VALIDATION OF A MATHEMATICS READINESS ASSESSMENT

A Thesis

Presented to the faculty of the
Department of Graduate and Professional Studies in Education
California State University, Sacramento

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF ARTS

in

Child Development

by

Jennifer Leigh Travis

SPRING 2016
VALIDATION OF A MATHEMATICS READINESS ASSESSMENT

A Thesis

by

Jennifer Leigh Travis

Approved by:

Dr. Sheri E. Hembree Committee Chair

Dr. Kristen Alexander Second Reader

April 28, 2016 Date

iii
Student: Jennifer Leigh Travis

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

Susan Heredia, Ph.D., Chair

Date: 4.26.16

Department of Graduate and Professional Studies in Education
Abstract

of

VALIDATION OF A MATHEMATICS READINESS ASSESSMENT

by

Jennifer Travis

The current study was a validation of a newly developed measure, the Early Numeracy Assessment (ENA). The ENA was created to be used in early education settings to test foundational numeracy skills that may support kindergarten academic success. The ENA was designed to aid efforts to evaluate readiness programs working to minimize the achievement gap for at-risk students, by measuring informal numeracy skills (Hemphill, Vanneman, Rahme, 2011; Geoffroy, Côté, Giguère, Dionne, Zelazo, Tremblay, & Séguin, 2010; Magnuson, Meyers, Ruhm, & Waldfogel, 2004; Purpura & Lonigan, 2013). The participants were 99 children, aged 3-5 years, attending a state preschool program. The ENA was administered using a manual developed for this research. Participants were asked questions that measured one-to-one correspondence, relative quantities, mathematic operations, patterning, rote counting, and identifying one-digit numerals. The results indicate significant internal consistency and inter-rater reliability for the ENA. Convergent validity was established based on the correlations between the ENA and the DRDP-PS 2010. There were positive relationships between the ENA and a Kindergarten readiness assessment without significant correlations. With
the ENA age was not a predictor of achievement, supporting the need to consider informal skills of young children in regards to academic achievement, and not focus solely on age. With current kindergarten programs restricting and measuring school readiness based on age alone, an accurate understanding of development and ability is not taken into account. This finding is supports the need to move away from the age based measures of school readiness towards more environmental and ecological measures of school readiness.

Dr. Sheri Hembree

4/28/16

Date
Completing this thesis would have never happened without the guidance of my professors, help from friends, and support from my family.

My sincere gratitude is extended to my sponsor Dr. Sheri Hembree for the continuous support of this research, for her patience, motivation, enthusiasm, and immense knowledge. Her guidance helped me in all the time, research, and writing of this thesis.

I would like to thank Dr. Kristen Alexander as my second reader for her encouragement, insightful comments, and questions.

My sincere thanks also goes to Dr. Amber Gonzalez, for offering me opportunities to complete and present research, leading me and working on diverse exciting projects.

I recognize my fellow cohort members, I cannot imagine going through this process without your encouragement and friendship.

Last but not the least, I would like to thank my friends and family for their support and believing in me along the way. I am who I am because of the love and care of my family: my parents, grandparents, sisters and Skylar. I love you guys!
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>3</td>
</tr>
<tr>
<td>Significance of Study</td>
<td>3</td>
</tr>
<tr>
<td>Methods</td>
<td>7</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>9</td>
</tr>
<tr>
<td>Limitations</td>
<td>10</td>
</tr>
<tr>
<td>Organization of Study</td>
<td>11</td>
</tr>
<tr>
<td>2. REVIEW OF THE LITERATURE</td>
<td>12</td>
</tr>
<tr>
<td>School Readiness</td>
<td>13</td>
</tr>
<tr>
<td>School Readiness Assessments</td>
<td>19</td>
</tr>
<tr>
<td>Numeracy Skills</td>
<td>30</td>
</tr>
<tr>
<td>Developing and Testing Measures</td>
<td>36</td>
</tr>
<tr>
<td>3. METHODS</td>
<td>43</td>
</tr>
<tr>
<td>Development of Measure</td>
<td>43</td>
</tr>
<tr>
<td>Administration of the Instrument</td>
<td>49</td>
</tr>
</tbody>
</table>

viii
LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bridge between Early Numeracy Assessment, Desired Results Developmental Profile for Preschool 2010, and California Kindergarten Common Core Math Standards</td>
<td>24</td>
</tr>
<tr>
<td>2. Research Supported Key Early Mathematics Skills</td>
<td>29</td>
</tr>
<tr>
<td>3. Participant Demographics</td>
<td>50</td>
</tr>
<tr>
<td>4. Early Numeracy Assessment (ENA) Descriptive Statistics</td>
<td>57</td>
</tr>
<tr>
<td>5. Internal Consistency of ENA</td>
<td>59</td>
</tr>
<tr>
<td>6. DRDP-PS (2010) and ENA Correlations</td>
<td>61</td>
</tr>
<tr>
<td>7. Kindergarten Sheet and ENA Correlations</td>
<td>62</td>
</tr>
</tbody>
</table>
INTRODUCTION

Statement of the Problem

When children enter into K-12 education with a limited knowledge base, research indicates they may remain academically behind their peers through 4th grade, creating an achievement gap (Geoffroy, Côté, Giguère, Dionne, Zelazo, Tremblay, & ... Séguin, 2010; Hemphill, Vanneman, Rahmen, 2011; Magnuson, Meyers, Ruhm, & Waldfogel, 2004; McKinsey & Company, 2009). Once a child falls behind academically, he/she is more likely to experience lower wages, increased health problems, and greater rates of incarceration and high school dropout affecting the nation as a whole by propagating a decreased workforce and labor market (Knapp & Woolverton, 1995; Lee, 2002; McKinsey & Company, 2009).

When an achievement gap exists in mathematics achievement, life-long social and developmental abilities, juvenile arrests for violent and nonviolent offenses, and grade retention and special education placement can be a result (Reynolds, Temple, Robertson, & Mann, 2001). These result effects are perpetuated onto future generations in that socioeconomic status (SES) predicts future academic achievement (Hackman & Farah, 2009; Knapp & Woolverton, 1995; McKinsey and Company, 2009; Reardon, 2011). An achievement gap in math skills, has a more significant effect than that of gap of reading skills and therefore important to consider (Baglici, Codding, & Tryon, 2010; Duncan,
Theoretically, if young children are entering kindergarten more ready for learning, then the achievement gap will lessen. Research indicates that when children participate in high quality early education consisting of developmentally appropriate practices and ongoing observational measurement of students' development, they are better prepared for kindergarten, sometimes even despite risk factors at home (Bauchmüller, Görtz, & Rasmussen, 2014; Côté, Mongeau, Japel, Xu, Séguin, & Tremblay, 2013; Early, Maxwell, Burchinal, Bender, Ebanks, Henry, & ... Zill, 2007; Schuyler Center, 2012; Schweinhart, 2013).

Creating and funding programs that increase readiness is one way in which policy makers and educators can address the achievement gap. However, there is a need for efficient, early, valid assessments of formal academic skills and concepts to measure the quality of the programs. The effectiveness of curriculum and practices can be measured with valid and reliable assessments. There are many different types of assessment styles currently used in early childhood education (ECE) classrooms including; standardized formal assessments, informal observational assessments, and performance assessments (Biddle, García-Nevárez, Henderson, Valero-Keérick, 2014; Leary, 2004; Wortham, 2005).

However, such assessments may not be developmentally appropriate (Wortham, 2005), may not account for cultural and language differences (Kusserow, 2012; Leary, 2004; Wortham, 2005), are often untested or are administered incorrectly (Biddle et al.,
The goal of this research was to create and validate an efficient and easy-to-use tool, the Early Numeracy Assessment (ENA). This tool may conceivably address the need for an effective measure of pre-academic skills associated with school mathematics readiness, and aid in the development and evaluation of readiness programs designed to narrow the achievement gap.

**Purpose of the Study**

The purpose of this study was to develop and administer the ENA in preschool settings to determine if it was reliable and valid to measure early numeracy skills predictive and supportive of kindergarten academic success. The ENA differs from typical classroom assessment by being rooted in, and supported by, current developmental research as well as linked to state standards. To determine if the ENA is effective and applicable in the classroom it was tested for accuracy and consistency of measuring early numeracy skills.

**Significance of the Study**

**Mathematics Readiness**

School readiness consists of the proficiencies of a child upon entering school. Skills contributing to academic success include: physical development, social and
emotional development, language development, methods of learning, and overall
cognition (Commodari, 2013; Snow, 2006). The concept of school readiness includes the
fundamentals of reading, writing, mathematics skills, emotional and impulse control, and
coopération skills (Commodari, 2013).

The goal for using the ENA and other similar assessments in early childhood
education settings is to measure academic achievement and/or school readiness. Current
practices in measuring school readiness include a type of screening measure to determine
if a child is academically prepared to enter kindergarten (Snow, 2006). However, these
school readiness screening tests raise concern due to lack of established validity in
predicting a child's success in kindergarten (Snow, 2006).

The measure developed in the current study specifically considers mathematics
readiness of preschool aged children from 3 to 5 years old. The ENA measures informal
mathematics skills that research indicates have a long term effect on later formal
mathematics skills (Purpura et al., 2013; Reynolds et al., 2001). Informal math skills are
typically fostered in homes with parents and through experiences that are often
spontaneous and unconventional that may come up through play such as conversations,
or sorting toys or objects (Purpura et al., 2013). Formal math skills are skills taught
explicitly such as patterns or addition and subtraction; these skills require a foundation of
informal math skills (Purpura et al., 2013). Children from low SES homes are at risk for
not meeting necessary informal academic skills due to decreased experiences that would
otherwise foster these unprompted learning interactions with parents such as limited time,
high levels of environmental factors, or lack of knowledge of child development and
learning (Bauchmüller, Götz, & Rasmussen, 2014; Côté, Mongeau, Japel, Xu, Séguin, & Tremblay, 2013; Schweinhart, 2013).

To address the need of families with children that are not able to foster these skills at home, ECE classrooms or child care may be an option. However, these programs may be expensive. Therefore, for families with limited resources programs such as Head Start and state preschool programs have been funded (Johnson, Martin, & Brooks-Gunn, 2013). These programs have been subsidized to catch low socioeconomic families that would typically be having these difficulties and better prepare their children for kindergarten and academic success. Hence, the target population for the testing of the ENA were children attending a state preschool program.

**Early Numeracy**

The focus of the current measure is early numeracy. Early numeracy is the ability to organize and classify quantities and objects based on characteristics and amounts (California Department of Education, 2010; Claeens & Engel, 2013; Purpura et al., 2013). Purpura et al. (2013) found that early numeral knowledge is the foundation upon which future mathematics is built. The ENA was developed to assess key early numeracy skills, one-to-one correspondence, rote counting, patterning, classification, mathematical operations, identifying one-digit numerals, in a valid and reliable way.

In ECE classrooms, intentional teaching of more complex concepts is a typical practice (Linder, Powers-Costello, & Stegelin, 2011; Varol & Farran, 2006). It is generally inferred that informal math skills were taught at home and include counting.
structure (i.e. counting sets, counting forward and backward, subitizing, estimation), relationship between quantities or numbers, and composing and decomposing sets (Purpura & Lonigan, 2013). Based on this inference informal math skills are unnecessary to teach in the classroom. Still, the intentional teaching of formal concepts in preschool is in need of an assessment that has established reliability and validity as well.

Current assessments may not address the informal knowledge that children may not have gained at home as a baseline for future learning. Without these informal foundational skills, children are at risk of falling behind academically. Without an accurate measurement tool, the cause for this gap may go unidentified. Assessment of early numeracy skills is necessary due to the long term stability of mathematic skills over other academic domains in relation to academic achievement (Siegler, Duncan, Davis-Kean, Duckworth, Claesens, Engel, Susperreguy, & Chen, 2012).

