constructs of spring "emergent/early literacy" and "math" achievement).

Like the EFA, I conducted CFA with MLR estimation using Mplus version 7.3 (Muthén & Muthén, 1998-2012) using the remaining 50% random subsample as recommended when confirming a priori hypotheses (Boomsma, 2000). The easyCBM-matched subsample was used for CFA of the spring easyCBM benchmark measures. Model fit to the observed covariance matrix of the measured variables was evaluated based on the reasonableness of the standardized factor loadings, threshold values for the CBRS items, and available model fit information criteria (Akaike, 1973; Hu & Bentler, 1999; Schwartz, 1978).

Structural equation modeling. The effects of entering kindergarten learning-related behaviors and academic achievement skills on student achievement in the spring were modeled concurrently to allow for the effect of one to be estimated while controlling for the others. For example, the effect of entering self-regulation behaviors was estimated beyond that accounted for by social-interpersonal behaviors and initial academic skill proficiency. Furthermore, complex structural relations, essentially regression paths, could then be identified between observed and unobserved variables (Byrne, 2012; Kaplan, 2009; Kline, 2010), while controlling for students' demographic characteristics.

All SEM analysis were conducted using Mplus Version 7.3 (Muthén & Muthén, 1998-2012). One key assumption for estimation in SEM with maximum likelihood (ML) estimation is that all endogenous outcome variables are multivariate normal (Arbuckle, 1996; Byrne, 2012; Kline, 2010). Prior to conducting SEM analyses, I inspected the descriptive statistics (see Table 5) along with univariate and bivariate distributions and scatterplots of the easyCBM benchmark outcomes and determined that the normality assumption was likely met, given skew < 2 and kurtosis < 7 for all benchmark measures (West, Finch, & Curran, 1995), though some distributions visually appeared non-normal (i.e., LNF, LSF). Despite the apparent nonseverity of the violation based on West and colleagues' criteria, I deemed the MLR estimation procedure more appropriate to account for observed non-normality (Muthén & Muthén, 2015b). Though ML is generally adept at handling minor indications of non-normality (Byrne, 2012), MLR more accurately estimates parameters and standard errors when such violations are beyond minor, and gives essentially identical estimates compared to traditional ML when the violations are slight (Muthén & Muthén, 2015b). Lastly, in practice, EFA, CFA and SEM appear relatively robust to non-normality as long as other assumptions (e.g., sufficient sample size, linearity) are met (Gorsuch, 1983). For these reasons OKA and easyCBM spring benchmark data were deemed sufficiently normal for all planned analyses.

Model building. Initially, I specified an SEM without demographic covariates (see Table D.1 in Appendix D) to model the structural relations of the OKA and spring achievement prior to controlling for student characteristics. The selection of variables and a priori relations specified in this SEM were based on the following: (a) acquisition and participation mechanisms of learning and their interrelation as theorized by Sfard (1998), (b) previous empirical research identifying and documenting the correlative and predictive relations between important early literacy and math skills over the kindergarten school year, as described in the literature synthesis, (c) previous analysis of OKA piloting data from Tindal et al. (Manuscript submitted for publication) that provided preliminary evidence of the kindergarten entry factor structures and their relation as operationalized by the OKA, and (d) empirical results from EFA and CFA in this study examining and confirming the factor structure and relation between the OKA measures. I then specified a series of two follow-up SEM, adding all demographic covariates initially (see Table D.2 in Appendix D), and then removing those that were nonsignificant in a single step to arrive at the final model (see Table 10).

Given the findings displayed in the upper measurement model portion all SEM (see Figure 1), I used EFA and CFA results to specify that items 1-10 from the CBRS loaded onto the *Self-regulation* latent predictor (labeled SR), while items 11-15 loaded onto a *Social-interpersonal* latent predictor (labeled SI). Together these two learning-related behavioral factors comprised the *Approaches to Learning* segment of the OKA, and are akin to the participatory behaviors in the metaphor Sfard (1998) described. A third latent factor *Academic Skill Proficiency* (labeled Skill) is comprised of the early literacy (LNF and LSF) and math (Numbers and Operations) academic achievement measures in the OKA, and is akin to acquisition skills in Sfard's metaphor.

I allowed the three latent factors comprising the OKA to correlate for three reasons. First, from a construct validity viewpoint, the items and measures comprise the entire OKA entry battery and are administered over a relatively short period of time. Second, adding covariance between the three components of the OKA has basis in theory given Sfard's argument that participation (behavioral skills) and acquisition (academic skill) are inherently related mechanisms of learning. Last, the correlations among latent factors in the CFA were low-moderate to strong further justifying correlation between the factors in the measurement model portion of the SEM. Of note is that in an SEM framework allows for the correlation between the entry behavioral and academic skill components of the OKA to be examined while simultaneously estimating their effect on students' spring achievement (Byrne, 2012; Kaplan, 2009; Kline, 2010).

Given the findings displayed in the bottom measurement model portion of the SEM (see Figure 1), the *Emergent Literacy* latent factor (labeled EL) serves as an endogenous outcome variable and is comprised of the three measures of spring early/emergent literacy, LSF, PSF, and WRF, with residual errors left uncorrelated. The spring *NCTM Math* benchmark (labeled Math) is included as a standalone continuous endogenous outcome variable. The *Emergent Literacy* latent factor and *NCTM Math* variable were specified as separate endogenous outcome variables because administrations of these benchmark assessments occurred over a broader range of time (about 3 months) than did administration of the OKA battery, and because the student population taking the emergent literacy and the math benchmarks was distinct as evidenced by the differences in *n*-size (Table 5). Furthermore, I documented a moderate positive relation between the latent factors ($r = .51$) using CFA, which I deemed strong enough to correlate the outcomes, but not strong enough to justify specifying a single spring achievement latent factor. Lastly, previous researchers found that such skills are distinct and positively related in young students (e.g., Gersten et al., 2005; Graney, Missall, Martínez, & Bergstrom, 2009; VanDerHeyden, Witt, Naquin, & Noell, 2001).

After the initial SEM model was specified and examined for consistency with EFA and CFA results, student-level demographic covariates (Sex, Nonwhite-Hispanic, Nonwhite-Non-Hispanic, Disability Status, Economic Disadvantage, and Limited English Proficiency) were added to the SEM as correlated predictors of both the *Emergent Literacy* latent factor and *NCTM Math* outcome variables. I removed non-significant demographic covariates in a single step. Finally, I compared nested models based on available fit indices, as well as the reasonableness of estimated parameters and effects, including their consistency with EFA and CFA results to arrive at a final SEM to answer Research Question 2.

***Model fit evaluation*.** Because typical goodness of fit indices used in covariance analyses techniques like SEM are not available when using MLR estimation with categorical outcome variables (Muthén & Muthén, 2015c), I compared adjacent-nested models using the AIC and BIC fit indices (Akaike, 1973; Schwartz, 1978), as well as a chi-square difference test based on loglikelihood values and scaling correction factors available with MLR estimation in Mplus 7.3. For the latter, significant values indicate a better fitting model (Muthén & Muthén, 2015a).

CHAPTER III

RESULTS

I first analyzed OKA data descriptively to ensure that data were sufficiently normal and reliable (see Tables 2, 3, and 5). On average, students entering public kindergarten named over 18 letters per minute, sounded nearly 7 letters per minute, and answered about 8 of the 16 math problems correctly (see Table 2). Standard deviations were greater than the means for the two early literacy measures (LNF and LSF). Students' total learning-related behavior rating scores averaged 54.85 (out of 75 possible) on the CBRS, 35.35 (out of 50 possible) for the self-regulation portion of the CBRS, and 19.51 (out of 25 possible) for items comprising the social-interpersonal portion. At the item level, students average ratings ranged from 3.00 to 3.91 across all 15 learning-related behavioral items, 3.00 to 3.71 for self-regulation items, and 3.89 to 3.92 for the social-interpersonal items (see Table 4).

Descriptive statistics for the OKA measures are shown in Tables B.1 and B.2 in Appendix B for demographic groups included as covariates in structural modeling. Of note was the disparity in kindergarten entry performance on the academic achievement measures (LNF, LSF and Numbers and Operations) for Nonwhite/Hispanic students, and for students identified as economically disadvantaged, as having a disability, and as having limited English proficiency—each of whom performed below their respective group peers. The disparity in performance was most substantial for the two early literacy measures, whereas differences in numeracy performance between groups were much smaller. For example, Nonwhite/Hispanic students averaged just fewer than 10 letter names per minute compared to white and Nonwhite/non-Hispanic students who averaged over 20 letters named. Students identified as having limited English proficiency averaged over 7 letter names and about 2 letter sounds per minute, whereas those identified as English proficient named and sounded about 21 and 8 letters, respectively. Between-group differences for total-rating averages of learning-related behaviors on the CBRS were much smaller across the demographic groups. Students identified as having a disability averaged about six-tenths of a point lower compared to their peers across all items on the behavioral rating scale—this was the largest behavioral-rating difference for any demographic group.

Descriptive statistics for the easyCBM spring benchmark measures are shown in Table 5, and for demographic covariates included in structural modeling in Table B.3 in Appendix B. On average, students in the spring sounded about 34 letters per minute, segmented over 40 phonemes per minute, read almost 15 words per minute, and answered about 36 of 45 math problems correctly. Standard deviations were quite large for the early/emergent literacy measures, with the mean and standard deviation nearly the same for the WRF measure. The largest differences in spring performance were between students identified as having a disability and of limited English proficiency, who performed below their respective group peers on each of the seasonal benchmark achievement measures.

Exploratory Factor Analysis: Examining the Factor Structure of the OKA

The internal consistency of the CBRS was sufficiently high (Cronbach's alpha = .96). The complete correlation matrix among items and measures in the OKA is displayed in Table A.1 in Appendix A. Overall, bivariate correlations ranged from low to high (0.13 to .90) across all items and measures. Bivariate correlations between the 15 learning-related behavioral items comprising the CBRS ranged from moderate to strong (.57 to .90) as well as for the entry academic achievement measures (LNF, LSF, and Numbers and Operations; .51 to .76) comprising the OKA.

The Kaiser-Meyer-Oklin value was appropriately high (.96) and Bartlett's Test of Sphericity was significant (χ^2 = 356628.01(153), p < .001). I specified and compared two exploratory models. The three-factor solution, in which learning-related behavioral items from the CBRS separated into two separate factors (*Self-regulation* and *Social-interpersonal*) and the academic achievement measures loaded to a single factor, better represented the observed OKA data in all respects. The chi-square difference test indicated that the three-factor solution fit significantly better than a two-factor solution (χ^2 = 20231.83(16), p < .001). Likewise, AIC and BIC criteria were both substantially lower for the three-factor solution (see Table 6).

Table 6

Model Fit Information Criteria for Two- and Three-Factor Solutions for OKA Battery

Information Criteria	Two-factor	Three-factor
AIC*	908972.51	885773.86
BIC**	909773.67	886701.95

Note. *Akaike (1973); **Schwartz (1978).

Measured variable (MV) separation in the three-factor solution was reasonable and interpretable compared to the two-factor solution. Communality values, representing the proportion of each MV variance explained by the three extracted factors are shown in Table 7. Values across all OKA items/measures ranged from .42 to .95, with values ranging from .63 to .87, .73 to .95, and .42 to .76 for the *Self-regulation*, *Social-interpersonal*, and *Academic Skill Proficiency* factors.

The rotated pattern and structure matrices, representing the linear combination of variables and the correlation between the items/measures and extracted factors, respectively, are also displayed in Table 7. Primary factor loadings and standard errors are bolded (Floyd & Widaman, 1995). Estimates are precise as indicated by the small standard errors across the extracted factors. The pattern of factor loadings is distinct. The *Self-regulation* factor was comprised of the following CBRS items: follows, completes, successfully, attempts, concentrates, responds, time, finds, errors, and returns. These ten items loaded strongly and differentially on the *Self-regulation* factor with minimal cross loading. The second factor, *Social-interpersonal*, consisted of the five remaining items from the CBRS: share, cooperative, turns, complies, and fuss. Three of these five items (share, cooperative, and turns) loaded strongly and differentially on the *Social-interpersonal* factor, whereas two items (complies and fuss) showed minor cross loading with the *Self-regulation* factor. The third factor, *Academic Skill Proficiency,* was comprised of the academic achievement measures from the OKA: LNF, LSF and Numbers and Operations. All three measures loaded differentially on the *Academic Skill Proficiency* factor, with the two early literacy measures loading quite stronger than Numbers and Operations. The correlation between *Self-regulation* and *Social-interpersonal* was strong (.70, $p < .05$), and between *Self-regulation* and *Academic Skill Proficiency* the correlation was moderate (.42, $p < .05$). The correlation between *Social-interpersonal* and *Academic Skill Proficiency* was very low (.05, $p < .05$).

Table 7

Communalities, Pattern and Structure Matrices for EFA Random Subsample for OKA Battery (n = 20,585).

Item	Comm-unalities	Pattern Matrix			Structure Matrix		
		Self-regulation	Social-interpersonal	Academic Skills	Self-regulation	Social-interpersonal	Academic Skills
follows	0.73	**0.94 (0.01)**	0.00	-0.37	**0.78**	0.64	0.02
completes	0.87	**1.00 (0.01)**	-0.11	0.01	**0.93**	0.59	0.42
successfully	0.85	**1.01 (0.01)**	-0.16	0.04	**0.91**	0.55	0.45
attempts	0.69	**0.87 (0.01)**	-0.07	0.03	**0.83**	0.53	0.39
concentrates	0.80	**0.92 (0.01)**	0.03	-0.11	**0.89**	0.66	0.28
responds	0.86	**0.93 (0.01)**	0.05	-0.11	**0.92**	0.69	0.28
time	0.80	**0.91 (0.01)**	0.02	-0.07	**0.89**	0.65	0.31
finds	0.81	**0.94 (0.01)**	-0.01	-0.08	**0.90**	0.64	0.32
errors	0.63	**0.87 (0.01)**	-0.15	0.05	**0.79**	0.46	0.40
returns	0.80	**0.90 (0.01)**	0.02	-0.06	**0.89**	0.65	0.32
share	0.93	0.00	**0.95 (0.00)**	0.09	0.70	**0.96**	0.13
cooperative	0.95	-0.01	**0.97 (0.00)**	0.10	0.71	**0.97**	0.14
turns	0.95	0.02	**0.96 (0.00)**	0.10	0.72	**0.97**	0.15
complies	0.80	0.30	**0.66 (0.01)**	-0.01	0.76	**0.87**	0.14
fuss	0.73	0.25	**0.66 (0.01)**	-0.01	0.71	**0.84**	0.13
LNF	0.76	0.03	0.03	**0.85 (0.01)**	0.41	0.10	**0.87**
LSF	0.68	-0.01	0.05	**0.83 (0.01)**	0.37	0.08	**0.83**
Math	0.42	0.13	-0.03	**0.59 (0.01)**	0.36	0.09	**0.64**

Note. OKA where: LNF = Letter Names Fluency, LSF = Letter Sounds Fluency, Math = Numbers and Operations, and item abbreviations for the CBRS behavioral rating segment. Primary factor loadings for the three extracted factors (*Self-regulation*, *Social-interpersonal*, and *Academic Skill Proficiency*) are bolded with standard errors shown in parentheses (Preacher & MacCallum, 2003).

Confirmatory Factor Analysis: Verifying the Factor Structure of the OKA and easyCBM Spring Benchmarks

I specified three unidimensional CFA models, which provided support for the factor structures extracted in the EFA (*Self-regulation*, *Social-interpersonal*, and *Academic Skill Proficiency*). Unstandardized and standardized parameter estimates for the unidimensional analyses are displayed by factor in Tables C.1, C.2, and C.3 in Appendix C. Of note is the strong and statistically significant loading of all CBRS items and academic achievement measures (LNF and LSF loaded more strongly than Numbers and Operations). Additionally, threshold values for items comprising the *Self-regulation* and *Social-interpersonal* latent factors (not reported for the unidimensional CFA), or the expected value of a given continuous latent factor at which an individual transitions from one rating-scale point to the subsequent/adjacent point (i.e., 1 to 2, 2 to 3, 3 to 4 and 4 to 5 on the ordinal 5-point scale of the CBRS) increased in a regular pattern from negative to positive, with the most extreme ratings typically reflecting the largest increase on the expected latent factor.

Results for the CFA concurrently estimating the factor loadings and interrelation of the three latent factors are presented in Table 8. Factor loadings, also interpreted as the correlation between each item/measure and the associated latent factor, were all significant ($p < .001$), strong and positive, mirroring EFA results. Standardized loadings for *Self-regulation* factor ranged from .81 to .94 for the first ten items on the CBRS (follows, completes, successfully, attempts, concentrates, responds, time, finds, errors, and returns). For the *Social-interpersonal* factor, correlations ranged from .88 to .98 for the remaining five items (share, cooperative, turns, complies, and fuss). Correlations ranged from .61 to .91 for the *Academic Skill Proficiency* latent factor (LNF, LSF and Math). Consistent with the EFA findings, the two early literacy measures (LNF and LSF) loaded most strongly on the achievement factor, with LNF explaining most of the variance in the factor, and math loading somewhat less strongly than both early literacy measures. Threshold values for categorical items comprising the *Self-regulation* and *Social-interpersonal* latent factors again increased in a regular pattern from negative to positive moving up the 5-point rating scale. The most extreme ratings generally reflected larger increases on the expected latent factor compared to more central/frequent values.

The correlations between the latent factors were all significant ($p < .001$) and of a similar magnitude and direction to those found in the EFA. The correlation between *Self-regulation* and *Social-interpersonal* was strong (.79, $SE = .00$, $p < .001$). The correlation between *Self-regulation* and *Academic Skill Proficiency* was moderate (.39, $SE = .01$ $p < .001$), and the correlation between *Social-interpersonal* and *Academic Skill Proficiency* was low (.20, $SE = .01$, $p < .05$).

Results for the CFA determining the structure of the spring *Emergent Literacy* latent factor (included as one of the two endogenous outcomes in the SEM) comprised of the early/emergent literacy benchmarks are presented in Table 9. All relations were significant, positive, and strong, with the LSF measure loading the most strongly compared to PSF and WRF. The correlation between *Emergent Literacy* latent factor and the continuous *NCTM Math* variable was moderate, $r = .51$, $SE = .03$, $p < .001$.

Table 8

Unstandardized and Standardized Loadings for CFA Random Subsample for the OKA Battery

	Self-regulation		Social-interpersonal		Academic Skills	
CBRS Item / Measure	Unstandardized	Standardized	Unstandardized	Standardized	Unstandardized	Standardized
follows	3.36 (0.04)	0.88				
completes	4.81 (0.07)	0.94				
successfully	4.14 (0.06)	0.92				
attempts	2.88 (0.04)	0.85				
concentrates	4.18 (0.05)	0.92				
responds	5.20 (0.07)	0.94				
time	4.03 (0.05)	0.91				
finds	4.14 (0.05)	0.92				
errors	2.48 (0.03)	0.81				
returns	4.04 (0.05)	0.91				
share			7.18 (0.15)	0.97		
cooperative			8.79 (0.24)	0.98		
turns			9.63 (0.28)	0.98		
complies			3.86 (0.05)	0.91		
fuss			3.36 (0.05)	0.88		
LNF					15.15 (0.10)	0.91
LSF					8.00 (0.08)	0.83
Math					1.93 (0.02)	0.61

Note. n = 20,585. CBRS items and academic achievement measures specified to load on a single factor (*Self-regulation, Social-interpersonal,* or *Academic Skills*) based on three-factor solution results in EFA. All parameter estimates significant, $p < .001$.

Table 9

Unstandardized and Standardized Loadings for easyCBM-matched Subsample for the easyCBM Spring Benchmarks

Spring measure	Unstandardized	Standardized
LSF	14.05 (0.28)	0.93
PSF	9.37 (0.27)	0.60
WRF	9.37 (0.23)	0.65

Note. easyCBM spring benchmarks where: LSF = Letter Sounds Fluency; PSF = Phoneme Segmenting Fluency; WRF = Word Reading Fluency. All parameter estimates, $p < .001$.

Structural Equation Modeling: Spring Early/Emergent Literacy and Math

I specified and compared three SEM to arrive at the final model. I have displayed unstandardized and standardized estimates and associated standard errors for all exogenous and endogenous variables in the final SEM in Table 10. I also have reported the same information for the initial two models in Tables D.1 and D.2 in Appendix D. Finally, I have included only the standardized relations between latent and observed variables included in the structural (path analysis) portion of the final SEM in Figure 1 (McDonald & Ho, 2002).

Incremental model fit indices (AIC and BIC) showed as student-level demographic covariates were added and removed when nonsignificant that model fit improved. The final SEM (Model 3, with Economic Disadvantage and Nonwhite-Hispanic removed from predicting spring *Emergent Literacy* and *NCTM Math*, and Sex removed from predicting *Emergent Literacy*) fit the observed data the best (Table 11).