The Need for Effective Measurement of Readiness

Many measures of readiness, including mathematics readiness, are not tested rigorously (Scott-Little & Niemeyer, 2001). Scott-Little and Niemeyer (2001) compiled and checked the validity of 39 assessment tools used for children in kindergarten. Of the 39 assessments, three were used to measure kindergarten readiness. Of those three, two were tested and considered valid based on content, construct, concurrent, and predictive validity (Scott-Little & Niemeyer, 2001). Based on these findings and similar research, there is a need for more reliable and valid assessments for early readiness (Kilday & Kinzie, 2008; Scheeringa & Haslett, 2010; Wortham, 2005).
Further, Neuman and Roskos (2005) emphasized the importance of using standards supported by current research and that are clearly stated and understood not only by early childhood educators but by the general public. ENA items were developed using current early education and Kindergarten standards and current research on numeracy development. In addition to measuring individual student performance and to develop a better understanding of individual performance in a classroom, effectiveness of the readiness program must be evaluated. Measuring content knowledge of students shapes the quality of the educational practices of the classroom by improving the experiences and skills gained by the students. Using tools like the ENA can assess mathematics readiness programs and can be used for program evaluation, curriculum and lesson planning, as well as individual readiness.

Methods

The purpose of the current study was to test how accurately the ENA measures numeracy development of preschool children by establishing its validity and reliability with a quantitative approach. Reliability is the consistency of the test (Leary, 2004; Wortham, 2005), whereas validity refers to how accurately the assessment measures the proposed constructs (Leary, 2004; Wortham, 2005). The current study sought to establish reliability by measuring internal consistency, and inter-rater reliability, as well as concurrent and predictive validity.
Measure Development

The ENA was developed based on principles rooted in developmental research that support the ENA’s ability to assess formal and informal numeracy skills. The developmental research was aligned with the California Kindergarten Common Core (CKCC). The alignment occurred in an attempt to make the ENA predictive of academic achievement in Kindergarten.

Once the ENA was developed, it was piloted with a smaller group of participants from the State Preschool program. The pilot testing provided an opportunity to reorganize the structure of the ENA, change the language used, reevaluate the materials used, and test the Spanish translation. Piloting of the tool also allowed for the administrator to work with a small group of participants to determine how they responded to the prompts and questions. Changes were made after the piloting to streamline the assessment, which resulted in the measure consisting of fewer items than the original, improved Spanish translation, and revised language for the questions asked by the administrator.

Administration of the Measure

Next, the measure was administered to 99 preschoolers between the ages of three and five years attending public state preschool centers in Northern California. The state preschool program admits families based solely on financial need, and all participants were from low-SES households. The program’s annual income ceiling that qualifies the families for the program ranges from $3,000-$5,900 for a family of two to 12 individuals. The results of the administered assessment were compared to the participants’
performance on the preschool teacher’s completion of the Desired Results Developmental Profile for Preschoolers 2010 (DRDP-PS2010) and to Kindergarten performance data after a semester in the program. The DRDP-PS2010 measurement of similarly defined concepts; number sense of quantity and counting, number sense of mathematical operations, classification, and patterning were compared to those of the ENA. The kindergarten data measured counting by ones, identifying numbers 0-30, one-to-one correspondence, comparing sets of objects, and adding numbers. These skills were then compared to the ENA to calculate reliability and validity.

**Definition of Terms**

Several terms are used in the following study. *Achievement Gap* refers to a situation in which one group of students performs less well than another group of students on measures of academic success such as standardized tests, grades, high school completion rates, and college enrollment and completion rates (Reardon, 2013). The term *School Readiness*, the enhancement of domain specific skills that contributing to overall competencies of school achievement was also used (Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov, & ... Japel, 2007). More specifically *Mathematics Readiness*, the ability of an individual to master certain standardized mathematics skills in a given context was discussed (Lee, Autry, Fox, & Williams, 2008), along with *Numeracy*, an understanding of the number system including the ability to solve number
problems in a variety of contexts in relation to school readiness (Australian Mathematics Teacher, 2012).

Terms used in describing the methodologies of this study were also used. First,\textit{ Validity} refers to the truth or accuracy of a statement or concept (Cosby & Bates, 2012; Leary, 2004). \textit{Construct Validity} refers to the accuracy of how well a tool measures what it is meant to measure, of findings based on statistical data (i.e., test scores; Teglasi, Nebbergall, & Newman, 2012). \textit{Concurrent Validity} denotes how well measurements on one test correspond to those on a similar test, of the ways that groups of individuals vary on an assessment in predictable ways (Cosby & Bates, 2012; Leary, 2004). The third type of validity discussed is \textit{Predictive Validity}, or the ability of a tool to predict future abilities or outcomes (Cosby & Bates, 2012).

Second were terms associated with \textit{Reliability}, or the consistency of a measure (Cosby & Bates, 2012). \textit{Interrater Agreement} is a type of reliability that is determined by the agreement of the results of an assessment between two test observers or results raters (Cosby & Bates, 2012). \textit{Internal consistency} is also discussed as the occurrence of multiple items of similar constructs producing similar results (Cosby & Bates, 2012).

\textbf{Limitations}

The current study was conducted with a sample that was not widely generalizable due to lack of diversity when compared with the overall population of children in Northern California aged 3.5 years. Although the research participants did represent
different languages and cultures, the sample was limited to lower income students. The differing languages and cultures represented in the sample also provided a limitation as well. The researcher did not speak any language other than English fluently, and was not a substitute for a native Spanish speaker and therefore may have affected the interaction with participants. As for the other languages presented by the participants, their home languages and potential cultural variances were not addressed, and should be considered in future research.

**Organization of Study**

Chapter 1 has served as an introduction to the research study and has described the purpose of the study and an overview of the methods employed. Chapter 2 is a review of literature related to school readiness and how school readiness is measured, indicating the need for early assessment tools that are both reliable and valid so that knowledge is accurately measured. Chapter 2 also includes information on the ENA and how the development and subsequent research can be used in early classroom settings to measure numeracy knowledge in order to identify children at risk of performing poorly academically increasing the achievement gap. Chapter 3 describes the methods that were used to gather the data that was used to test the validity and reliability of the ENA. Chapter 4 provides the results of tests of reliability and validity of the measure. Chapter 5 provides a discussion of the results presented in Chapter 4, as well recommendations for future directions for research and measure development for early numeracy.
Chapter 2

REVIEW OF THE LITERATURE

Over 500,000 children enter the California school system through kindergarten enrollment each year (California Department of Education, 2014). With increased rates of kindergarten and transitional kindergarten enrollment, educators and researchers have increasingly focused on the construct of “school readiness.” Although there is no consensus on a definition of “school readiness,” broadly defined, the term refers to the abilities and competencies a child has upon entering school that predict of later achievement (Espinoza & García, 2012; Snow, 2006; Wesley & Buysse, 2003).

The goal for the current study was to develop and validate a measure for determining one aspect of kindergarten readiness; readiness for mathematics. Developmental research was used to establish predictors of academic achievement in mathematics. These predictors were incorporated into a numeracy assessment aligned with current ECE and kindergarten standards and developmental practices. The assessment tool was then administered to a group of preschool children and tested for validity and reliability.

The following is a review of relevant research on school readiness, early numeracy, and mathematics readiness. In order to identify current issues and trends in early education school readiness was discussed. School readiness was then reviewed through the lens of early assessment. Developing a deeper understanding of early assessment trends and issues, which support the need of early education assessments, was
reviewed. The results indicating the need for increased statistical analysis to determine reliability and validity of current tools was considered presenting the need for the development of the Early Numeracy Assessment (ENA).

School Readiness

The concept of school readiness has become a driving force to the pedagogical shift of Kindergarten. Kindergarten was originally an introduction to schooling; a socializing agent for young children (Moravčík & Feeney, 2012; Wesley & Buyssse, 2003). Kindergarten curriculum is now more likely to include academic topics and standards and be accompanied by standardized testing. These standards and tests mean that children entering kindergarten need to be prepared to “hit the ground running” academically. For most children this means preschool experience, and research supports that preschool is beneficial for later academic achievement (Barnett, 1998; Hill, Gormley, & Adelstein, 2015; Schweinhart, & Weikart, 1998).

Readiness is a dynamic concept that is difficult to define because of varying perspectives on development and the purpose of early education. Various pedagogical philosophies exist regarding what content should be taught in ECE and when that content should be understood by students (File, 2012; Hatch, 2012; Helm, 2012; Janesick, 2003; Micheal-Luna & Helmeta, 2009; Moravčík & Feeney, 2012; Spodek & Brown, 1993; Stremmel, 2012). These differing frameworks make it difficult to establish universal curriculum or content across early education settings. Whereas in public K-12 programs,
state and federal standards act as the readiness norm, early childhood programs do not necessarily subscribe to state or federal standards, depending on their funding sources.

Historically the concept of school readiness has been applied with a maturational or biological perspective (Carlton & Winsler, 1999). However, more recently systems and transactional models are being used as a framework for readiness. The current assessment incorporates developmental concepts within a transactional or systems-based theoretical framework that takes into account maturational-based key skills related to standards-based math success at school entry. These standards are an important aspect when measuring school readiness because standards are the premise for school readiness.

Academic achievement is measured based on a child's ability to demonstrate understanding of concepts dictated by the state standards. Children experiencing little to no early educational experience, especially children from at-risk backgrounds, are more likely to fall behind in kindergarten by not meeting these standards (Geoffrey, Côté, Giguère, Dionne, Zelazo, Tremblay, & ... Séguin, 2010; Espinoza & García, 2012; Hemphill, Vanneman, & Rahman, 2011; Keys, Farkas, Burchinal, Duncan, Vandell, Li, ... Howes, 2013; Lamy as cited in Pianta & Barnett, 2012; Ryan, Fauth, & Brooks-Gunn as cited in Saracho & Spodek, 2013). Children from at-risk backgrounds are more likely to lack the knowledge and experience that their peers have upon school entry, creating an achievement gap (Geoffrey et al., 2010; Espinoza & García, 2012; Hemphill et al., 2011; Keys et al., 2013; Magnuson, Meyers, Ruhm, & Waldfogel, 2004).

Creating standards for readiness and developing Early Childhood Education (ECE) programs designed to increase readiness are ways in which policy makers and
Educators can meet the goal of decreasing the achievement gap. Theoretically, if young children are entering kindergarten more ready for learning, then the achievement gap will lessen. ECE stakeholders can meet this goal through development of efficient, reliable, and valid assessment of program effectiveness. The assessment tool created and validated in the current study was developed to meet this assessment need.

If children demonstrate understanding of the state standards, then they are considered ready for school. In other words, in our current academic pedagogy, standards dictate what school readiness looks like and what academic achievement is. Standards are met in schools by using developmental perspectives like the maturational perspectives, transactional perspectives, and systems perspectives, which are described next.

Maturational Perspectives on Readiness

Gesell's (1925) maturational theory considered children as being able to learn certain concepts only after reaching a certain age and biological predisposition (Aldridge & Goldman, 2014; Kagan, 1990). This perspective is aligned with Piaget's developmental viewpoint, which views the child as becoming more academically capable as age increases (Miller, 2011). A majority of American schools and teachers follow this model, that readiness is a result of developmental age (Carlon & Winsler, 1999). With the inception of compulsory schooling came a debate about what age children should be required to attend school (Saow, 2006).

Currently, in the state of California and in many other states, there is a transition from an optional school entry at four or five years of age to a strict admittance of five
year olds to kindergarten (Snow, 2006). This transition is solely age-based. In California, school entry at five years of age is currently being addressed with the implementation of Transitional Kindergarten programs (TK) (California Department of Education, 2014). TK is a mix of preschool and kindergarten where the children attend a longer day that includes a structure and curriculum more similar to kindergarten yet it is more play based than typical kindergarten. TK is an example of how dominant maturational perspectives are in public school systems; admittance is based solely on age and not any other characteristics or abilities. Despite the dominance of the maturational perspective in defining school readiness, age by itself is not a predictor of academic success (Carlton & Winsler, 1999).

Children learn in the classroom setting regardless of age, particularly when learning formal skills. That is, despite trends in the use of maturational perspective to inform academic pedagogy, it is not the most generalizable theory to use due to environmental and experience factors. According to Purpura et al. (2013), there are two types of skills that children learn; formal and informal skills. Informal skills are concepts that are typically taught at home in the years leading up to formal schooling, whereas formal skills are skills not necessarily taught at home that require intentional teaching (Purpura et al., 2013; Purpura & Lonigan, 2013). Informal skills “affect both the child and social environment, early academic skills… are linked to subsequent academic achievement because they provide the foundation for positive classroom adaptation” (Duncaň, Dowsett, Claessens, Magnuson, Huston, Klebano, & … Jäpe, 2007, p. 2):
Depending on previous experience, children of any age will learn in the classroom based on exposure to different informal skills in their homes. This implies that even if a child is older, he or she may not be ready to learn all of the standards-based content because of lack of exposure to informal skill development. That is especially true for children who have grown up in low-SES households. Children from economically disadvantaged homes are less likely to have the experiences and informal training, that a child growing up in an economically advantaged household may have experienced (Bauchmüller, Görtz, & Rasmussen, 2014; Côté, Mongeau, Japel, Xu, Séguin, & Tremblay, 2013; Early et al., 2007; Purpura et al., 2013; Schweinhart, 2013).

Increasingly, regardless of age of entry, children are not necessarily coming into kindergarten with the informal skills that prepare them to meet the standards. ECE programs thus ideally address the development of these informal skills in children who have not developed them. and the assessment developed in the current study addresses this disconnect by measuring skills that are foundational and informal—skills that are predictive of achievement of kindergarten standards.

**Transactional and Systems Perspectives on Readiness**

Models and approaches to understanding school readiness are increasingly transactional or systems based in nature (Snow, 2006). Transactional models of school readiness view readiness as being shaped by, and understood within, the context of the school environment (Snow, 2006). In other words, transactions occur between the contextual framework of education in ECE programs, kindergarten standards, and the
student's role and purpose in the classroom. This in turn affects the how and what students are taught and learn.

According to Snow (2006) a transactional or ecological model of school readiness includes readiness of the schools themselves. This includes the quality of the schools and programs that students will be entering, and how they aid young children from home to kindergarten (Snow, 2006), as well as the degree of accessibility for all children (Kagan, 1990). School readiness through a systems theory lens is consistent with Bronfenbrenner's ecological systems theory (Miller, 2009). In this perspective, an individual's readiness is a product of their interactions within a particular school setting, community, societal climate, as well as those throughout their lifespan.

There are many external and internal influences at work within these contexts that contribute to an individual's academic performance. As a result, children interact with their environment based on individual characteristics such as development and life experiences, which affects the rate at which children process new information and collect feedback (Duncan et al., 2007). For example, when children are raised in homes with many stressors, they tend to have more difficulties academically.

The processes of bio-ecological systems theory and transactional models of development are similar to a readiness model presented by Graue (2006). Graue (1998) suggests that children should be studied in context because of the influence that context and environment have on an individual child's behavior. Graue (2006) argues that readiness is produced by social groups, institutions, individuals, and businesses, rather than the individual, and that assessing readiness is difficult because age is not the only
determinant of readiness; there are community responsibilities as well. The development of formal and informal skills is, therefore due to the interplay of child, family and community.

Consistent with this model, the readiness assessment developed in the current study would be used in a way that measures programs, rather than solely to document individual readiness. By identifying trends in student performance on the skills assessed, program-wide changes that increase readiness may be implemented. These trends may also inform curriculum used in the classroom.

The transactional and systems theoretical perspectives support the idea that an individual child's performance on a readiness assessment may not be the most accurate measure for placing children or excluding them from experiences. Indeed, many school readiness assessments are not necessarily valid when predicting actual academic performance (Carlton & Winsler, 1999; Scott-Little & Niemeyer, 2001). Carlton and Winsler (1999) argue that this is because school readiness assessments tend to measure developmental milestones or academic knowledge. Thus, the current assessment measures pre-academic conceptual knowledge based on empirically supported developmentally based milestones in numeracy development.

**School Readiness Assessments**

Current national policy emphasizes the importance of literacy, language, and mathematics (Wesley & Buysse, 2003) in public schools. Current research is influencing
federal policy through implementation of assessments for children in ECE settings, using observational and developmentally appropriate methods of assessment (Espinoza & Garcia, 2012). Federal and national initiatives are also focusing on the alignment of ECE standards with K-12 standards. With this focus comes a need to report how to accurately measure a child’s readiness relative to future academic achievement based on future standards (Neuman & Roskos, 2005). Measures must be consistent and accurate to be used in preschool classrooms and ultimately, any measure of readiness must be linked to achievement in school, i.e. standards and performance in school. Many measures used to assess readiness do not meet this basic standard. Despite the need for a link between readiness and achievement, current measures of readiness are not necessarily linked to actual school achievement (La Paro & Pianta, 2000; Niemeyer & Scott-Little, 2001; Snow, 2006). According to La Paro and Pianta (2000) school readiness assessments provide estimates of individual students’ abilities; however they are untested with regard to predictive validity. In other words, estimates are provided with testing but the use of these estimates may not be useful in determining how a student will perform.

According to Kilday and Kinzie (2008), there is a need for research to identify what math skills and assessments are appropriate for preschool classrooms. Kilday and Kinzie (2008) therefore conducted an analysis of nine of the most used math assessments. Of the nine assessments reviewed, all of them were lacking theoretical bases or goals of the assessment itself. Of the nine assessments, only one of them reported validity. The current research was intended to bridge the gap between research and practice by
providing validation of an assessment with a base in theory and specific developmental goals.

**Use and Misuse of Readiness Measures.**

La Paro and Pianta (2000) state that many schools have developed kindergarten readiness assessments that are used to test whether children should postpone attending kindergarten. However, there is less concern over whether the institution is ready for children or in using these assessments to inform program content for children. This is a particular concern for children in low-SES homes who may not be entering school with a sufficient level of informal knowledge. There is a need to evaluate whether or not readiness programs are actually supporting children’s skill development.

Using readiness measures solely to evaluate deficits in children, or worse, to exclude children from programs, is problematic in that it does not address the larger issues related to program quality and meeting children where they are developmentally and supporting their learning (Carlton & Winsler, 1999). Further, measurements of readiness in preschools can and should be used for program evaluation, curriculum development, and lesson planning.

Any effective assessment of readiness should have multiple characteristics including assessing individual (informal) skills in a reliable and valid way, be tied to actual readiness in the kindergarten classroom and, be useful in evaluating program content and effectiveness. The ENA developed in the current study is grounded in developmental research, and may potentially be useful as a tool to predict individual
kindergarten achievement, but ideally would also be useful in shaping curricula, and assessing program effectiveness.

Standards-Based Readiness

According to Snow (2006), national attention has increasingly placed importance on the development and implementation of state level standards for ECE as a means of preparing children for school entry. National content standards are attempts to standardize preschool content and assessment to align with K-12 standards, and have deemed the content areas of language and literacy and mathematics predominate (Neuman & Roskos, 2005). Federal and state level policy makers are striving for “standards based accountability” to raise the academic readiness of young children (Brown, 2007). However, Neuman and Roskos (2005) emphasized the importance of using standards that are supported by current research and that are clearly stated and understood by not only early childhood educators but by the public.

In California, the California Preschool Learning Foundations (PLFs) are generally accepted academic concepts considered important to fostering kindergarten readiness. The PLF explicitly states that the domains included in the three volume document are crucial to “strengthen preschool education and school readiness and to close the achievement gap” (California Department of Education, 2008, pp. xi):

The California PLF (2008) lists mathematic skills associated with the domain as number sense, algebra functions (classification and patterning), measurement, geometry, and mathematical reasoning. Such standards are not meant to be used as an assessment
tool according to the PLF manual, but more as a reference for how to set up environments and teach certain expected academic and/or developmental concepts. Whereas federally and state funded programs are required to use the assessment Desired Results Developmental Profile for Preschoolers 2010 (DRDP-PS 2010) to measure academic development occurring in the classroom.

The DRDP-PS 2010 was created by California Department of Education (CDE), and the Early Education and Support Division and is a universal tool for assessing development in preschool classrooms (Kareltz, Parrish, Yamada, & Wilson, 2010). Preschool teachers’ completion of the DRDP-PS 2010 was used in the current study to establish concurrent validity of the ENA. The DRDP-PS 2010 measures number sense quantity and counting, number sense mathematical operations, classification, measurement, shapes, and patterning.

In order to build the bridge from Preschool to Kindergarten, the CKCC-MS is the currently practiced standards system in California public schools. Incorporation of the CKCC-MS into the development of the ENA was necessary in an attempt to foster the predictive validity of the ENA. The skills measured started at a developmentally appropriate level to catch the informal concepts that the student would learn, up to concepts that would be measured in kindergarten by the CKCC-MS. According to the CKCC-MS (California Department of Education, 2014) critical areas of instruction should include written numerals, counting sets, comparing sets, simple equations, and representation of quantities.
The ENA represents an attempt to tie the PLF, DRDP-2010, CKCC-MS, and current research in an assessment. The numeracy skills measure by this assessment tool combines the previously listed set of standards and research into the following categories: mathematics operations, relative quantities, one-to-one correspondence, patterning, classification, rote counting; and identifying one-digit numerals. The connections between the three items are identified in Table 2.

Table 1

*Bridge between ENA, Desired Results Developmental Profile for Preschool 2010 (DRDP-PS 2010), and California Kindergarten Common Core Mathematic Standards (CKCC-MS)*

<table>
<thead>
<tr>
<th>ENA</th>
<th>DRDP-PS 2010</th>
<th>CKCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics Operations</td>
<td>Number sense of quantity and counting</td>
<td>Operations and Algebraic Thinking</td>
</tr>
<tr>
<td>Relative Quantities</td>
<td>Number sense of mathematical operations</td>
<td>Counting and Cardinality</td>
</tr>
<tr>
<td>One-to-one Correspondence</td>
<td>Number sense of quantity and counting</td>
<td>Counting and Cardinality</td>
</tr>
<tr>
<td>Patterning</td>
<td>Patterning</td>
<td>Measurement and Data</td>
</tr>
<tr>
<td>Classification</td>
<td>Classification</td>
<td>Measurement and Data</td>
</tr>
<tr>
<td>Rote Counting</td>
<td>Number sense of quantity and counting</td>
<td>Counting and Cardinality</td>
</tr>
<tr>
<td>Identifying Numerals</td>
<td>Number sense of quantity and counting</td>
<td>Counting and Cardinality</td>
</tr>
</tbody>
</table>
Early Numeracy and Mathematics Readiness

Early numeracy, which is counting abilities, numeral estimations, and logical operations, is the foundation on which future mathematics is built (Claessens & Engel, 2013; Duncan et al., 2007; Ginsburg & Amit, 2008; Kléémans, Peeters, Segers, & Verhoeven, 2012; Purpura et al., 2013). Early math skills predict later mathematics achievement from kindergarten through middle school (Romano, Bachishin, Pagani, & Kohen, 2010; Sophian, 2013). Mathematics skills are more predictive of academic success than reading/language and attention, because such skills include both a conceptual and procedural concept (Duncan et al., 2007), that is, the cognitive tools and structural process of understanding math concepts are transferable to other content areas, such as later reading skills (Romano et al., 2010). Considering early numeracy is also important because mathematics skills are more stable than those of other academic domains in relation to academic achievement (Siegler, Duncan, Davis-Kean, Duckworth, Claesens, Engel, Susperreguy, & Chen, 2012) and predictive of general readiness scores (Kléémans et al., 2012).

The assessment evaluated in the current study focuses on mathematic readiness; more specifically numeracy development as a precursor for future academic success. The skills measured by this assessment address the developmental abilities of informal skills that may be supportive of later mathematics learning and achievement. Purpura et al. (2013) discussed the concept of informal and formal mathematics knowledge. Informal mathematics knowledge are the basic numeracy skills that young children develop that foster and support later formal mathematics knowledge thought in school (Purpura et al.,
The current assessment tool measures the informal mathematics skills associated with later formal mathematics skills. These informal skills are typically fostered in homes in everyday experiences with parents and other adults.