Table 10

Unstandardized and Standardized Parameter Estimates for the Final SEM (Model 3)

Structural Model				
Factor		Factor	Unstandardized	Standardized
Spring Emergent Literacy	<--	Academic Skills	1.19 (0.04)	0.74
Spring Emergent Literacy	<--	Self-regulation	0.20 (0.04)	0.12
Spring Emergent Literacy	<--	Social-interpersonal	-0.09 (0.03)*	-0.06*
Spring Emergent Literacy	<--	Nonwhite/Non-Hispanic	0.27 (0.05)	0.17
Spring Emergent Literacy	<--	Disability	-0.58 (0.06)	-0.36
Spring Emergent Literacy	<--	LEP	0.19 (0.05)	0.12
Residual variance for (spring) Emergent Literacy			--	0.38
Spring NCTM Math	<--	Academic Skills	2.05 (0.10)	0.32
Spring NCTM Math	<--	Self-regulation	1.62 (0.18)	0.25
Spring NCTM Math	<--	Social-interpersonal	-0.66 (0.16)	-0.10
Spring NCTM Math	<--	Female	0.57 (0.17)*	0.09*
Spring NCTM Math	<--	Nonwhite/Non-Hispanic	-1.07 (0.28)	-0.17
Spring NCTM Math	<--	Disability	-2.35 (0.38)	-0.37
Spring NCTM Math	<--	LEP	-3.65 (0.45)	-0.57
Residual variance for (spring) NCTM Math			30.80 (1.20)	0.75
Measurement Model				
Spring benchmark		Factor	Unstandardized	Standardized
LSF	<--	Emergent Literacy	6.54 (0.22)	0.71
PSF	<--	Emergent Literacy	4.90 (0.22)	0.51
WRF	<--	Emergent Literacy	7.47 (0.13)	0.85
CBRS item		Factor	Unstandardized	Standardized
follows	<--	Self-regulation	3.39 (0.03)	0.88
completes	<--	Self-regulation	4.83 (0.05)	0.94
successfully	<--	Self-regulation	4.19 (0.04)	0.92
attempts	<--	Self-regulation	2.89 (0.03)	0.85
concentrates	<--	Self-regulation	4.21 (0.04)	0.92
responds	<--	Self-regulation	5.18 (0.05)	0.94
time	<--	Self-regulation	4.08 (0.04)	0.91
finds	<--	Self-regulation	4.23 (0.04)	0.92
errors	<--	Self-regulation	2.47 (0.02)	0.81
returns	<--	Self-regulation	4.08 (0.04)	0.91
CBRS item		Factor	Unstandardized	Standardized
share	<--	Social-interpersonal	7.19 (0.11)	0.97
cooperative	<--	Social-interpersonal	8.79 (0.17)	0.98
turns	<--	Social-interpersonal	9.75 (0.20)	0.98
complies	<--	Social-interpersonal	3.88 (0.04)	0.91
fuss	<--	Social-interpersonal	3.34 (0.03)	0.88
Entry measure		Factor	Unstandardized	Standardized
LNF	<--	Academic Skills	15.14 (0.07)	0.91
LSF	<--	Academic Skills	8.06 (0.06)	0.83
Math	<--	Academic Skills	1.96 (0.02)	0.62

Note. *$p < .01$; all others, $p < .001$. LSF = Letter Sounds Fluency; PSF = Phoneme Segmenting Fluency; WRF = Word Reading Fluency; LNF = Letter Names Fluency. Residual variances for all easyCBM achievement measures are shown in Figure 1.

The chi-square difference test indicated that adding demographic covariates significantly improved model fit, and that removing non-significant demographic covariates also significantly improved model fit.

Table 11

Model Fit Information Criteria for Specified SEM

Fit Criteria	Model 1	Model 2	Model 3
AIC	1967148.50	1966800.65	1966794.38
BIC	1968054.18	1967438.00	1967404.49
Chi-square	--	319.85*	3.47*

Note. Chi-square difference test statistics compare the adjacent/nested model, and are based on loglikelihood values and scaling correction factors available with MLR estimation in Mplus 7.3, in which significant values indicate a better fitting model (Muthén & Muthén, 2015a), *$p < .05$.

Factor loadings and correlations in the final SEM remained consistent with those found in EFA and CFA results, with all measures of the OKA and several student-level demographic covariates having significant effects on spring early/emergent literacy and math achievement (Table 10 and Figure 1). Standardized factor loadings (i.e., correlations) between the OKA *Self-regulation* latent factor and the specified CBRS indicators were strong, ranging from .81 to .94 for the ten teacher-rated items. Correlations between the OKA *Social-interpersonal* latent factor and its indicators were also strong, ranging from .88 to .98 for the five remaining CBRS items. Correlations between the OKA *Academic Skill Proficiency* and the LNF and LSF were very strong, .91 and .83, respectively, and strong .62 for the Numbers and Operations, indicating that the *Academic Skill Proficiency* and the three observed variables were strongly related, though the early literacy measures were most strongly associated with the latent factor. The residual variances of the LNF, LSF and Numbers and Operations indicators were .18, .31, and .62, respectively (Figure 1). Correlations between the *Emergent Literacy* endogenous outcome and the LSF and WRF indicators were .71 and .85, respectively, and .51 for the PSF indicator; therefore, the *Emergent Literacy* latent factor and the three observed variables were strongly related, though PSF less so relative to LSF and WRF. The residual variances of the LSF, PSF and WRF indicators were .49, .74, and .28, respectively.

For the OKA battery, the correlation between *Self-regulation* and *Social-interpersonal* was .79, indicating a strong positive relation between the two entry learning-related behavioral factors (Figure 1). The correlation between the *Self-regulation* and *Social-interpersonal* and the *Academic Skill Proficiency* factor was .39 and .20, respectively, indicating moderate and low positive relations between entry achievement and the two learning-related behavioral factors. For the endogenous outcomes, the correlation between *Emergent Literacy* and *NCTM Math* was .32, indicating a low-moderate relation between spring early/emergent literacy and math achievement.

Spring early/emergent literacy. For the final SEM, the reference group was white, male students identified as not having a disability and as being English proficient. Each one standard deviation increase in *Skill Proficiency* was associated with, on average, a .74 standard deviation increase in spring *Emergent Literacy*, holding all else constant (Table 10 and Figure 1). Students' entering *Self-regulation* and *Social-interpersonal* behaviors as rated by their teacher were associated with, on average, a .12

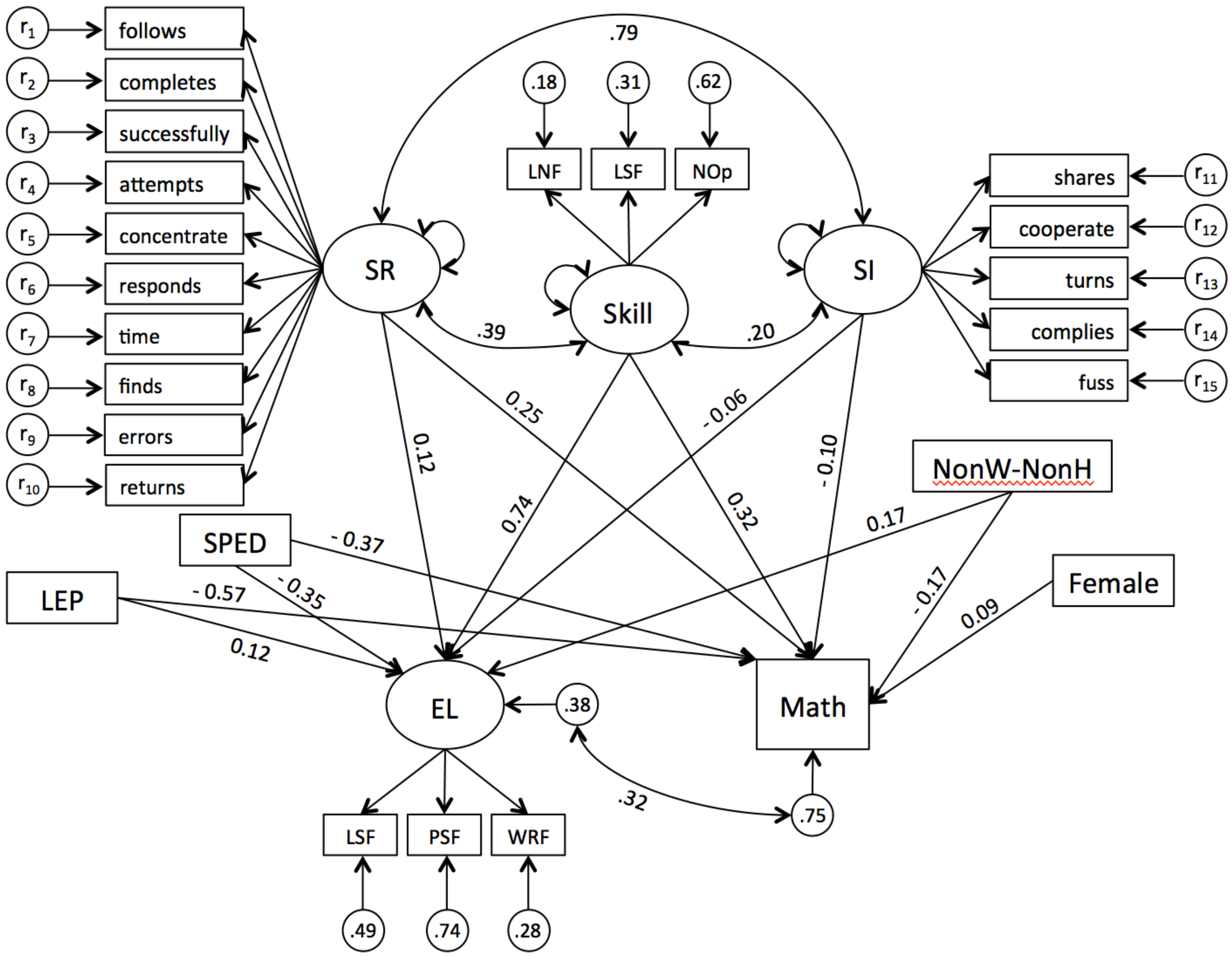

Figure 1. Standardized results for final SEM: (a) OKA exogenous predictors of *Self-regulation* (SR), *Social-interpersonal* (SI), and *Academic Skill Proficiency* (Skill)—comprised of Letter Names Fluency (LNF), Letter Sounds Fluency (LSF) and Numbers and Operations (NOp), and (b) easyCBM spring benchmark endogenous outcomes of *Emergent Literacy* (EL)—comprised of LSF = Letter Sounds Fluency, PSF = Phoneme Segmenting Fluency, and WRF = Word Reading Fluency, and *NCTM Math* (Math).

standard deviation increase and a -.06 decrease, respectively, in spring *Emergent Literacy*, controlling for all other predictors.

Standardized results for student-level demographic covariates are based on standardization of the endogenous literacy outcome. Nonwhite/Non-Hispanic students performed, on average, .17 standard deviations higher compared to white students in spring *Emergent Literacy*, holding all else constant. Being a student identified as having a disability was associated with a -.35 standard deviation decrease in spring *Emergent Literacy*, holding all else constant. Students identified as having limited English proficiency was associated with, on average, .12 standard deviations increase on spring *Emergent Literacy*, controlling for all else. The residual variance of the *Emergent Literacy* latent factor was .38, indicating that the final SEM explained a good proportion of variance (.62) in the *Emergent Literacy* endogenous outcome.