Clark, Sheffield, Wiebe, and Espy (2013) have argued that math skills develop in young children through the course of everyday interactions with adults. However, children from low SES homes are at risk for not gaining these informal academic skills through every day experiences (Bauchmüller, Gørts, & Rasmussen, 2014; Côté et al., 2013; Early et al., 2007; Purpura et al., 2013; Schweinhart, 2013). Because children from low-SES households are more likely to support to attain informal mathematics skills (Bauchmüller, Gørts, & Rasmussen, 2014; Côté, Mongeau, Japel, Xu, Séguin, & Tremblay, 2013; Early et al., 2007; Purpura et al., 2013; Schweinhart, 2013), the sample of the current study is comprised of students attending school readiness preschool programs specifically targeting low-SES households. In these ECE classrooms, intentional teaching or reinforcement of these concepts is the typical practice in order to build the foundational skills needed for future academic achievement (Linder, Powers-Costello, & Stegelin, 2011; Purpura et al., 2013; Varol & Farran, 2006).

Current research demonstrates the benefits of high quality early education for the development of school readiness skills (Espinoza & García, 2012; Geoffroy et al., 2010; Hemphill et al., 2011; Keys et al., 2013; Lamy, 2012; Ryan, Fauth, & Brooks-Gunn, 2013). Significant relationships exist between the qualities of the classroom and acquisition of mathematical skills for children who entered the programs with few academic skills (Keys et al., 2013). For example, Keys et al. (2013) measured teacher and
classroom intentionality and relationships and found that higher quality programs fostered mathematics development. Mathematics skill development was supported by classroom instructional quality despite maternal education and ethnicity (Keys et al., 2013). This research demonstrates the importance of measuring the effectiveness of readiness programs in order to ensure program quality. The effectiveness of the programs significantly influences a student’s academic achievement, including mathematics achievement.

There are many different types of assessments currently used in ECE classrooms including standardized formal assessments, informal observational assessments, and performance assessments (Biddle, Garcia-Nevarez, Henderson, & Valero-Kerrick, 2014; Leary, 2004; Wortham, 2005). However, such assessments may not be developmentally appropriate (Wortham, 2005), may not account for cultural and language differences (Kusserow, 2012; Leary, 2004; Wortham, 2005), are often untested or are administered incorrectly or with biases (Biddle et al., 2014; Leary, 2004), and may be often cumbersome for teachers and other professionals to use. Therefore, the goal of this study was to create and validate an efficient and easy-to-use tool that measures one important and meaningful aspect of academic development, early numeracy. This tool may be used to address the need for a measure of academic skills associated with school mathematics readiness, and aid in the evaluation of readiness programs designed to narrow the achievement gap.

The current study’s assessment tool addresses numeracy skills that have been established by previous research as predicting later mathematics achievement. One such
skill is early numeral knowledge (Clæsens & Engel, 2013; Purpura et al., 2013).

Numeral knowledge consists of multiple components to mathematic knowledge. Numeral knowledge is the ability of an individual to organize and classify quantities and objects based on characteristics and amounts (California Department of Education, 2010; Clæsens & Engel, 2013; Purpura et al., 2013). Purpura et al. (2013) found that early numeral knowledge is the foundation upon which future mathematics is built. Based on findings from various studies, Table 1 shows a sample of math skills that are empirically established as important building blocks for school math achievement.

The skills measured by the ENA are those that support numerical knowledge; including: one-to-one correspondence, rote counting, patterning, classification, mathematic operations, identifying one-digit numerals, and using non-standard units to compare objects (California Department of Education, 2008; Clæsens & Engel, 2013; Clark, Sheffield, Wiebe, Espy, 2013; Clements, Sarama & DiBiasé, 2004; Ginsburg & Amit, 2008; Linder, Powers-Costello, Stegelin, 2011; Purpura et al., 2013; Varol & Farron, 2006; Wang, 2009). These skills have been reevaluated and for the current numeracy assessment tool the skills rote counting, one-to-one correspondence, relative quantities, mathematic operations, patterning, identifying one-digit numerals, and classification, will be assessed. Each of these skills is described in the following section.
### Table 2

**Research Supported Key-Early Mathematic Skills.** This table illustrates math skills supported by current research and documents.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Definition</th>
<th>Research Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-to-one correspondence</td>
<td>Ability to associate an item and quantity to a number. For example, pointing to one cracker and labeling it one, then pointing to another cracker and labeling it two.</td>
<td>California Department of Education (2010); Purpura et al. (2013); Linder, Powers-Castille, Stegelin (2011); Wang (2009)</td>
</tr>
<tr>
<td>Rote counting</td>
<td>Ability to say the names of numbers aloud and in order. For example, a child counting from one to ten aloud.</td>
<td>Purpura et al. (2013); Claesens &amp; Engel (2013); Clark et al. (2013); Linder et al. (2014); Wang (2009)</td>
</tr>
<tr>
<td>Patterning</td>
<td>Ability to create a repeating sequence. For example, a child putting a red triangle then a blue square, then a red triangle followed by a blue square.</td>
<td>California Department of Education (2010)</td>
</tr>
<tr>
<td>Classification</td>
<td>Ability to sort objects based on similar characteristics. For example, a child putting the entire toy horses in one row, then all of the pigs in another.</td>
<td>California Department of Education (2010); Linder et al. (2011)</td>
</tr>
<tr>
<td>Mathematic operations</td>
<td>Ability to perform simple arithmetic problems. For example, if the child has three cats and 2 dogs, they can identify which has more and how many more or less.</td>
<td>California Department of Education (2010); Purpura et al. (2013); Claesens &amp; Engel (2013); Clark et al. (2013); Linder et al. (2011); Varol &amp; Farron (2006); Wang (2009)</td>
</tr>
<tr>
<td>Identifying one-digit numerals</td>
<td>Ability to label a number by viewing it its written characteristics.</td>
<td>Purpura et al. (2013); Clark et al. (2013); Linder et al. (2011); Wang (2009)</td>
</tr>
<tr>
<td>Non-standard units to compare objects</td>
<td>Ability to describe attributes of an item by comparing it to non-typical measurement units. For example, using small blocks to measure how long a large block is.</td>
<td>California Department of Education (2010), Claesens &amp; Engel, 2013</td>
</tr>
</tbody>
</table>
Numeracy Skills

Classification

Classification is the skill linked to sorting and grouping objects or concepts into logical group (California Department of Education, 2008; Charlesworth & Lind, 2010; Clementś, 2004; Linder et al. 2011). This is a foundational skill linked to adding and subtracting, building on the conceptual knowledge. The conceptual knowledge is developed through object sorting and solving problems abstractly and/or concretely (California Department of Education, 2008; Charlesworth & Lind, 2010; Markman, Cox, & Machida, 1981). According to Markman et al. (1981) early categorization skills affect a child’s ability to form more complex categorical representations.

The process of classifying objects requires the use of mental structures that support the creation of complex relationships and is predictive of algebraic thinking (California Department of Education, 2008; Richardson, 2008). The CKCC-MS suggest that in kindergarten students should be able to count the number of objects in categories, categories in which students have created or sorted (California Department of Education, 2010). The ENA developed for this study addresses this concept through the classification of objects based on color. The materials consisted of two different colored bears and corresponding cups, blue and red. The participants were asked to put the bears away in the cup that is “the same color.”
Identifying One-Digit Numerals

For the current assessment, identifying one-digit numerals is considered as the ability to name the symbol associated with the numbers one to ten (Clark et al., 2013; Clements, 2004; Linder et al., 2011; Purpura et al. 2013; Wang, 2009). Identifying the symbol of an amount, a number, is the skill needed for students to be able to develop an understanding of an amount and the symbol, or numeral, associate with that amount (Charlesworth & Lind, 2010). This skill is useful for students to be able to record or represent information (Richardson, 2008).

The CKCC-MS suggests that students should have developed the ability to compare numerals between the symbols of one through 10 (California Department of Education, 2010). The ability to count with one-to-one correspondence and in turn identify the amount of items, cardinal number, is the developmental skill predictive of comparison numerical symbols. The assessment tool developed for this study measured the participants’ ability to name the numbers one through ten, not consecutively ordered.

Mathematics Operations

Mathematics operations can be defined as the ability to solve simple arithmetic problems (CDE, 2010; Claesens & Engel, 2013; Clark, Sheffield, Wiebe, Espy, 2013; Clements, 2004; Linder, Powerš-Coștălo, Stegelin, 2011; Värel & Fairén, 2006; Wang, 2009). Purpura and Lonigan (2013) found that mathematics operations is one of three primary domains that key to future mathematic success. This skill includes composing and decomposing of amounts, adding and/or subtracting to develop an understanding of
parts and whole (Charlesworth & Lind, 2010; Clements, 2004). In kindergarten, children are expected to develop understanding of addition and subtraction through numerical representation as well as in word problems with objects or actions (California Department of Education, 2014). This also consists of an explanation or representation of the process of addition or subtraction (California Department of Education, 2014).

The PLF (2008) suggests that preschool aged children potentially should be able to use simple math equations such as being able to identify when one object is added or taken away, and that it affects the total amount of items. At the preschool age, this consists of understanding that counting will be a part of the problem solving process (Clements, 2004). In the current assessment, participants are asked a series of addition and subtraction problems to measure mathematics operation development. This included one question that asked the children to add two bears, add three and one, and add three to six.

**One-to-One Correspondence**

For the current assessment study, one-to-one correspondence was defined as the ability to associate an item and a number to a set of objects (Clements, 2004; Linder et al. 2011; Pulipura et al., 2013; Wang, 2009). The skill of one-to-one correspondence is a foundation for other skills such as grouping, sorting, classifying, and scientific inquiry that varies by developmental understanding (Charlesworth & Lind, 2010; Clements, 2004). Richardson (2008) argues that until a child is able to count with one-to-one correspondence, he or she does not fully understand the concept counting as a tool.
According to Muldoon, Lewis, and Francis (2007) the ability to accurately count one-to-one is a supportive skill to being able to understand the meaning of what a number means, that to say the word seven, or to see the number seven, the child will know that it means seven objects.

The PLF states that preschool is when children begin to develop the understanding of cardinality using one-to-one correspondence, the understanding that the last item counted is the total amount counted (California Department of Education, 2008). The CKCC-MS propose that kindergarteners should be able to count in order to find out the number of objects by pairing a number with an object, also that the last object counted tells the total number of objects (California Department of Education, 2014). In the current assessment the participants are asked to count bears. The bears were lined up on the table following the mathematical operations and relative quantities questions. The participants were asked to touch each bear on the head while counting them.

**Patterning**

Patterning is the ability to sequence items in a repeating order (CDE, 2008; Clements, 2004). According to Richardson (2008) mathematics is the study of patterns, that the whole field was founded on a pattern system. Patterning is a fundamental skill associated with sequencing, spatial skills, and ordering (Charlesworth & Lind, 2010; Richardson, 2008). Lee, Ng, Bull, Pe, & Ho (2011) found that early patterning skills are also predictive of algebra tasks because it teaches children to compute and identify rules within the system. It is an essential ability associated with predicting and making sense of
situations, and/or finding order (California Department of Education, 2008; Richardson, 2008). Specifically, patterns are precursors for mathematics generally and algebra specifically (California Department of Education, 2008).

The assessment tool developed for this research measures how the participants either create their own pattern or finish a pattern prompted by the assessor. The participants are given the opportunity to demonstrate their current understanding of patterns on their own by creating a pattern, or the children are given the beginning of a pattern to finish. For example, the participants were prompted by the assessor when the assessor laid out an ABAB pattern for the participant to continue, or the participant created their own, which may have been an ABAB pattern or another structure.