Spring math. For the reference group, each one standard deviation increase in students' *Skill Proficiency* was associated with, on average, a .32 standard deviation increase in spring *NCTM Math* performance, controlling for all else (Table 10 and Figure 1). Students' entry *Self-regulation* and *Social-interpersonal* behaviors were associated with, on average, a .25 standard deviation increase and a -.10 standard deviation decrease in spring *NCTM Math* scores, respectively, holding all else constant.

Female students scored, on average, .09 standard deviations higher than their male peers on spring *NCTM Math*, controlling for all else. Nonwhite/Non-Hispanic students' performed, on average, -.17 standard deviations lower than white students on spring *NCTM Math*., holding all else constant Being identified as having a disability was associated with a -.37 standard deviation decrease in spring *NCTM Math* achievement, holding all else constant. Students identified as having limited English proficiency performed, on average, -.57 standard deviations lower (equivalent to -3.65 points) than their counterparts identified as English proficient on spring *NCTM Math*, holding all else constant. The residual variance of *NCTM Math* was .75, indicating that the final SEM explained a small proportion of variance (.25) in the *NCTM Math* endogenous outcome.

CHAPTER IV

DISCUSSION

I had two main purposes for this dissertation. The first was to document and verify the underlying latent dimensions and interrelations of the skill components comprising Oregon's mandated kindergarten entrance assessment (OKA) using a statewide sample of students (Oregon Department of Education, 2013b). The second was to investigate the relation of these entry skill components to the level of early/emergent literacy and mathematics achievement measured in the spring of the same kindergarten school year while simultaneously accounting for student-level demographic characteristics.

This study adds to the previous research base in two important ways. First, this study sought to critically examine a state's model of kindergarten entry by investigating the individual and combination of skills measured by the assessment and the latent constructs underlying them (Brown & Moore, 2012; Fabrigar et al., 1999). It is the only known study to attempt to examine a state-mandated kindergarten entrance assessment in this manner. In this study, the results from 2013-2014 administration of the OKA provide a snapshot of specific learning-related behavioral and academic proficiency skills for a large statewide sample of kindergartners. I reasoned that if the underlying dimensions of the state's entry assessment could be modeled and confirmed, inferences about the learning-related behavioral and early academic skills that entering Oregon kindergartners possess as measured by the OKA could be drawn.

Secondly, this study extended the reach of the OKA by examining its association to achievement skills measured later in the same school year for a closely representative portion of the statewide sample. Using structural modeling techniques (Byrne, 2012; Kaplan, 2009; Kline, 2010), I was able to concurrently isolate and estimate the relations between students' entering latent skills and their spring early/emergent literacy and mathematics skills, while accounting for student demographic characteristics. Findings from these analyses offer an opportunity for additional insights (beyond preparedness for kindergarten) into the complex interplay of skills present over the initial year of public schooling in Oregon and that cut across pre-K-12 learning more generally. Although mechanisms are not explicitly outlined in ODE publications, one intention behind the implementation of the OKA is as a basis for decision-making. This intention is perhaps, in part, grounded in current federal policy and funding, which places importance on providing access to high quality early learning opportunities and on improving alignment across PK-12 public systems (U.S. Department of Education, 2013, 2014, 2015). Results from structural modeling may offer some insight into the utility of the OKA as such a guide.

Substantive Findings

Exploratory and confirmatory factor analyses appear to validate the state's model of kindergarten entry skills, which allows for inferences regarding Oregon students' entering skills as measured by the OKA. Distinct and parsimonious latent factors emerged from the analyses, a prerequisite for drawing appropriate test-based inferences based on student performance (Justice et al., 2005). I identified two learning-related behavioral skill factors, *Self-regulation* and *Social-interpersonal*. The distinct pattern of separation of CBRS items onto these two factors, comprising the Approaches to Learning measure on the OKA, matches what the CBRS purports to measure (Bronson et al., 1990) and the sub-scores that the state formally and publicly reports (Oregon Department of Education, 2013c, 2014). Additionally, the behavioral factors, which underlie learning-related skills found important in early learning contexts and beyond (e.g., Ladd et al., 1999; McClelland et al., 2006; McClelland et al., 2000), were strongly and positively related to one another as postulated theoretically (e.g., Bandura, 1991; Sfard, 1998) and as documented empirically in early classroom settings (e.g., Finn, 1993; McClelland & Morrison, 2003).

The third factor underlying the OKA, termed *Academic Skills Proficiency*, was comprised of the easyCBM progress-monitoring measures (LNF, LSF, and Numbers and Operations). Letter naming and sounding measures loaded most strongly on the latent proficiency skill factor, indicating the factor is predominantly typified by indicators of early literacy skill. Together, performance on these measures provides a snapshot of early literacy and numeracy abilities as students enter public schooling. The OKA thus measures skills that researchers have found interrelated and appropriate for students in early academic settings (e.g., Foegen et al., 2007; Gersten et al., 2012; Ritchey & Speece, 2006), and that are predictive of later achievement in kindergarten and beyond (e.g., Ritchey, 2008; Seethaler & Fuchs, 2011; Speece et al., 2004).

Importantly, the observed pattern and strength of the OKA factor structures and relations corroborated the findings of Tindal et al. (Manuscript submitted for publication), who documented essentially the same entry model using data from the 2012-2013 OKA pilot. The substantive finding that emerges from the exploratory and confirmatory analyses of the 2013-2014 OKA are based on a statewide sample of kindergarten students—providing evidence that the state's model of kindergarten entry skills likely generalizes to the broader Oregon kindergarten student population.

In addition to confirming the state's kindergarten entry model, structural modeling documented statistically significant and practically meaningful relations between some of students' entering skills and their spring early/emergent literacy and math achievement. The effects of the learning-related behavioral factors on spring early/emergent were statistically significant, and a curious negative relation was observed for social-interpersonal behaviors, similar to that found by Tindal et al. (Manuscript submitted for publication). However, given the small magnitude of the effects of learning-related behaviors estimated in this study, they may lack practical utility in informing instructional decision-making, a stated purpose behind implementation of the OKA (Oregon Department of Education, 2013c).

The small (and perhaps practically inconsequential) relation found between the learning-related behavioral factors and spring early/emergent literacy achievement runs counter to what many researchers have previously observed, including those who used the CBRS, though in smaller study samples. Similar learning-related skills have been predictive of early/emergent literacy and math achievement later in kindergarten and early elementary (McClelland et al., 2006), including when controlling for prior achievement (Ladd et al., 1999; McClelland & Morrison, 2003) and student-level demographic characteristics (Finn, 1993). However, some research appears to support the lack of influence of social learning-related behaviors on later achievement. Using data from the Early Childhood Longitudinal Study, Kindergarten dataset (ECLS-K), Claessens, Duncan, and Engel (2009) found virtually no impacts of socioemotional skills on fifth grade reading and mathematics achievement beyond that of students' entry performance, except for children's capacity to pay attention (a self-regulatory skill).

Two issues could be considered in terms of learning-related behavioral entry skills and the small effects related to later achievement observed in this study. First, teachers were rating students, very early, within about the first month of school. Despite the state's training parameters, it is quite possible that teachers did not have enough experience with their students or the CBRS that early in the year to rate their learning-related behaviors with the nuance needed for predicting later achievement. A lack of student familiarity with both their students and the behavioral rating scale may account for the small variance and negative skew across the distributions of CBRS items. For example, students averaged 3.89 to 3.92 (on the five-point rating scale) across the five CBRS items representing the social-interpersonal factor, with similar patterns (though slightly lower average scores) observed for those items representing the self-regulation factor. Perhaps including parents as raters of their students' learning-related behaviors would add some distinction to behavioral ratings and improve the utility of the ratings in terms of predicting later achievement skills. Research like this is already being conducted with the OKA in Central Point School District in southern Oregon, where both teacher and parent ratings of learning-related behaviors are being used to establish criteria for further student evaluation (Rowley, 2015).

Second, *executive functioning*, a more comprehensive set of skills that involve one's ability to plan, self-monitor, and self-manage (which includes self-regulation and additional skills), and *working memory*, the system by which we temporarily store, access and utilize information to carry out complex cognitive tasks like learning and reasoning, are not directly measured as part of the OKA battery and have been shown to be positively related to achievement. For example, Alloway and Alloway (2010) found that working memory at five years of age was the best predictor of literacy and numeracy skills 6 years later, above and beyond that accounted for by IQ in kindergarten. Similarly, Bull, Espy, and Wiebe (2008) found that higher levels of executive functioning was associated with "immediate head starts" in reading and math in preschool, and that these advantages were maintained across the first three years of schooling to age seven. Including a broader set of learning-related behaviors might better characterize entering kindergartners' learning-related skillsets and improve the utility of that portion of the OKA in predicting later academic performance.

Despite similar self-regulatory and social-interpersonal learning-related behavioral skills predicting proximal and distal literacy and mathematics achievement in previous research, it was by modeling the effects of the *Academic Proficiency Skill* factor that the strongest relation was found with spring early/emergent literacy achievement. This finding is perhaps not surprising given that early alphabetic skills (such as LNF and LSF, those dominating the entry academic skill latent factor of the OKA) have been strongly related to higher-order early/emergent literacy skills measured later in time, such as the indicators of the spring latent outcome in this study (LSF, PSF, and WRF; Cummings et al., 2011; Linklater et al., 2009; Ritchey, 2008; Ritchey & Speece, 2006). Controlling for the learning-related behaviors and student demographic characteristics, a one standard deviation increase in entering academic proficiency corresponded to nearly three-fourths of a standard deviation increase in spring early/emergent literacy performance. The strong effect of entering academic proficiency on spring early/emergent literacy achievement is noteworthy considering the hypersensitivity of letter naming skills (Francis et al., 1996; Paris, 2005) and possible floor effects of the letter sounding measure (Catts et al., 2009)—both of which have the potential to limit the predictive-concordant utility of the OKA, but do not seem to do so in this study.

The effects of kindergarten entry skills on spring math performance were smaller and less informative compared to how they related to spring early/emergent literacy. While the effect of entering *Academic Proficiency Skill* was again greater than the effects of either learning-related behavioral factor, the disparity between the effects was far smaller. On average, a one standard deviation increase in OKA *Academic Proficiency Skill* performance was associated with about a third of a standard deviation increase in spring math achievement, equivalent to just over a 2 points on the raw score scale for the NCTM Math assessment (out of a possible total score of 45). *Self-regulation* had a positive effect on spring math achievement skills, associated with about a quarter of a standard deviation increase in spring math scores (equivalent to roughly 1.62 points on the raw scale), while the small though curious negative relation between *Social-interpersonal* behavioral skills and spring achievement was once again observed with respect to math performance. In this study the OKA, especially the *Academic Proficiency Skill* portion of the battery, explained a large proportion of variation in spring early/emergent literacy performance, but offered little in the way of explaining spring mathematics achievement for the portion of the statewide kindergarten population in the analytic sample. This disparity in explained variance between spring early/emergent literacy and math, however, may have been a result of the OKA being comprised of two early literacy measure and just one short mathematics measure targeting a single a single numeracy skillset, whereas the spring assessment was a benchmark measure that included many more items targeting a broader range of math skillsets addressing the kindergarten NCTM Focal Point standards.

Examining the relational effects of student demographic characteristics on spring early/emergent literacy achievement revealed disparities even while controlling for prior (entering) academic proficiency and learning-related behavioral skills. Most prominent was the effect for students identified as having a disability, which was associated with over a third of a standard deviation decrease in spring emergent/early literacy achievement. This observed negative relation is noteworthy because students identified with disabilities averaged far lower scores on both early literacy measures on the OKA, and thus the negative relation can be viewed as the initial gap in performance growing for these students from fall to spring. The negative effect of disability, however, needs to be interpreted in light of the identification process in which students with more academically debilitating, low incidence conditions (e.g., intellectual disabilities) are typically identified at this age, whereas higher incidence disabilities (e.g., learning disability) are typically not identified until the later in the kindergarten school year, or further ahead in first and second grades.

Additional concern might be justifiable regarding students of Nonwhite/Hispanic decent and those who were economically disadvantaged. Though the effects of Nonwhite/Hispanic and Economic Disadvantage on spring early/emergent literacy were not statistically significant, on average, these students performed well below their respective peers on the two early literacy measures of the OKA. Thus, the gap in performance observed upon kindergarten entry appears to have not closed over the school year for either student demographic group. Conversely, identification as being limited in English language proficiency was associated with over a tenth of a standard deviation increase in spring early/emergent literacy achievement compared to their English language proficient peers, an important consideration given their lower initial performance on the early literacy measures of the OKA.

Disparities in spring math performance based on demographic characteristics were also present. The largest negative effects were associated with students identified as having disabilities and of limited English proficiency. These two groups, respectively, averaged over a third of a standard deviation and almost six-tenths of a standard deviation lower than their corresponding peers (equivalent to -2.35 and -3.65 points on the raw scale, respectively), controlling for all else. These negative relations are likely concerning in that average initial numeracy performance on the OKA was not substantially different for students in these two demographic groups—less than two points (out of 16 possible) for both groups (see Table B.1)—thus, a slight gap in entering early numeracy performance appears to have gotten larger over the kindergarten school year for students identified as having disabilities and of limited English proficiency in the sample.

Student-level demographic characteristics appear to affect spring early/emergent literacy and mathematics achievement even after controlling of kindergarten entry learning-related behavioral and academic proficiency skills for the portion of the statewide kindergarten population included in structural analyses. However, it is important to note that in the current study I analyzed only the main effects of demographic characteristics on spring literacy and math achievement. Such modeling might be too simplistic a view within the context of examining entering skills and their relation to later achievement. Specifically, demographic covariates might more appropriately be modeled as mediating/moderating effects in relation to entry skills and later achievement. Investigating whether (and how) the effects of different entry skills on later achievement change when demographic characteristics are accounted for in this manner is an important consideration. Given a stated intention behind implementation of the OKA is to identify achievement gaps among children in various demographic groups early on (Oregon Department of Education, 2013c), decision-makers would be wise to consider the combined influence of entry skills and demographic characteristics on indicators of early/emergent literacy and math achievement over the initial year of public schooling that are shown to be important for higher-order skill mastery later in elementary schooling.

Limitations and Future Research

The weak research design is a main limitation of this dissertation study. I analyzed extant data in this study that were not collected using any type of experimental or quasi-experimental design, and thus, causal inferences are not warranted. The academic achievement measures comprising the OKA and serving as spring achievement outcomes come from the same interim-formative assessment system (easyCBM) and were designed for interim-formative benchmark screening and progress-monitoring, typically as part of school improvement initiatives such as Response to Intervention (RTI; Alonzo et al., 2006). Inferences drawn about school preparedness based on performance on these measures for the state-mandated OKA should be made carefully (if at all) given such inferences fall outside of the intended purpose of the original measure development.

Specifically, caution should be taken when considering the relations of the underlying skillsets comprising the OKA, and in terms of their relation to early/emergent literacy and math achievement later in the kindergarten year. While the patterns of skill relations are certainly interpretable and may generalize across Oregon kindergartners in the full sample in this study, observed relations were for a single cohort and performing within a single year of schooling (2013-2014). Contextual concerns around implementation and administration of the OKA likely further limit inferences around skill relations. For example, contextual issues such as this being the first statewide administration of the OKA, teachers unfamiliarity with their students combined with rating the frequency of learning-related behaviors, along with the fact that teachers hand-entered data into a state-run website introduce error into the extant data analyzed in this study and is unaccounted for in statistical modeling. Whether the relations estimated here generalize over time across subsequent Oregon kindergarten cohorts and across kindergarten populations in other states are questions that need answered using more tightly controlled research designs.

Another limitation is that observed relations between students' entry skills and later spring achievement were drawn for a much smaller portion of the statewide sample (< 20% with respect to the spring early/emergent literacy measures and about 10% for math). Though initial performance and demographic characteristics were very similar between the full analytic sample and the easyCBM-matched subsample, important distinctions and questions remain. First, students who are assessed using interim benchmark screening and progress-monitoring measures, the students who make up the easyCBM-matched subsample, are likely substantively different from general state population of kindergartners who may or may not take part in such interim-formative assessment in RTI improvement contexts. Outside of their demographic characteristics and their performance on the initial statewide field test of the OKA, we know very little about the full sample of Oregon kindergarten students and how they compare to those in the easyCBM-matched subsample. Second, while initial performance on the OKA and demographics characteristics were remarkably similar between the statewide sample and matched subsample, especially given the lack of sampling design (see Tables 1 and 3), the easyCBM-matched subsample had about 6% more white students and 6% fewer Nonwhite/Hispanic students. How this difference in demographic make-up between the full sample and matched subsample affects inferences about the broader statewide kindergarten population was not examined in this study.

Further, whether the observed skill relations and performance disparities between demographic groups persist over time and across future kindergarten cohorts in Oregon should be examined in future research if the OKA is to be used as a reliable and valid indicator of performance gaps or as a guide to decision-making. Structural modeling in the current study assumes invariance of factor structures and pattern loading across demographic groups. This assumption may not be tenable for either the full analytic sample or across time in future Oregon kindergarten populations. Previous work that examined the easyCBM math benchmark measures, albeit it for test forms in grades 3-5, suggested non-invariance of factor structures across educational setting, language and ethnic groups (Nese, Anderson, & Tindal, 2010, May). I found no published studies that evaluated the invariance of the CBRS across demographic populations. Future work should specifically examine the sensitivity of measures comprising the OKA and their invariance across key demographic groups.

The lack of specificity at the student level is another important limitation of my findings, especially in light of two of the stated purposes for the OKA to identify observable entering performance gaps and guide instructional decision-making (Oregon Department of Education, 2013a, 2013c). Though the purpose of this study was to investigate the underlying constructs of a state-mandated kindergarten entry assessment and their relation to important early/emergent literacy and math achievement at the end of the school year, caution should be taken when considering the inferences I'm drawing in with respect to the average kindergarten performance across the state and those that can/should be drawn for more localized settings—within districts, schools, classrooms—and certainly at the individual student level.

With the state's entry model confirmed, the capacity to draw inferences around the level of kindergartners' entering learning-related behavioral and academic proficiency skills for the statewide population of kindergartners appears defensible—though capacity should clearly not be confused with assuming the appropriateness and consequences of such inferences and associated decisions for the general and especially more specific student populations (Kane, 1992; Messick, 1994). For example, characterizing how individual students or groups of demographically similar students perform on the OKA is a distinctly different issue from determining if such students or groups of students are "prepared" for kindergarten or need further screening and/or intervention following school entrance because they are deemed "at-risk" under some criteria. As Gersten et al. (1995) argued with respect to using math scores as a means to determine risk and guide instruction, test-based inferences should demonstrate adequate consequential validity in the context in which they are operationalized.

Lastly, questions remain about kindergarten students' math performance and how it is impacted by the skills with which students enter school. My findings do not explain the variation in kindergarten students' end-of-year math performance on a measure that assesses broad skillsets (as outlined by the NCTM Focal Point Standards—Numbers and Operations, Geometry, and Measurement), and that are shown to be important for later math success (Gersten et al., 2005; VanDerHeyden et al., 2006). Future research should address whether the learning-related behavioral and academic proficiency skills measured in the OKA are related to later math achievement in subsequent cohorts. Further, the ODE might consider whether assessing additional mathematical skillsets (beyond simply entering early numeracy skills) should be incorporated into the entry assessment battery to improve its utility in predicting later math performance. For example, Clements et al. (2008) developed and validated the REMA to assess a broad range of math skills across five content domains they argued were developmentally appropriate for preschool and kindergarten children. Weiland et al. (2012) later developed and validated a condensed form more conductive to classroom use. Perhaps measuring a more diverse array of math-related skillsets using an assessment that is similar in length to the Numbers and Operations test form currently a part of the OKA would more appropriately target entering kindergarten students' math skills and serve as a basis for identifying achievement/performance gaps and influencing decision-making. Further yet, including additional learning-related behavioral skillsets like executive functioning and working memory, shown to be related to later math achievement (Alloway & Alloway, 2010; Bull et al., 2008), might also improve the utility of the OKA in this regard. An important consideration would be that measuring additional math skills and/or learning-related behaviors as part of the state's entry assessment battery would have to be balanced against the need to keep the entry assessment battery short and facile for kindergarten teachers who are just getting to know their students personally while determining their instructional and learning-related needs.

Key questions arise from the discussion of highlighted limitations. For example, can data from the OKA be used to consistently identify and address achievement (or performance) gaps between (demographic) groups of students over time? How should OKA data influence decision-making at the state level and more localized levels like districts, schools, and classrooms? Should classroom teachers use OKA data to guide their instruction, in what manner should this be done—and further, what is the impact of doing so? How do results from the OKA impact the way in which publicly funded PK-12 learning systems are aligned and improved in Oregon? Such questions are critical for policymakers and researchers to investigate and for educators to carefully consider if the OKA is to validly serve its stated purposes and positively impact students, teachers and education policy and learning systems in Oregon.

Conclusions

The findings in this study provide evidence that Oregon has implemented a kindergarten assessment that assesses learning-related behavioral and early academic proficiency skills that are both interrelated and that have statistically significant and practically important associations to key indicators of literacy and math achievement later in the kindergarten school year. However, these findings are not without consequential limitations. In this light, data from the OKA must be used and interpreted in a manner consistent with its purpose of providing a snapshot of entering students' skills and as a means to identify early achievement gaps and guide instructional-decision-making, with the outcomes of doing so carefully documented and examined. While the OKA assesses entry skills that a wide base of research have shown present in young children and important to later school success, this finding must be considered within the intersection of issues surrounding education measurement, assessment, instruction, and policy.