Relative Quantities

Charlesworth and Lind (2010) argue that in identifying relative quantities children find relationships consisting of attributes in determining the association of groups or sets. Purpura and Lonigan (2013) found that the skill domain of relative quantities is causally linked to arithmetic performance in young children. However, the understanding of the concepts of more, less, and same are directly correlated with language development (Purpura & Lonigan, 2013). If a child does not have an understanding of the language used in such word problems they cannot demonstrate their knowledge accurately (Purpura & Lonigan, 2013).

The PLF states that around the age of 5 years of age children should be able to compare groups using the terms more, less, or same as (California Department of
The CKCC-MS considers these skills as the ability to label and compare attributes (California Department of Education, 2014). More specifically, and directly reflected in the assessment, is a child’s ability to label objects in a group as greater than, less than, or equal to another group of objects by using numeracy skills such as counting (California Department of Education, 2014; Clements, 2004). In the current assessment, participants were asked to identify the relative quantities of each set of objects as they completed operations tasks. The participants were asked to identify which of two sets of objects was “more”, “less” or “the same”.

**Rote Counting**

Rote counting is the act of verbally counting consecutively from one without any tools (CDE, 2010; Claesens & Engel, 2013; Clark et al., 2013; Clements, 2004; Linder et al., 2011; Varol & Farron, 2006; Wang, 2009). It is the skill that is used by almost every other numeracy skill and is the premise for one-to-one correspondence and identifying quantities (California Department of Education, 2008; Charlesworth & Lind, 2010). A student’s understanding of numbers and quantities is highly reliant on their ability to count by memory (California Department of Education, 2008). The CKCC-MS suggest that children in kindergarten should be able to rote count with correct number names and in the correct sequence up to 100; this includes counting by tens as well as ones (California Department of Education, 2010). In addition, those kindergarteners should have developed the ability to count and understand the concept of cardinality based on counting.
In the current study, participants were asked to start with number one and "count as high" as they can as a measure of rote counting. The administrator asked each participant to count; the highest number that they counted too was used as the measurement. If the participant had a difficult time starting to count the administrator began counting up to 5 to start the process for the participant. The participant was able to count with English or Spanish.

Developing and Testing Measures

There is a need for more reliable and valid assessments for early readiness (Kilday & Kinzie, 2008; Scheeringa & Haslèt, 2010; Wortham, 2005). According to Carlton and Winsler (1999) many assessments, both formal (measurement tools) and informal (developed by teachers and districts), are not tested for reliability. Scott-Little and Niemeyer (2001) compiled and checked the validity of 39 assessment tools used for children in kindergarten. Of the 39 assessments, three were used to measure kindergartern readiness. Of those three, two were tested and considered valid based on content, construct, concurrent, and predictive validity (Scott-Little & Niemeyer, 2001). Of these three assessments reliability was not measured for each, and the amount of time to complete each assessment ranged from 10-30 minutes. Therefore, the ENA was tested for reliability and maintained a much shorter implementation period, of around 5 minutes.

The goal of the current study was to test the validity and reliability of the Early Numeracy Assessment (ENA) developed by the author. Items for the measure were
developed through consultation of scholarly research on numeracy development and learning standards. These items were refined through piloting and then administered to a preschool sample. Then results of the administered assessment were evaluated for internal consistency, inter-rater reliability, and concurrent and predictive validity.

There are established procedures for developing a psychometrically sound assessment. The following section discusses procedures that were employed to establish reliability and validity of the ENA, as well as practical and training considerations for administering tests in the preschool environment. Camilleri and Botting (2013) argue that with a structured and scripted administration, videotaping allows for differing assessment administrators to attain comparable responses, or score the same child multiple times. In the current study the assessor used a script and video recorded every assessment in order to standardize the assessment and to minimize variance in the administration of the ENA.

**Establishing Reliability and Validity**

Reliability is established in regards to the consistency of the test across settings (Leary, 2004; Wortham, 2005), whereas validity refers to how accurately the assessment measures the purported constructs (Leary, 2004; Wortham, 2005). In order to test reliability, the researchers measured interrater agreement for items and internal consistency of the items on the scale.

**Inter-Rater Agreement.** Inter-rater agreement, a type of reliability, is determined by the degree of agreement in observation coding between two independent raters (Cozby & Bates, 2012). Inter-rater reliability can be established when assessments are scripted
and videotaped, as in the current study, as in this research. To compute inter-rater agreement, the scripted videotaped sessions were viewed by a second individual. The second rater then coded the participants’ responses from the video, which were then compared to the responses coded by the original test administrator.

**Internal Consistency.** When there are multiple similar items used in a scale, internal consistency of the items may also be assessed: Internal consistency refers to similarity between items measuring the same construct (Cozby & Bates, 2012). This research used internal consistency to assess the consistency between mathematic operations, relative quantities, one-to-one correspondence, patterning, rote counting, and identifying one-digit numerals.

For example, Camilleri and Botting (2013) examined correlations between different items of the Dynamic Assessment of Word Learning. They used Spearman’s correlation coefficients because scoring was not on consistent scales across the measure. When positive correlations were determined or found to be statistically significant then internal consistency was decidedly established (Camilleri & Botting, 2013). Similar processes were used to measure internal consistency for the current measure. Each skill measured on the ENA will be examined in relation to the other skills measured by the ENA. That is, internal consistency will be determined between patterning and one-to-one correspondence, as well as patterning and mathematic operations. Cronbach’s alpha was also calculated to assess overall item consistency for the scale.

**Validity.** Validity refers to the “truth or accuracy” of information (Cosby & Bates, 2012, p. 69). Essentially, establishing validity means ensuring that the assessment
measures or predicts what it is theoretically intended to measure or predict (Cosby & Bates, 2012; Leary, 2004; Wortham, 2005). One aspect of validity, *content validity*, is established when there is a good "match" between test items and subject area being assessed. In the current study, *content validity* was considered and improved by consulting readiness and school standards and developmental literature in making decisions about ENA items.

Other forms of validity are criterion-related. For example, Vujnovic, Fabiano, Waschbush, Pelham, Greiner, Linke, Gormley, and Buck (2014) investigated the reliability and validity of new measure of challenging behaviors in the classroom and the teacher’s responses, the Student Behavior Teacher Response (SBTR). The measure was administered in a Head Start program. Vujnovic et al. (2014) measured convergent validity by computing Pearson correlations between the SBTR and a similar measure called CLASS. Convergent validity, how the skills measured are comparable to similar skills measured by another assessment, was measured in the current study by comparing the results of the ENA with the results of similar constructs of the DRDP-2010 (Cosby & Bates, 2012; Leary, 2004).

The ENA was developed with the DRDP-2010 and the California PLF in mind. The language and content was incorporated along with current research and California CKCC-MS. The skills on the ENA almost mirror those on the DRDP-2010 measuring classification, patterning, mathematic operations, one-to-one correspondence, and rote counting. Each tool measures these skills differently but the definition of the skills is similar. Finally, predictive validity, the degree to which a measure predicts what it is
designed to measure, was partially tested by comparing ENA scores with Kindergarten teachers' assessments of mathematics skills.

**Training and Administration Issues**

Assessments of readiness can be useful for many purposes, to measure the effectiveness of a program, children's strengths, and areas for improvement (Ackerman & Coley, 2012). However, assessing young children presents potential issues such as using tools that are age appropriate and allow the child to accurately display their abilities, the attention span of young children is not very long, and often their learning is sporadic and inconsistent (Ackerman & Coley, 2012). The limits of assessment in this age group were understood going into the development of the ENA. Thus, the ENA was designed to assess numeracy skills using a concise and play-based procedure, that can easily be administered and re-administered in a short test period. The ENA is not based on subjective observational ratings; rather, the administrator can measure the students' participation in a standardized way.

As with any measure, a teacher or test administrator needs to be properly trained in administering the assessment, scoring the assessment, and rating the assessment for the measure to produce accurate results (Ackerman & Coley, 2012). However, the ENA is relatively easy to administer and requires minimal training. Further, teacher training, education, and understanding of child development play a significant role in the appropriate use assessment results (Ackerman & Coley, 2012). Therefore, the numeracy assessment developed for this research is accompanied with a teacher manual for teacher
training to prevent confounds in variance of rating and administering the ENA (see Appendix A).

Similarly, in order for an assessment to be accurate, it must be administered successfully in the home language of the child (Bornman, Sevcik, Romski, & Pae, 2010; Espinosa & Gutierrez-Clellen, 2013; Gullo, 2013; Pena, Gillam, Bedore, Bohman, 2011). However, translation alone is not sufficient, cultural considerations must be made to make a comprehensive measure (Biddle et al., 2014). This is due to cultural context? of a word, which if translated incorrectly, may be interpreted as incorrect, insensitive, or threatening (Bornman et al., 2010). In addition to lexical cultural considerations, home life should be considered particularly due to cultural biases innate within the assessment (Espinoza, 2005). Espinoza (2005) argues that a teacher must be sensitive of the needs and stages of language of the child and by working collaboratively with the child’s family to develop a more thorough snapshot of both their home language and second language.

Dual Language Learners (DLLs) experience different levels of comprehension in both their native language and second language, and that individuality needs to be considered when assessing children. Espinosa and Clellen (2013) argue that knowledge is mediated by language, and therefore it is impossible to get an accurate account of a DLL concept understanding without using the child’s first and second language. The importance of using the child’s home language as well as their second is that a child may have more understanding in one language or the other based on task or concept (Espíñosa & Clellen, 2013). For example, a child may have stronger counting skills in English (their second language) but there mathematical reasoning skills may be stronger in Spanish.
(first language). Therefore, when administering the ENA the participants were able to use which ever language they felt more confident using.

The overarching goal of the current research was to assist in assessment of school mathematiës readiness at both an individual and program level. The assessment tool developed for the current study, the Early Numeracy Assessment, differs from the typically used classroom readiness assessment in that it is grounded in, and supported by, developmental research on numeracy development, is tied to school standards, is reliable and easy to use, and is tied to assessment of skill in Kindergarten. Further, the tool is designed to be more efficient and accurate when measuring numeracy skills in the preschool classroom. The following section delineates the methods used to develop and test the Early Numeracy Assessment.
Chapter 3

METHODS

The aim of the current study was to validate a new early numeracy assessment, rooted in developmental research and current standards, for use in kindergarten readiness programs. The assessment was administered to a group of preschoolers and the results were tested for validity and reliability. The development and testing of the assessment occurred in phases: (a) development of the measures, (b) administering the assessment, (c) testing for reliability and validity.

Development of Measure

Review of Literature and Mathematics Standards

Information gained from developmental literature was used to create items for the assessment tool. Using developmentally appropriate practices that are supported by research is a key factor in assessment development (Espinoza & García, 2012). Developmental literature was also used to select skills with possible links to future academic achievement (Claessens & Engel, 2013; Duncan et al., 2007; Ginsburg & Amit, 2008; Purpura et al., 2013).

Based on an initial literature review, important numeracy skills related to school readiness include: (a) mathematics operations, (b) relative quantities, (c) one-to-one correspondence, (d) patterning, (e) classification, (f) rote counting, and (g) identifying
one-digit numerals (California Department of Education, 2010; Claesens & Engel, 2013; Clark et al., 2013; Linder, Powers-Castillo, Stegelin, 2011; Purpura et al., 2013; Varol & Farron, 2006; Wang, 2009). These skills are measured because of the core competency of the topic in relation to future academic achievement. The assessment tool measures the informal, beginning stages of each of these numeracy skills, the earliest stages of mathematics numeracy learning that typically would be occurring through everyday interactions of young children (Table 1).

School performance is increasingly standards-driven (Duncan et al., 2007; Hemphill et al., 2011). In California, the relevant standards consist of the California Preschool Foundations and the California Common Core (California Department of Education, 2010; California Department of Education, 2014). Thus, these standards were considered in developing the numeracy assessment (Table 2).