Recently, the federal government released a summary report on preschool in America (U.S. Department of Education, 2015), finding that almost 60% of four-year-olds nationwide and nearly 80% in Oregon did not attend a public preschool program (National Institute for Early Education Research, 2013). Despite being the overall largest and fastest-growing minority, Latino students in the U.S. had the lowest public preschool participation rate (about 40%) compared to African American (50%) and White (53%) children, while economically disadvantaged students also lag behind their peers in preschool participation. Citing longitudinal research that documented, on average, minority students and children falling below the federal poverty line entered kindergarten with lower (pre)reading and mathematics skills compared to their peers (Mulligan, Hastedt, & McCarroll, 2012), the U.S. Department of Education argued that high quality early learning opportunities need to be accessible to all children.

Similar disparities in performance were observed upon entry into kindergarten and at the end of the kindergarten school year for specific demographic groups of Oregon kindergartners in this study—disparities perhaps linked to the lack of (quality) early learning opportunities and preschool for many children in the state. Devising innovative ways to provide high quality early learning access for all children, so that students enter public schools with the acquisition (achievement) and participatory (learning-related behavioral) skills they need for future success, is a challenge nationwide and especially in Oregon where the majority of young children do not attend preschool. Although the OKA might provide information to help meet this challenge, given the entry assessment provides information about the skills kindergartners' possess very early on in their public schooling experience, it is but one piece of a complex puzzle of characterizing "readiness for schooling", and for improving and aligning educational systems across the early learning and K-12 continuum and beyond.

APPENDIX A

COMPLETE EFA CORRELATION MATRIX

Table A.1

Bivariate Correlation Matrix for the 2013-2014 OKA (Full Analytic Sample)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	LNF	LSF	NOp
1	-	.73	.68	.60	.78	.80	.71	.70	.57	.70	.67	.68	.68	.73	.67	.23	.20	.21
2	-	-	.85	.74	.75	.79	.75	.78	.68	.74	.56	.57	.58	.61	.57	.34	.29	.31
3	-	-	-	.75	.72	.75	.75	.75	.68	.72	.54	.54	.55	.58	.54	.36	.31	.32
4	-	-	-	-	.66	.69	.69	.69	.63	.67	.51	.53	.53	.57	.53	.31	.27	.28
5	-	-	-	-	-	.83	.76	.73	.64	.76	.62	.62	.62	.65	.61	.27	.23	.24
6	-	-	-	-	-	-	.77	.77	.64	.77	.63	.65	.65	.70	.64	.27	.23	.25
7	-	-	-	-	-	-	-	.76	.66	.74	.61	.61	.62	.64	.60	.27	.23	.25
8	-	-	-	-	-	-	-	-	.67	.76	.60	.60	.61	.63	.60	.27	.23	.25
9	-	-	-	-	-	-	-	-	-	.70	.47	.47	.48	.48	.47	.34	.30	.30
10	-	-	-	-	-	-	-	-	-	-	.61	.61	.62	.63	.60	.27	.23	.25
11	-	-	-	-	-	-	-	-	-	-	-	.88	.88	.75	.73	.15	.13	.13
12	-	-	-	-	-	-	-	-	-	-	-	-	.90	.77	.73	.15	.13	.14
13	-	-	-	-	-	-	-	-	-	-	-	-	-	.78	.74	.16	.14	.14
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.81	.17	.14	.16
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.15	.13	.14
LNF	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.76	.56
LSF	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.51
NOp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note. All correlations (pairwise deletion) are significant at the $p < 0.01$ level (2-tailed). Overall, bivariate correlations ranged from .13 to .90. Bivariate correlations between items comprising the *Self-regulation* latent factor (items 1-10) ranged from .57 to .85, from .73 to .90 for the *Social-interpersonal* latent factor (items 11-15), and from .51 to .76 for the three academic achievement measures (LNF, LSF and Numbers and Operations [NOp]) comprising the *Academic Skill Proficiency* latent factor of the OKA.

APPENDIX B

DESCRIPTIVE STATISTICS BY DEMOGRAPHIC GROUP

Table B.1

Descriptive Statistics by Demographics for OKA Total Scores: Achievement Measures

	LNF		LSF		Math	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sex						
Female	19.25	16.49	7.07	9.79	7.99	3.05
Male	17.77	16.88	6.39	9.61	8.05	3.27
Race/Ethnicity						
White	20.94	16.40	7.79	10.00	8.41	3.12
Nonwhite/Hispanic	9.81	13.34	2.92	6.37	6.83	2.85
Nonwhite/Non-Hispanic	22.04	18.16	8.20	11.24	8.28	3.38
Economic Disadvantage						
Not Disadvantaged	24.22	16.97	9.81	11.10	8.87	3.20
Disadvantaged	13.40	14.70	3.95	7.2	7.27	2.93
Disability						
No Disability	19.15	16.80	7.07	9.90	8.14	3.15
Disability	12.11	14.34	3.35	6.67	6.86	3.08
LEP						
Not limited	20.94	16.68	7.79	10.18	8.36	3.15
Limited	7.34	11.51	1.78	4.62	6.46	2.74

Note. Full analytic kindergarten sample where: LNF = Letter Names Fluency, LSF = Letter Sounds Fluency, and Math = NCTM Numbers and Operations, pairwise deletion.

Table B.2

Descriptive Statistics by Demographics for OKA Total Scores: Approaches to Learning

	Self-regulation		Social-interpersonal		Total Score	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sex						
Female	37.39	7.87	20.38	3.99	57.77	11.13
Male	33.40	8.66	18.67	4.54	52.07	12.41
Race/Ethnicity						
White	35.61	8.54	19.51	4.40	55.12	12.21
Nonwhite/Hispanic	34.51	8.35	19.51	4.25	54.02	11.83
Nonwhite/Non-Hispanic	35.63	8.62	19.46	4.40	55.09	12.29
Economic Disadvantage						
Not Disadvantaged	36.87	8.29	20.06	4.22	56.93	11.78
Disadvantaged	34.02	8.49	19.02	4.43	53.04	12.17
Disability						
No Disability	35.96	8.17	19.76	4.19	55.72	11.61
Disability	29.46	9.50	17.05	5.23	46.51	13.93
Limited English Proficiency						
Not limited	35.62	8.53	19.53	4.39	55.15	12.19
Limited	34.15	8.37	19.38	4.26	53.53	11.85

Note. Full analytic kindergarten sample where average scores are given for summative item totals for the two learning-related behavioral factors extracted in EFA and for the total, pairwise deletion.

Table B.3

Descriptive Statistics by Demographics for Spring easyCBM Benchmarks

	LSF		PSF		WRF		Math	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sex								
Female	34.51	14.73	41.68	15.12	15.44	14.42	36.51	6.25
Male	32.67	14.93	38.67	15.74	14.29	14.25	36.23	6.60
Race/Ethnicity								
White	33.61	14.29	40.96	15.29	14.80	13.84	36.87	5.98
Nonwhite/ Hispanic	31.00	15.37	36.93	15.57	10.46	10.04	33.76	7.93
Nonwhite/ Non-Hispanic	36.11	16.10	40.19	15.96	19.53	18.15	35.98	6.64
Econ Disadv								
Not Disadv	35.56	14.26	41.53	15.30	18.50	16.45	37.65	5.77
Disadv	31.84	15.15	38.98	15.60	11.63	11.24	35.24	6.77
Disability								
No Disability	34.56	14.43	41.45	14.70	15.56	14.55	36.72	6.18
Disability	24.55	15.71	28.30	17.71	8.36	10.27	32.56	7.81
LEP								
Not limited	33.85	14.57	40.85	15.37	15.43	14.72	36.77	6.09
Limited	31.67	16.62	35.21	15.67	10.73	10.44	31.27	8.25

Note. Full analytic kindergarten sample where average total scores are given for spring easyCBM benchmarks: LSF = Letter Sounds Fluency, PSF = Phoneme Segmenting Fluency, WRF = Word Reading Fluency, Math = NCTM Math, and LEP = Limited English Proficiency, pairwise deletion.

APPENDIX C

UNIDIMENSIONAL CFA RESULTS

Table C.1

Unstandardized and Standardized Factor Loadings for the OKA: Self-regulation Factor

Item	Unstandardized	Standardized
successfully	3.24 (0.04)	0.87
completes	4.95 (0.07)	0.94
time	4.25 (0.06)	0.92
responds	2.89 (0.04)	0.85
attempts	4.13 (0.05)	0.92
errors	5.08 (0.07)	0.94
returns	3.99 (0.05)	0.91
concentrates	4.14 (0.05)	0.92
finds	2.51 (0.03)	0.81
follows	4.00 (0.05)	0.91

Note. n = 20,585. Items specified loading on a single factor (*Self-regulation*) based on three-factor solution in EFA. All parameter estimates, $p < .001$.

Table C.2

Unstandardized and Standardized Factor Loadings for the OKA: Social-interpersonal Factor

Item	Unstandardized	Standardized
share	7.17 (0.16)	0.97
turns	8.84 (0.26)	0.98
cooperative	9.75 (0.31)	0.98
complies	3.69 (0.05)	0.90
fuss	3.26 (0.05)	0.87

Note. n = 20,585. Items specified to load on a single factor (*Self-regulation*) based on three-factor solution in EFA. All parameter estimates, $p < .001$. $\chi^2 = 10593.74(3010)$, $p < .001$.

Table C.3

Unstandardized and Standardized Factor Loadings for the OKA: Academic Skill Proficiency Factor

Item	Unstandardized	Standardized
LNF	15.20 (0.11)	0.91
LSF	7.99 (0.08)	0.83
Math (Numbers and Operations)	1.91 (0.02)	0.61

Note. n = 20,585. Items specified to load on a single factor (*Self-regulation*) based on three-factor solution in EFA. All parameter estimates, $p < .001$.

APPENDIX D

PRELIMINARY SEM RESULTS

Table D.1

Unstandardized and Standardized Parameter Estimates for the Initial SEM (Model 1)

Factor		Factor	Unstandardized	Standardized
Spring NCTM Math	<--	Academic Skills	2.02 (0.10)	0.36
Spring NCTM Math	<--	Self-regulation	1.63 (0.18)	0.25
Spring NCTM Math	<--	Social-interpersonal	-0.67 (0.16)	-0.11
Residual variance for (spring) NCTM Math			32.23 (1.28)	0.78
Spring Emergent Literacy	<--	Academic Skills	1.19 (0.04)	0.74
Spring Emergent Literacy	<--	Self-regulation	0.20 (0.04)	0.14
Spring Emergent Literacy	<--	Social-interpersonal	-0.09 (0.03)*	-0.05*
Residual variance for (spring) Emergent Literacy			--	0.38
Spring benchmark		Factor	Unstandardized	Standardized
LSF	<--	Emergent Literacy	6.59 (0.23)	0.71
PSF	<--	Emergent Literacy	4.94 (0.23)	0.51
WRF	<--	Emergent Literacy	7.47 (0.13)	0.85
CBRS item		Factor	Unstandardized	Standardized
follows	<--	Self-regulation	3.39 (0.03)	0.88
completes	<--	Self-regulation	4.83 (0.05)	0.94
successfully	<--	Self-regulation	4.19 (0.04)	0.92
attempts	<--	Self-regulation	2.89 (0.03)	0.85
concentrates	<--	Self-regulation	4.21 (0.04)	0.92
responds	<--	Self-regulation	5.18 (0.05)	0.94
time	<--	Self-regulation	4.08 (0.04)	0.91
finds	<--	Self-regulation	4.23 (0.04)	0.92
errors	<--	Self-regulation	2.47 (0.02)	0.81
returns	<--	Self-regulation	4.08 (0.04)	0.91
CBRS item		Factor	Unstandardized	Standardized
share	<--	Social-interpersonal	7.19 (0.11)	0.97
cooperative	<--	Social-interpersonal	8.79 (0.17)	0.98
turns	<--	Social-interpersonal	9.75 (0.20)	0.98
complies	<--	Social-interpersonal	3.88 (0.04)	0.91
fuss	<--	Social-interpersonal	3.34 (0.03)	0.88
Entry measure		Factor	Unstandardized	Standardized
LNF	<--	Academic Skills	15.14 (0.07)	0.91
LSF	<--	Academic Skills	8.06 (0.06)	0.83
Math	<--	Academic Skills	1.96 (0.02)	0.62

Note. **p* < .01; all others *p* < .001. *Self-regulation* - *Social-interpersonal, r* = .79; *Academic Skills* - *Self-regulation, r* = .39; *Academic Skills* - *Social-interpersonal*, *r* = .20; and *Emergent Literacy* - *NCTM Math*, *r* = .32.

Table D.2

Unstandardized and Standardized Parameter Estimates for the SEM with All Demographic Covariates Included (Model 2)

Factor		Factor	Unstandardized	Standardized
Spring Emergent Literacy	<--	Academic Skills	1.18 (0.04)	0.73
Spring Emergent Literacy	<--	Self-regulation	0.20 (0.04)	0.12
Spring Emergent Literacy	<--	Social-interpersonal	-0.09 (0.03)*	-0.06*
Spring Emergent Literacy	<--	Female	0.02 (0.03)†	0.01†
Spring Emergent Literacy	<--	Nonwhite-Hispanic	-0.05 (0.05)†	-0.03†
Spring Emergent Literacy	<--	Nonwhite/Non-Hispanic	0.26 (0.05)	0.16
Spring Emergent Literacy	<--	Disability	-0.59 (0.06)	-0.36
Spring Emergent Literacy	<--	Economic Disadvantage	-0.01 (0.04)†	-0.00†
Spring Emergent Literacy	<--	LEP	0.22 (0.06)	0.14
Residual variance for (spring) Emergent Literacy			--	0.38
Spring NCTM Math	<--	Academic Skills	2.02 (0.10)	0.31
Spring NCTM Math	<--	Self-regulation	1.62 (0.18)	0.25
Spring NCTM Math	<--	Social-interpersonal	-0.67 (0.16)	-0.10
Spring NCTM Math	<--	Female	0.58 (0.18)*	0.09*
Spring NCTM Math	<--	Nonwhite-Hispanic	-0.41 (0.37)†	-0.06†
Spring NCTM Math	<--	Nonwhite/Non-Hispanic	-1.14 (0.28)	-0.18
Spring NCTM Math	<--	Disability	-2.36 (0.38)	-0.37
Spring Emergent Literacy	<--	Economic Disadvantage	-0.16 (0.18)†	-0.03†
Spring NCTM Math	<--	LEP	-3.33 (0.54)	-0.52
Residual variance for (spring) NCTM Math			30.80 (1.19)	0.75
Spring benchmark		Factor	Unstandardized	Standardized
LSF	<--	Emergent Literacy	6.54 (0.22)	0.71
PSF	<--	Emergent Literacy	4.90 (0.22)	0.51
WRF	<--	Emergent Literacy	7.47 (0.13)	0.85
CBRS item		Factor	Unstandardized	Standardized
follows	<--	Self-regulation	3.39 (0.03)	0.88
completes	<--	Self-regulation	4.83 (0.05)	0.94
successfully	<--	Self-regulation	4.19 (0.04)	0.92
attempts	<--	Self-regulation	2.89 (0.03)	0.85
concentrates	<--	Self-regulation	4.21 (0.04)	0.92
responds	<--	Self-regulation	5.18 (0.05)	0.94
time	<--	Self-regulation	4.08 (0.04)	0.91
finds	<--	Self-regulation	4.23 (0.04)	0.92
errors	<--	Self-regulation	2.47 (0.02)	0.81
returns	<--	Self-regulation	4.08 (0.04)	0.91
CBRS item		Factor	Unstandardized	Standardized
share	<--	Social-interpersonal	7.19 (0.11)	0.97
cooperative	<--	Social-interpersonal	8.79 (0.17)	0.98
turns	<--	Social-interpersonal	9.75 (0.20)	0.98
complies	<--	Social-interpersonal	3.88 (0.04)	0.91
fuss	<--	Social-interpersonal	3.34 (0.03)	0.88
Entry measure		Factor	Unstandardized	Standardized
LNF	<--	Academic Skills	15.14 (0.07)	0.91
LSF	<--	Academic Skills	8.06 (0.06)	0.83
Math	<--	Academic Skills	1.96 (0.02)	0.62

Note. $^{\dagger}p > .05$; $*p < .01$; all others $p < .001$. *Self-regulation - Social-interpersonal, r =* .79; *Academic Skills - Self-regulation, r* = .39; *Academic Skills - Social-interpersonal*, *r* = .20; and *Emergent Literacy* - *NCTM Math*, *r* = .32.

REFERENCES CITED

Abt Associates. (1988). Evaluation of Project Giant Step (Technical Progress Report No. 4). Cambridge, MA: Author.

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*(106), 20-29.

Alonzo, J., Lai, C. F., & Tindal, G. (2009a). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 2 (Technical Report No. 0920). Eugene, OR: Behavioral Research and Teaching: University of Oregon.

Alonzo, J., Lai, C. F., & Tindal, G. (2009b). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 3 (Technical Report No. 0902). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Lai, C. F., & Tindal, G. (2009c). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 4 (Technical Report No. 0903). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2007a). The development of early literacy measures for use in a progress monitoring assessment system: Letter names, letter sounds and phoneme segmenting (Technical Report No. 39). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2007b). The development of word and passage reading fluency measures in a progress monitoring assessment system (Technical Report No. 40). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009a). Alternate form and test-retest reliability of easyCBM reading measures (Technical Report No. 0906). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009b). Alternate form and test-tetest reliability of easyCBM reading measures (Technical Report No. 0906). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009c). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 1 (Technical Report No. 0919). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009d). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Kindergarten (Techincal Report No. 0921). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). easyCBM online progress monitoring assessment system.

Anderson, D., Lai, C. F., Nese, J. F. T., Park, B. J., Sáez, L., Jamgochian, E. M., . . . Tindal, G. (2010). Technical adequacy of the easyCBM primary-level mathematics measures (Grades K-2), 2009-2010 Version (Technical Report No. 1006). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Park, B. J., Lai, C. F., Alonzo, J., & Tindal, G. (2012). An Examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 1 (Technical Report No. 1216). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243-277). Mahwah, NJ: Lawrence Erlbaum.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ:: Prentice-Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248-287. doi: 10.1016/0749-5978(91)90022-L

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology, 35*, 61–79.

Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling: A Multidisciplinary Journal, 7*, 461-483. doi: 10.1207/S15328007SEM0703_6

Brigance, A. (1985). Brigance Preschool Screen. North Billerica, MA: Curriculum Associates.

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

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

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*(2), 253–282.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361-379): The Guilford Press.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*(3), 205-228.

Byrne, B. M. (2012). *Structural equation modeling with Mplus*. New York: Routledge.

Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M. S., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities, 42*(2), 163-176. doi: 10.1177/0022219408326219

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review, 28*, 415-427.

Clements, D. H., Sarama, J. H., & Lieu, X. H. (2008). Development of a measure of early mathematics achievement using the Rasch model: The Research-based Early Maths Assessment. *Educational Psychology, 28*(4), 457-482. doi: 10.1080/01443410701777272

Clements, D. H., Wilson, D. C., & Sarama, J. H. (2004). Young children's composition of geometric figures: A learning trajectory. *Mathematical Thinking and Learning, 6*(2), 163-184. doi: 10.1207/s15327833mtl0602_5

Connors-Tadros, L. (2014). Fast fact: Information and resources on developing state policy on Kindergarten Entry Assessment (KEA). New Brunswick, NJ: Center on Enhancing Early Learning Outcomes.

Cooper, D., & Farran, D. C. (1988). Behavioral risk in kindergarten. *Early Childhood Research Quarterly, 3*(1), 1-19. doi: 10.1016/0885-2006(88)90026-9

CTB McGraw-Hill. (2010). TerraNova3 Complete Battery. Monterey, CA.

Cummings, K., Kaminski, R., Good, R., & O'Neal, M. (2011). Assessing phonemic awareness in preschool and kindergarten: Development and initial validation of first sound fluency. *Assessment for Effective Intervention, 36*(2), 94-106.

Downing, S. M. (2006a). Selected-response item formats in test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 287-301). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Downing, S. M. (2006b). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3-25). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299. doi: 10.1037/1082-989X.4.3.272

Finn, J. (1993). School engagement and students at risk. Washington, DC: U.S. Department of Education, National Center for Educational Statistics.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical sssessment instruments. *Psychological Assessment, 7*(3), 286-299. doi: 10.1037/1040-3590.7.3.286

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *Journal of Special Education, 41*, 121-139. doi: 10.1177/00224669070410020101

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3-17. doi: 10.1037/0022-0663.88.1.3

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71*(1), 7-21.

Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education, 33*(18), 18-28. doi: 10.1177/002246699903300102

Gersten, R., Clarke, B., Jordan, N. C., Newman-Gonchar, R., Haymond, K., & Wilkins, C. (2012). Universal screening in mathematics for the primary grades: Beginnings of a research base. *Exceptional Children, 78*(4), 423-445.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*, 293-304.

Gersten, R., Keating, T., & Irvin, L. (1995). The burden of proof: Validity as improvement of instructional practice. *Exceptional Children, 61*(6), 510-519.

Ginsburg, H. P., & Baroody, A. J. (2005). Test of Early Mathematics Ability (3rd ed.). Austin, TX: Pro-ed.