Existing Numeracy Assessments

Unlike the tendency of current assessments to measure general developmental level, the assessment tool developed for the current study was not designed to assess participants' broader abilities, but specific numeracy skills that change over time. Each task in the assessment is similar to those included in many early math assessments, for example, the research-based early mathematics assessment (REMA) developed by Clements, Sarama and Liu (2008). The 7 skills were incorporated from early assessments and research by condensing all of the previously established measurable skills into a fundamental basic category of numeracy. For example, comparing amounts, counting
strategies, arithmetic, subitizing, number sense: quantities and county, number sense: mathematic operations, etc. were able to be lumped into mathematic operations and relative quantities.

The goal of the current assessment, however, was to be easy to use in the preschool environment. Thus, every effort was made to create an instrument that was engaging for young children, of short duration, that required few materials, and that required minimal training on the part of the assessor/teacher.

**Piloting**

The first step in administering the assessment tool was piloting the measure. Initial piloting of items or sets of items was conducted informally by the author, along with a draft script. The piloting occurred within a three-week period at two of six State Preschool sites, where parental consent for observation and educational assessment was provided upon participant registration in the program. The assessment was conducted as part of a fun, game-based activity during 22 of participants’ typical assessment period.

Varol and Farran (2006) note that there is strong support for using objects when teaching and assessing mathematic concepts. The initial assessment required two cups of plastic bears, one cup with 10 red bears, and one blue cup with 10 blue bears. The assessor asked participants to complete 7 different tasks associated with 7 items on the ENA. A complete list of materials can be found in the ENA Test Manual (see Appendix A).
As a result of this piloting alterations were made to the configuration and organization of the assessment. The original assessment tool took roughly twice as long as intended. As a result of piloting, items were re-ordered and the script was altered to improve clarity and flow of the assessment, and shorten the time required for assessment. The assessment script was also translated into Spanish using an online tool, and then checked and corrected by two different native Spanish speaking preschool teachers for accuracy (see Appendix E).

The ENA Measure

The resulting measure (see Appendix F) was comprised of 6 tasks representing 7 numeracy concepts: mathematics operations, relative quantities, one-to-one correspondence, patterning, classification, rote counting, and identifying one digit numerals. This order item order allowed for the most efficient use of time with each participant.

The first task assessed both Mathematics Operations and Relative Quantities. The two skills were combined within the questioning, which was measured by the participants' response on 3 questions. In the first question, the participant was asked, “I am going to put one blue bear here and one red bear here. Which color has more bears or is there the same amount? How do you know? How many bears are there all together?” The first question, “which color has...? How do you know” was used to measure relative quantities. The second question, “how many are there all together” was used to measure mathematical quantities. The two questions were asked together in order to save time.
from transitioning too many times during the assessment. A complete English script breaks down each set of questions individually (see Appendix D).

If the participant answered the first question correctly, they were asked a follow-up question. If the participant did not answer the follow-up question correctly, they were scored with a 1 for relative quantities. If the participants answered the second question correctly, they were asked a follow-up question. If the participant did not answer the follow-up question correctly, they were scored a 1 for mathematic operations. If the participant did not answer either of these questions correctly, they were scored a 0. A coding key was created to parallel the score sheet in-order to maintain consistency in scoring (see Appendix C).

The second questions consisted of asking the participant “I have three blue bears and one red bear. Which color has less? How do you know? How many bears are there all together?” If the participant answered these questions correctly the third sets of question were asked. If the participant did not correctly answer the third set of questions they were scored with a 2 for both relative quantities and mathematic operations.

The third question for mathematic operations and relative quantities was assessed by the response the child provided after asking, “I have three red bears and six blue bears. Which color has more? How do you know? How many bears are there all together?” If the participant correctly answered the third set of questions they were given a score of 3 for both mathematic operations and relative quantities.

One-to-one Correspondence was measured by prompting and observing the participant count two rows of twenty bears. If the child correctly counted from 1 to 4
bears correctly, he/she was given a score of 1. If the child counted from 1 to 10 bears correctly, the score given was a 2. If the child correctly counted all 20 bears with one-to-one correspondence, the score of 3 was given.

For Patterning, the child was asked to complete an ABAB or an AABAAB pattern with the bears. If the participant did not create or complete a pattern, they received a score of 0. A score of 1 was achieved if the child created a pattern with prompting, if the administrator started a pattern while the participant completed the pattern. Level 2 was reached if the child was able to create an ABAB pattern without the administrator beginning the pattern. The third level was achieved if the child could complete an AABAAB pattern after the assessor began the first series of the pattern.

Classification was measured by asking the child to sort the two colored bears into their corresponding colored cups. The item was scored 0/1. If the participant did not sort the bears based on the corresponding color and cup they did not demonstrate an understanding of classification and received a “0”. A “1” was given if the child classified the bears correctly.

Next, the child completed a Rote Counting task that asked participants to count by memory as many numbers as they possibly could. Rote counted was scored as 1 when the participant correctly counted from 1 to 10, a 2 when the participant correctly counted from 1 to 30, and children were scored at a level 3 if they correctly counted from 1 to over 30. Whereas, some participants did count all the way to one hundred or well over 30 they were also scored with a 3.
The skill of Identifying One Digit Numerals was scored based on the participants' ability to correctly identify a series of written numbers 10 in random order. A 1 was scored if the child correctly identified up to 4 numbers correctly. A 2 was achieved if the child correctly identified between 5 and 7 numbers correctly. A 3 was scored if the child correctly named 8 to 10 numbers.

Administration of the Instrument

Participants

The participants were 99 preschoolers between the ages of three and five years (M = 58.23 months) attending public state preschool centers. The state preschool program admits families based solely on financial need. Therefore, all participants were from low-SES households. Annual income for qualifying families of two to 12 ranges from $3,000 to $5,900 per family. Of the 100 participants, 67% were English Learners. Home languages consisted of English, Nepali, Punjabi, Spanish, and Urdu. Access to this sample was available because the author was a lead teacher in the program. (The statistics describing the participants' characteristics can be found in Table 3).
Table 3

*Participant Demographics*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>50</td>
<td>50.5</td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
<td>49.5</td>
</tr>
<tr>
<td>Participants with IEP</td>
<td>11</td>
<td>11.5</td>
</tr>
<tr>
<td>Home Language:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>29</td>
<td>30.2</td>
</tr>
<tr>
<td>Spanish</td>
<td>50</td>
<td>52.1</td>
</tr>
<tr>
<td>Punjabi</td>
<td>6</td>
<td>6.3</td>
</tr>
<tr>
<td>Urdu</td>
<td>4</td>
<td>4.2</td>
</tr>
<tr>
<td>Nepali</td>
<td>2</td>
<td>2.1</td>
</tr>
<tr>
<td>English/Spanish</td>
<td>5</td>
<td>5.2</td>
</tr>
<tr>
<td>Unknown</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Consent for the participants was obtained from parents upon the child’s enrollment in the program. Parents consented to their children being assessed, observed, and videotaped in the classroom as a part of the typical testing and assessment schedule. Parents not wishing for their child to be videotaped were permitted to opt out of video recording. Only those sites with consenting teachers and signed consent from parents participated in the validation study. There were no participants excluded due to lack of consent. However, one site did not participate fully, as the teachers only allowed the assessments to occur for approximately 45 minutes due to a scheduling conflict, limiting the number of children participating at that site.
Procedures

All assessment occurred in the participants' home classrooms, the classrooms where they spent the majority of their school time. Each participant was asked if he or she would like "to play a game" with the assessor, and if they agreed to participate, the assessment was given. The assessment took place in a relatively quiet area of the classroom to prevent as much interruption as possible, typically at a small table away from the center of the classroom. The materials used for the assessment were: (a) 10 identical plastic blue bears, (b) 10 identical plastic red bears, (c) 1 laminated strip of numbers 1-10, (d) scoring sheet, (e) script. The testing area was prepared by clearing off the table to prevent distraction; the only items on the table were the script and the assessment materials.

A video camera was set up on a tripod behind the administrator to record the participants' responses, with a general bird's eye view of the questions and responses. This was done to ensure that the participants remained generally unidentifiable on the recording. Video recordings were used to check assessment accuracy and test for instrument reliability.

Before the questions from the script were asked, the assessor asked the children their age and asked casual questions to gain an understanding as to what language the assessment should be completed. If a participant answered the general questions in Spanish or showed difficulty using English, and if the home language was Spanish, the Spanish version was used. The participants were then asked a series of questions that can be found on the script (see Appendix D).
For each question the assessor marked off the participants' response on the score sheet (see Appendix F). As the questions were answered the assessor immediately recorded the participants' responses. When all of the questions were answered, or the participant was finished participating, the participant was excused to rejoin the classroom activities. The assessment sheet was photographed and stored digitally for later coding and analysis.

When the participant was finished with ENA their responses were coded, except for classification which was scored as 0 (incorrect) or 1 (completed correctly), each item was scored from 0 to 3. Depending on the participants' response to the increasing difficulty of the prompts, the participants were rated based on the highest level of successful completion of the task. That is, for level one on the ENA the participants' response was coded as a one. If the participant reached the third level of a skill on the ENA their response was coded as a 3. The codes from the individual items were summed to obtain a total scale score, ranging from 0 to 18.

Other Assessments

**Desired Results Developmental Profile.** To measure the validity of the ENA data collected using the Desired Results Developmental Profile for Preschoolers (DRDP-PS 2010) was used. The DRDP-PS 2010 is used to measure overall learning and development. To complete the DRDP-PS 2010 teachers observe and collect samples of work that demonstrates student understanding of domain concepts. The DRDP-PS 2010
measures multiple domains of child development including: self and social, language and literacy, English language, cognitive, mathematical, physical, and health.

For each domain there are multiple skills measured using a five-point hierarchical scale. The performance of each skill is rated on a scale from not yet at first level, exploring, developing, building, and integrating. For the current study, only the skills measured in the mathematics domain were used. The Math Development domain of the DRDP measured: number sense—quantity and counting, number sense—mathematical operations, classification, measurement, shapes, and patterning.

As part of the regular program assessment, DRDP-PS 2010 was completed for each child by the State Preschool classroom teacher and assistant teacher 2 times throughout the year. Both teachers were trained several times on DRDP-PS 2010 assessment throughout their career by multiple entities. Previous research indicates a 93% inter-rater reliability for the DRDP-R for an earlier version of the DRDP-PS 2010 (Karelitz, Parrish, Yamada, and Wilson, 2010). The intra-class correlations coefficient for the DRDP-R for preschool was also significant (Karelitz et al., 2010).

**Basic Kindergarten Math Concepts Recording Sheet.** Once the participants entered Kindergarten another assessment was conducted by the Kindergarten teachers. At the end of their first semester the Basic Kindergarten Math Concepts Recording Sheet was completed. This sheet was completed by the teacher via observation and documentation of the student's coursework. The teachers collected this data during small group activities and from the student responses on paper and pencil work. For this
research the Basic Kindergarten Math Concepts Recording Sheet was not available for all the original participants, therefore a reduced sample size \((n=15)\).

The first and second trimester the Kindergarten assessment recorded counting by ones, identifying numbers, counting unifix cubes, comparing 2 or more sets of objects using the language fewer, equal, or more, and adding objects. Each object was scored on a 4-point scale. The scale varied from the first trimester to the second trimester. At the end of the first trimester when the participant counted to twenty-nine they scored a 4. However, at the end of the second trimester if the participant counted until fifty they scored a 3; to score a 4 the participant needed to count to one hundred. Because of missing data, for the current study only 3 concepts were used to measure validity; counting unifix cubes (one-to-one correspondence), identifying numbers, and counting by ones. The Cronbach Alpha for mathematics readiness items was .93.

**Instrument Validation**

**Reliability**

Reliability was assessed in two ways: (a) through assessment of internal consistency of the measure and (b) tests of interrater agreement. Internal consistency refers to how well the items are related to each other (Cozby & Bates, 2012). Cronbach Alpha was used as a measure of internal consistency for the ENA.

Interrater reliability was measured by comparing the initial assessor's coding of the assessment with a second observer's coding of the assessment. For 5 (5%) randomly
selected participants, the first assessor administered the measure, took notes, and scored the child. A second assessor trained on how to administer the ENA, watched the videos of the assessments and recorded the participants' performance (see Appendix B). The two sets of codes were then compared, with Pearson correlations between total scores and Kappa as assessments of observer agreement.