Good, R. H., & Kaminski, R. A. (2003). Dynamic Indicators of Basic Early Literacy Skills administration and scoring guide (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Graney, S. B., Missall, K. N., Martínez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School Psychology, 47*(2), 121-142.

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

Haladyna, T. (2004). *Developing and validating multiple-choice test items* (Third ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. doi: 10.1080/10705519909540118

IBM Corp. (2012). IBM SPSS Statistics for Mac Version 21.0. Armonk, NY: IBM Corp.

Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2012). The alignment of the easyCBM grades K-2 math measures to the Common Core State Standards (Technical Report No. 1228). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Jackson, D. L., Gillaspy Jr., J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14*(1), 6-23.

Johnstone, C., Altman, J., & Thurlow, M. (2006). A state guide to the development of universally designed assessments. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Justice, L. M., Invernizzi, M., Geller, K., Sullivan, A. K., & Welsch, J. (2005). Descriptive-developmental performance of at-risk preschoolers on early literary tasks. *Reading Psychology, 26*(1), 1-25. doi: 10.1080/02702710490897509

Kane, M. (1992). An argument-based approach to validity. *Psychological Bulletin, 112*, 527-535. doi: 10.1037/0033-2909.112.3.527

Kaplan, D. (2009). *Structural equation modeling: Foundations and extensions* (2nd ed.). Thousand Oaks, CA: Sage.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

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

Lai, C. F., Alonzo, J., & Tindal, G. (2009a). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 5 (Technical Report No. 0901). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009b). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 6 (Technical Report No. 0907). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009c). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 7 (Technical Report No. 0907). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009d). The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 8 (Technical Report No. 0904). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2013). easyCBM reading criterion related validity evidence: Grades K-1 (Technical Report No. 1309). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Nese, J. F. T., Jamgochian, E. M., Kamata, A., Anderson, D., Park, B. J., . . . Tindal, G. (2010). Technical adequacy of the easyCBM primary-level reading measures (Grades K-1), 2009-2010 version. (Technical Report No. 1003). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lembke, E., & Foegen, A. (2009a). Identifying early numeracy indicators for kindergarten and first-grade students. *Learning Disabilities Research & Practice, 24*(1), 12-20.

Lembke, E., & Foegen, A. (2009b). Indentifying early numeracy indicators for kindergarten and first-grade students. *Learning Disabilities Research & Practice, 24*(1), 12-20.

Linklater, D. L., O'Connor, R. E., & Palardy, G. J. (2009). Kindergarten literacy assessment of English Only and English language learner students: an examination of the predictive validity of three phonemic awareness measures. *Journal of School Psychology, 47*(6), 369–394. doi: 10.1016/j.jsp.2009.08.001

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

McClelland, M. M., & Morrison, F. J. (2003). The emergence of learning-related social skills in preschool children. *Early Childhood Research Quarterly, 18*, 206-224. doi: 10.1016/S0885-2006(03)00026-7

McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at-risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly, 15*(3), 307–329. doi: 10.1016/S0885-2006(00)00069-7

McConnell, S., McEvoy, M., & Priest, J. (2002). "Growing" measures for monitoring progress in early childhood education: A research and development process for individual growth and development. *Assessment for Effective Intervention, 27*(4), 3-14. doi: 10.1177/073724770202700402

McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64-82. doi: 10.1037//1082-989X.7.1.64

McGuinn, P. (2012). Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy, 26*(1), 136-159. doi: 10.1177/0895904811425911

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23. doi: 10.3102/0013189X023002013

Mulligan, G. M., Hastedt, S., & McCarroll, J. C. (2012). First-time kindergartners in 2010-11: First findings from the kindergarten rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11. Washington, DC: National Center for Education Statistics.

Muthén, L. K., & Muthén, B. O. (1998-2012). Mplus user's guide (Seventh ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2015a). Chi-Square difference testing using the Satorra-Bentler scaled chi-square. Retrieved March 7, 2015, from http://www.statmodel.com/chidiff.shtml

Muthén, L. K., & Muthén, B. O. (2015b). Mplus Discussion: Choice of estimator. Retrieved March 7, 2015, from http://www.statmodel.com/discussion/messages/23/54.html

Muthén, L. K., & Muthén, B. O. (2015c). Mplus Discussion: Fit indexes. Retrieved March 7, 2015, from http://www.statmodel.com/discussion/messages/22/72

National Institute for Early Education Research. (2013). 2013 State Preschool Yearbook. Washington, DC.

National Institutes of Child Health and Human Development. (2000). Report of National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction *Report of the Subgroups*. Washington, DC.

Nese, J. F. T., Anderson, D., & Tindal, G. (2010, May). *The invariance of the easyCBM mathematics measures across educational setting, language, and ethnic groups*. Paper presented at the National Council on Measurement in Education (NCME) Annual Meeting, Denver, CO.

Nese, J. F. T., Lai, C. F., Anderson, D., Park, B. J., Tindal, G., & Alonzo, J. (2010). The Alignment of easyCBM Math Measures to Curriculum Standards (Technical Report No. 1002). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110 § 115, 1425 Stat. (2002).

Nurss, J. R., & McGauvran, M. E. (1986). *The Metropolitan Readiness Tests*. New York: Harcourt Brace Jovanovich.

Executive Numbered Memo: 010-2012-13 - OAR 581-022-2130 - Kindergarten Assessment (2013a).

Oregon Department of Education. (2013b). Kindergarten Assessment 2013 ‐ 2014 [Press release]. Retrieved from http://www.ode.state.or.us/go/assessmentupdate

Oregon Department of Education. (2013c). Test Adminstration Manual 2013-2014: Appendix L - Kindergarten Assessment. Salem, OR: Office of Assessment and Information Services.

Oregon Department of Education. (2014). *2013-2014 Statewide Kindergarten Assessment Results*. Retrieved from: http://www.ode.state.or.us/search/page/?=3908

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184-202. doi: 10.1598/RRQ.40.2.3

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's Electric Factor Analysis Machine. *Understanding Statistics, 2*(1), 13-43. doi: 10.1207/S15328031US0201_02

Ritchey, K. D. (2008). Assessing letter sound knowledge: A comparison of letter sound fluency and nonsense word fluency. *Exceptional Children, 74*(4), 487-506.

Ritchey, K. D., & Speece, D. L. (2006). From letter names to word reading: The nascent role of sublexical fluency. *Contemporary Educational Psychology, 31*, 301-327. doi: doi:10.1016/j.cedpsych.2005.10.001

Rowley, B. (2015). *Kindergarten Assessment: Analysis of the Child Behavioral Rating Scale (CBRS).* University of Oregon, Eugene, OR.

Sáez, L., Irvin, P. S., Alonzo, J., & Tindal, G. (2012). Phoneme segmenting alignment with the Common Core Foundational Skills Standard Two: Grades K-1 (Technical Report No. 1227). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Seethaler, P., & Fuchs, L. (2011). Using curriculum-based measurement to monitor kindergarteners' mathematics development. *Assessment for Effective Intervention, 36*(4), 219-229. doi: DOI: 10.1177/1534508411413566

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4-13.

Speece, D. L., Ritchey, K. D., Cooper, D., Roth, F., & Schatschneider, C. (2004). Growth in early reading skills from kindergarten to third grade. *Contemporary Educational Psychology, 29*, 312-332. doi: doi:10.1016/j.cedpsych.2003.07.001

SPSS Inc. (2010). SPSS for Macintosh License Agreement. Chicago, IL: SPSS Inc.

The White House & Office of the Press Secretary. (2013). Fact sheet President Obama’s plan for early education for all Americans [Press release]. Retrieved from http://www.whitehouse.gov/the-press-office/2013/02/13/fact-sheet-president-obama-s-plan-early-education-all-americans

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved December 1, 2012, from: http://www.cehd.umn.edu/nceo/OnlinePubs/Synthesis44.html.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education (International Scholarly Research Network), 29*. doi: 10.1155/2013/958530

Tindal, G., Irvin, P. S., & Nese, J. F. T. (April 2013). *Learning to Read: A Review of Research on Growth in Reading Skills*. Paper presented at the National Council on Measurement in Education, SanFrancisco, CA.

Tindal, G., Irvin, P. S., & Nese, J. F. T. (Manuscript submitted for publication). Preliminary evidence for a state's kindergarten entry skill assessment.

U.S. Census Bureau. (2013). Table 1: Enrollment status of the population 3 years old and over, by sex, age, race, Hispanic origin, foreign born, and boreign-born parentage: October 2013 *Microsoft Excel*. Washington, DC: U.S. Department of Commerce.

U.S. Department of Education. (2002). *Reading First*. Washington, DC: Office of Elementary and Secondary Education.

U.S. Department of Education. (2008). A Nation Accountable: Twenty-five Years After A Nation at Risk. Washington, DC.

U.S. Department of Education. (2013). U.S. Department of Education awards more than $15.1 million in Enhanced Assessment Grants to develop or improve kindergarten entry assessments [Press release]. Retrieved from http://www.ed.gov/news/press-releases/us-department-education-awards-more-151-million-enhanced-assessment-grants-devel

U.S. Department of Education. (2014). *Early learning: America's middle class promise begins early*. Washington, DC: U.S. Department of Education.

U.S. Department of Education. (2015). A matter of equity: Preschool in America. Washington, DC: U.S. Department of Education.

VanDerHeyden, A. M. (2011). Technical adequacy of response to intervention decisions *Council fo Exceptional Children, 77*(3), 335-350.

VanDerHeyden, A. M., Broussard, C., & Cooley, A. (2006). Further development of measures of early math performance for preschoolers. *Journal of School Psychology, 44*(6), 533-553. doi: dx.doi.org/10.1016/j.jsp.2006.07.003

VanDerHeyden, A. M., Broussard, C., Fabre, M., Stanley, J., Legendre, J., & Creppell, R. (2004). Development and validation of curriculum-based measures of math performance for preschool children. *Journal of Early Intervention, 27*, 27-41. doi: 10.1177/105381510402700103

VanDerHeyden, A. M., Witt, J. C., Naquin, G., & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. *School Psychology Review, 30*, 363-382.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). The development of reading-related phonological processing abilities. *Developmental Psychology, 30*(1), 73–87.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). Comprehensive Test of Phonological Processing. Austin, TX:: Pro-Ed.

Weiland, C., Wolfe, C. B., Hurwitz, M. D., Clements, D. H., Sarama, J. H., & Yoshikawa, H. (2012). Early mathematics assessment: Validation of the short form of a prekindergarten and kindergarten mathematics measure. *Educational Psychology, 32*(3), 311-333. doi: 10.1080/01443410.2011.654190

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56-75). Newbury Park, CA: Sage.

The White House Summit on Early Childhood Education. (2014). Washington, DC.

Wray, K., Lai, C. F., Saez, L., Alonzo, J., & Tindal, G. (2014). easyCBM kindergarten beginning reading measures: Alternate form reliability and criterian validity with the SAT-10 (Technical report in preparation). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*, 370.