Validity

Convergent validity refers to how participants' results on one assessment are related to the results on a similar assessment (Cozby & Bates, 2012). In the current study, convergent validity was tested by comparing the ENA with mathematics skills measured by the Desired Results Developmental Profile Revised (DRDP) version 2010. Predictive Validity was assessed by comparing ENA scores with kindergarten teachers' ratings of the readiness at school entry. Pearson correlations between these three assessments were therefore examined.
Chapter 4

RESULTS

The Early Numeracy Assessment (ENA) was designed to measure informal numeracy skills of children in early childhood education settings. The goal of this research was to determine whether the ENA was valid and reliable for use in an early education setting. In order to determine the reliability and validity, internal consistency, interrater reliability, and convergent validity were tested. Data were collected from administrations of the ENA via the instructions provided in the training manual.

Preliminary Analysis

First, descriptive statistics were conducted. Table 4 shows the scaled range of the ENA, total number of participants, and the means and standard deviations of the skills measured on the ENA across all 6 classrooms. As shown in Table 4, the mean response fell within ±1 point from the mean for mathematics operations (M = 2.07, SD = .94), one-to-one correspondence (M = 1.93, SD = .94), and patterning (M = 1.73, SD = .96), whereas the measured performance of relative quantities (M = 1.64, SD = .42), counting (M = .92, SD = .65), and identification of numeral (M = 2.08, SD = 1.07) skills showed greater variability. There was no variability in classification performance, with all children demonstrating classification on the ENA task. Classification was therefore dropped for subsequent analyses.
Table 4

*Early Numeracy Assessment (ENA) Descriptive Statistics*

<table>
<thead>
<tr>
<th>ENA items</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mathemati Operations</td>
<td>0</td>
<td>3</td>
<td>2.07</td>
<td>0.94</td>
<td>99</td>
</tr>
<tr>
<td>2. Relative Quantities</td>
<td>0</td>
<td>3</td>
<td>1.64</td>
<td>1.42</td>
<td>99</td>
</tr>
<tr>
<td>3. One-to-One Correspondence</td>
<td>0</td>
<td>3</td>
<td>1.93</td>
<td>0.94</td>
<td>99</td>
</tr>
<tr>
<td>4. Patterning</td>
<td>0</td>
<td>3</td>
<td>1.73</td>
<td>0.96</td>
<td>99</td>
</tr>
<tr>
<td>5. Counting</td>
<td>0</td>
<td>3</td>
<td>0.92</td>
<td>0.65</td>
<td>99</td>
</tr>
<tr>
<td>6. Identifying Numerals</td>
<td>0</td>
<td>3</td>
<td>2.08</td>
<td>1.07</td>
<td>99</td>
</tr>
</tbody>
</table>

Next, participant demographics were considered as variables in relation to ENA items. T-test results indicated that the mean ENA score varied depending on whether the assessment was conducted in English (M= 3.45, SD= .66, n= 66) or Spanish (M=3.23, SD= .64, n= 30), t (94) = 1.71, p=.09. Participant age was linked to ENA performance on some items, including Mathematics Operations (r=.26, p<.05), Relative Quantities (r=.24, p<.05), and One-to-one Correspondence (r=.21, p<.05). There were no other significant associations between demographic variables and ENA scores.
Internal Consistency

Internal consistency is the congruency between items that measure similar skills (Cozby & Bates, 2005). Internal consistency was established through examining correlations between each skill measured by the ENA. Two-tailed Pearson's correlations were used to determine inter-correlations between the participants' performance on skills measured by the ENA. As shown in Table 5, there were significant correlations between all ENA items, ranging from .22 to .61. The Cronbach Alpha (α) coefficient for the total ENA scale was .77, showing sufficient internal consistency (Cozby & Bates, 2005).

Table 5

Internal Consistency of ENA

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mathematic Operations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Relative Quantities</td>
<td>.39**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. One-to-one Correspondence</td>
<td>.43**</td>
<td>.45**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Patterning</td>
<td>.32**</td>
<td>.22*</td>
<td>.39**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Rote Counting</td>
<td>.41**</td>
<td>.31**</td>
<td>.39**</td>
<td>.32**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Identifying Numerals</td>
<td>.43**</td>
<td>.24*</td>
<td>.55**</td>
<td>.26**</td>
<td>.54**</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.07</td>
<td>1.64</td>
<td>1.93</td>
<td>1.73</td>
<td>2.08</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.94</td>
<td>1.42</td>
<td>.94</td>
<td>.96</td>
<td>1.07</td>
<td></td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
Inter-Rater Agreement

To assess observer agreement, a randomly selected sample of $5\%$ of the video recordings of assessment sessions were coded by a second trained independent observer/coder. The independent coder was trained on the assessment by the author and coded the videotaped assessment with reference to the assessment manual. Percentage agreement for 30 observation pairs (5 participants with 6 ENA items) was 83% (range = 60% to 100%). The overall Pearson $r$ between observers on individual item scores for individual scores was .93 for the 30 agreement pairs. Inter-observer agreement was assessed using Cohen’s Kappa with a resulting overall Kappa of .77.

Convergent and Predictive Validity

Convergent validity is documented when performance on one measure is related to performance on a similar measure (Cozby & Bates, 2012). Measuring convergent validity enables a researcher to determine how similar the constructs measure the same skill, therefore determining accuracy of the measure. For this research a two-tailed Pearson’s correlation was used to determine the relationship between the participants’ performance on similar skills measured by the ENA and the Desired Results Developmental Profiles (DRDP-PS2010). For each skill measured on the ENA, the relationship to a similar construct on the DRDP-PS (2010) was determined. For example, the average performance on ENA skill identifying numerals was correlated to DRDP-PS...
(2010)’s quantities and counting, mathematic operations, classification, and patterning. Similarly, for every other skill measured by the ENA.

As shown in Table 6, there was a significant correlation between identifying numerals and DRDP classification, \( r = .58, p < .01 \). There was also a significant correlation between counting and DRDP classification, \( r = .46, p < .01 \). A third significant correlation between identifying numerals and DRDP mathematic operations, \( r = .43, p < .01 \). The remaining correlations between ENA and DRDP items were nonsignificant. Age was significantly correlated with all DRDP measures, the strongest correlation was with DRDP measurement and age, \( r = .48, p < .01 \).
Table 6

*DRDP-PS (2010) and ENA Correlations*

<table>
<thead>
<tr>
<th>ENA Mathematics Operation</th>
<th>Quantities and Counting</th>
<th>Mathematics Operations</th>
<th>Classification</th>
<th>Patterning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENA Relative Quantities</td>
<td>.34**</td>
<td>.38**</td>
<td>.27**</td>
<td>.37**</td>
</tr>
<tr>
<td>ENA One-to-one Correspondence</td>
<td>.19</td>
<td>.31**</td>
<td>.23*</td>
<td>.21*</td>
</tr>
<tr>
<td>ENA Patterning</td>
<td>.33**</td>
<td>.35**</td>
<td>.33**</td>
<td>.41**</td>
</tr>
<tr>
<td>ENA Classification</td>
<td>.12</td>
<td>.24*</td>
<td>.23*</td>
<td>.11</td>
</tr>
<tr>
<td>ENA Identifying Numerals</td>
<td>.28**</td>
<td>.41**</td>
<td>.46**</td>
<td>.39**</td>
</tr>
<tr>
<td></td>
<td>.39**</td>
<td>.43**</td>
<td>.58**</td>
<td>.41**</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01

**Predictive Validity**

To test for predictive validity, two-tailed Pearson correlations were used to determine the relationship between the participants' performance on similar skills measured by the ENA by Kindergarten teachers after their first semester. The Basic Kindergarten Math Concepts Recording Sheet (BKMCRS) was used to measure the predictive validity of the ENA. The internal consistency of the BKMCRS was determined to be statistically significant with a Pearson Correlation. There was statistical significance between counting and identifying numbers, $r=.66$, $p<.05$. 
As shown in Table 7, some 2-tailed correlations are moderate in size, but not statistically significant (n=11-15). There are positive relationships between some items, however these relationships are not strong. A weak correlation existed between ENA relative quantities and kindergarten data identifying numerals, r=.28. Whereas a moderate nonsignificant correlation was found between ENA identifying numerals and the kindergarten data identifying numerals r=.66.

Table 7

*Kindergarten Sheet and ENA Correlations*

<table>
<thead>
<tr>
<th></th>
<th>One-to One Correspondence</th>
<th>Identifying Numbers</th>
<th>Counting</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENA Mathematic Operations</td>
<td>-.19</td>
<td>-.20</td>
<td>.72</td>
</tr>
<tr>
<td>ENA Relative Quantities</td>
<td>.09</td>
<td>.31</td>
<td>.12</td>
</tr>
<tr>
<td>ENA One-to-one Correspondence</td>
<td>.09</td>
<td>.21</td>
<td>.26</td>
</tr>
<tr>
<td>ENA Patterning</td>
<td>.24</td>
<td>.30</td>
<td>.36</td>
</tr>
<tr>
<td>ENA Rote Counting</td>
<td>-.09</td>
<td>-.01</td>
<td>.45</td>
</tr>
<tr>
<td>ENA Identifying Numerals</td>
<td>.04</td>
<td>-.03</td>
<td>-.13</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>.07</strong></td>
<td><strong>.17</strong></td>
<td><strong>.22</strong></td>
</tr>
</tbody>
</table>

*p<.05, **p<.01
In the current study, the Early Numeracy Assessment (ENA) was developed and tested for reliability and validity in an attempt to fill the need for an assessment to be used in early childhood classrooms. The results indicate suitable validity and reliability of the initial assessment with respect to inter-item consistency, inter-rater reliability, and convergent validity as a measure of school readiness.

The results indicate that the skills measured on the ENA reflect similar skill sets for the participant’s early numeracy. As performance on one ENA skill improves so does performance on any other skill measured on the ENA (Leafy, 2004). However, the intercorrelations are not so high as to indicate that the two skills are too similar or to infer that the activity measured were too alike. This suggests a single numeracy factor or construct is being assessed. Similarly, interrater agreement was good, with adequate Kappa, percentage agreement, and interrater correlations.

These results indicate that the skills measured by the ENA may be reflective of developmentally expected levels of informal math skills of children who have attended preschool for eight to nine months. Each skill is related to another skill that was described in Table 1 as having predictive, foundational importance in early numeracy development (California Department of Education, 2010; Claesens & Engel, 2013; Clark et al., 2013; Linder, Powers-Castillo, Stegelin, 2011; Purpura et al., 2013; Varol & Farron, 2006; Wang, 2009).
Validity assessments also show that the ENA has promise as a measure of mathematics readiness. There were significant positive correlations between DRDP classification and the ENA identifying numerals and counting. These correlations indicate convergent validity, in that the two tools measure similar skills, and when the participant performed well on one they also performed well on the other. The results for each assessment are reflective of one another. That is, the way in which the participant was measured for content knowledge was similar for both tools.

There was however a much lower correlation of performance for the ENA than the DRDP based on age. For the DRDP as age increased, so did overall performance on the measure. That is, as the child was older they were rated higher on the DRDP. The DRDP was developed with both developmental research and on a continuum for learning and age, indicating the levels of the DRDP are meant to measure age based skills and learning (California Department of Education, 2010). This aligns with the findings in this research demonstrating the positive correlation between DRDP and age.

With the ENA non-significant relationships were found based on age; age was not a predictor of achievement. This supports the need to consider informal skills of young children and not focus on age (Purpura & Lonigan, 2013). When kindergarten programs restrict and measure school readiness based on age alone, they are not taking the whole scope of development and experience into account based on these results; similar to the push to move away from the maturational Perspective towards a more holistic theory of learning proposed by Snow (2006).
While promising, correlations between kindergarten scores and ENA items were nonsignificant. This may be partly due to a lack of statistical power in such a small sample (n=11-15). Without data supporting a strong relationship between the two assessments predictive validity cannot be determined. That is, it cannot be statistically supported the performing well on the ENA means that the student will perform well in Kindergarten. Future research could use more kindergarten data from more participants. With a larger sample size patterns in the data may be identifiable. Ongoing research over multiple years of participants may also be necessary. According to Hymel, Lemare, and Mckee (2011) establishing validity is an ongoing process that requires multiple testing.

Limitations and Future Changes to the Assessment

The participants for this research were children generally considered at-risk due to environmental factors such as low SES, low levels of parent education, and home language not being English. Because of these factors they are more likely to be not be experiencing the same kinds of in-home experiences and interactions that their middle to upper SES peers are experiencing (Geoffroy et al., 2010, Hemphill et al., 2011, Magnuson et al., 2004; McKinsey & Company, 2009). This also implies that informal math skills are not typically being fostered in their home environments despite their age (Purpura et al., 2013). According to this result age is not a direct predictor of performance, supporting a need to move away from the maturational perspective. Also
supporting the research findings that at-risk children are much more likely to fall behind academically perpetuating the achievement gap (Heimphill et al. 2011).

As previously discussed, the sample, while relatively diverse in terms of ethnicity, may lack generalizability due to lack of diversity when compared with the overall population of children in Northern California aged 3-5. If the population were more diverse with similar results the research may be more generalizable. Limitations were also present due to the primary languages of all of the participants which were not used. In the case of the Spanish speaking participants, the assessments were not given by a native Spanish speaker. This may be the reason that when the assessments were given in Spanish the average score was slightly lower than the ENA given in English. Research supports English only or primarily English classroom teaching does not sustain academic achievement throughout elementary school due to lack of support of first language (Méndez, Crais, Castro, & Kainz, 2015; Peña, et al, 2011). Since the classrooms were primarily English only classrooms this could be a reason that the English Language Learners underperformed on the ENA.

Another limitation arose due to the small sample size of the Basic Kindergarten Math Concepts Recording Sheet. With such a small sample size there is not enough data to strongly support or predict correlations and predictive validity. There are promising relationships between the skills measured on the ENA and the Kindergarten sheet, but nothing statistically significant and therefore the ability of the ENA to predict academic achievement would be better reflected with future research.
Patterning was revised due to the piloting of the ENA in both administration and scoring. Almost all participants were demonstrating an ability to continue an ABAB pattern, and so the presentation of the task was revised. Due to this revision a third prompt and score was created. It was originally a yes/no prompt; however this was not thorough enough of an indication of understanding. The revisions for this research seemed to capture a more accurate coding based on participant ability. Future research would include scaffolding for the participants, and asked more clearly for their participation, or remove the prompted/unprompted language, instead also have 3 separate types of patterns varying in difficulty.

The language used for relative quantities was another opportunity for revision. The semantics were confusing for the participants, instead of the words same, less, more the language equal, fewer, more would be more effective. This was observational and anecdotal based on the researcher having to reiterate and use other words to reflect the meaning of less and same. The original language used was not accurate for describing quantities. These changes have been included in the script and the manual. However, for the process of advancing the measure should be considered and possibly revised for future research if necessary, specifically considering the role of semantic understanding in numeracy fluency. Including consideration of changing how the children are prompted to identify the amount of bears. Future research can include asking them instead to indicate how many bears are displayed.

Additional changes that could be made to the ENA for future research would be to make sure that the bears are neatly lined up and spaced nicely for the one-to-one
correspondence, as well as changing the coding schema for the measurement of rote counting. The organization change, making sure the bears are lined up nicely are to ensure that each child is being provided with a similar set up. In this research the table used was not always a smooth surface and so the bears fell over and may have made counting more difficult for the participants.

Also, the bears were either presented in a long line or into two lines of five per each color. This set up may also have skewed the results because the participant may have counted in groups or sets instead of as a whole. This is supported by theoretical beliefs that at this age conservation of number is difficult due to not having fully developed logico-mathematical reasoning skills (Houdé, & Guichard, 2001). This change would consist of: 0 = less than 10; 1 = less than 20; 2 = less than 30; 3 = over 30. With this consistency a measurement of correlation between these two skills can be determined. Aligning this with rote counting will also allow for the measurement to possible be streamlined further if these two measures are found to measure the same ability.

Nevertheless, future research may consider future revisions. The revised ENA omits the classification skill as is, because almost all of the participants completed this task successfully, demonstrating no variability of skill level. Future research may include classification modified for the participants to sort items based on multiple characteristics, such as size and color.

A significant area for future research would be measuring the predictive validity of the ENA with an increased sample and possibly with reliable and valid kindergarten assessments. Conducting follow-up research once participants finish Kindergarten may
support the ability of the ENA to measure skills that are predictive of academic achievement and decreasing the achievement gap. This longitudinal study can be completed after additional years of schooling to determine the long-term effects of knowledge of early education numeracy skills.

Conclusions

Hemphill et al. (2011) found that the achievement gap increases and maintains over chronological time. When children enter into formal education with a smaller knowledge base, they will remain academically behind their peers with more experiences and prior knowledge (Geoffroy et al., 2010, Hemphill et al., 2011, Mångerson et al., 2004; McKinsey & Company, 2009). Across the nation, there is an achievement gap of an average of 25 points on standardized math tests and 30 points on literacy tests between children of different races as of fourth grade (Hemphill et al., 2011). This gap persists from fourth grade until eighth grade where the average score gap was 27 in mathematics and 25 in literacy assessments gap (Hemphill et al., 2011).

There is a need for an efficient and easy-to-use tool that measures important and meaningful aspects of academic development as they relate to a defined concept of school readiness to minimize this achievement gap. To understand if a student is academically prepared to succeed in kindergarten, and consequently begin to reduce the achievement gap in public schools, the students will need to be assessed. This assessment, however, needs to be concise, consistent, generalizable, and valid (Biddle et
al., 2014; Kusserow, 2012; Leary, 2004; Wortham, 2005). This study is an attempt to fill the need for this by validating and determining the reliability of measure numeracy skills predictive of school readiness. However, future research would benefit from longitudinal research measure whether the ENA can fill this need.

Future research expanding on this project would benefit from the use a larger sample, consisting of participants with varying socioeconomic status and classroom environments. Ideally, this larger more representative sample would be tested by native language speakers of the participant’s primary language to rule out confounds based on the language and translation of the ENA. The future research should also use the revised ENA to ensure consistency across the participants.
Appendix A

Early Numeracy Assessment Test Manual

EARLY NUMERACY ASSESSMENT (ENA) Test Manual

Sheri E. Hembree
Jennifer Travis

Draft 9/15/15
No part of this manual or test may be used or copied without permission
Early Numeracy Assessment (ENA)

The Early Numeracy Assessment is meant for practical measurement of informal number concepts in early childhood education (ECE) settings, including (a) relative quantities, (b) mathematics operations, (c) one-to-one correspondence, (d) patterning, (e) rote counting, and (f) identification of numerals. These numeracy skills are supported by empirical research as being predictive of future success in mathematics (Claessens & Engel, 2013; Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov, & ... Japel, 2007; Ginsburg & Amit, 2008; Purpura et al., 2013; Romano, Babchishin, Pagani, & Kohen, 2010; Sophian, 2013).

The ENA is a play-based standardized instrument that can be used throughout the school year to measure a preschooler’s numeracy knowledge. The assessment has been developed as an intentional, interactive, hands-on activity that is engaging for young children. The child is asked age appropriate questions of skills that are typical for the preschool age group, and provided the opportunity to demonstrate understanding of the informal skills by manipulating materials. This hands-on approach helps to moderate potential language barriers and allows the child to respond in a concrete and meaningful way.

This assessment has been specifically developed for use in ECE settings by teachers, administrators, or other trained individuals serving children between the ages of 3 and 5 years. The tool can be administered in the classroom during a regular school day, eliminating the need for teachers and/or students to leave the classroom. While the measure is standardized, it is of short duration and can be administered with minimal training. As a performance-based measure, it provides a method for teachers to gain more information than solely observing incidental behavior while the child is at play. The assessment results are also an opportunity to build a more thorough understanding of the child’s conceptual understanding of informal mathematics concepts.

As a measure of informal skills children may need upon school entry, the ENA may be useful in evaluating young children’s early mathematics readiness. The information gained by administering this tool may thus be useful in informing lesson planning, curriculum modifications, classroom environment, and material rotation related to early numeracy development. It also may be useful in evaluating ECE programs by providing information about program effectiveness in establishing mathematics readiness, and in creating recommendations for program improvement.
Preparing for the Assessment

Training
Before administering the assessment, it is necessary for assessors/teachers to be trained on use of the measure. They should familiarize themselves with the script, scoring process, and materials as described in this manual. As part of training, assessors/teachers will watch sample videotaped administrations, conduct practice sessions, preferably with co-teachers so that they can compare answers and check for understanding. Within any setting using the assessment, interrater agreement on scoring assessments should be established before assessment begins.

Materials
The following materials are needed for the assessment. Materials should be set up prior to beginning the assessment, and easily available during the assessment to avoid discontinuity in testing.

1. 10 identical plastic blue-bears in plastic cup
2. 10 identical plastic red bears in plastic cup
3. 1 laminated strip of numbers 1-10
4. Scoring sheet
5. Script “cheat sheet”

Establishing Rapport with the Child
The child’s comfort is of utmost important in obtaining accurate assessment. When children are more comfortable in a familiar setting with a familiar and friendly adult, they are more likely to demonstrate their understanding. Accordingly, the assessment is best administered in the child’s classroom or other familiar location with a familiar and trusted adult. If an outside assessor is used, that person must spend adequate time in establishing rapport with the child so that the child is comfortable during testing. Assessors should spend a few moments before beginning the assessment engaging the child in conversation about his/her day or his/her interests. Display a warm and engaging demeanor throughout the assessment, encouraging the child and praising his/her effort and attention.

Another important point to consider is the child’s home language. Even if a child is learning English there is always a possibility for confusion or misunderstanding. If a child’s dominant or first language is Spanish, then Spanish should be used to conduct the assessment, if possible. A Spanish translation of the script has been provided. The child’s regular teacher will likely know which language is best used. A way to measure a child’s familiarity with a language is to simply ask the child which language he or she prefers before starting. Asking the child some friendly questions before starting to casually test their language understanding is another option. For example, the assessor/teacher might
ask the child “how old are you? Did you just have a birthday?” or “are you excited about kindergarten?”

The assessment may be conducted in a quiet corner of the regular classroom, preferably away from organized activity. The child should be seated at the table facing the teacher/assessor. If the assessment will be videotaped, then the video should be set up ahead of time with a view of the table where materials will be manipulated. Be sure to have materials at hand so that the session goes smoothly.
Appendix B

Instructions for Administering the ENA

The assessment typically takes between 4 to 10 minutes, depending on the child’s responses. The activity should be presented as a game. If a child does not want to participate then it is suggested that the teacher/assessor, ask the child at a later time. Sometimes a child is not ready or may be engaged in another activity. To introduce the activity, say something like: “I am going to play a game with these blue and red bears. Would you like to play the game with me?” The following sections describe how to use the standard script and score the child’s responses.

The Script and Procedures

The assessment procedures and script are standardized, meaning that the instructions and language used should be the same for all children assessed. Some instructions and transitional language may be paraphrased but there are certain concepts and language that the assessment tool is measuring, and so the questions (italicized in this manual) must be stated as worded and in the order provided. The script should be kept where the assessor can read and see it during the testing. For example, the script can be taped to the wall in front of the table used for the testing area or placed on the table near the assessment. However, the assessor should be completely familiar with the script, and use the script only as a cue or guide.

The assessment is organized into 3 sections: Introducing the Assessment, Administering the Assessment, and Finishing the Assessment. Assessment instructions are provided and italicized language represents the standardized script and the underlined italicized is the Spanish script. This italicized script is to be intended to be asked exactly as worded.

I. Introducing the Assessment

Tell the child who you are and what you are going to do, and ask them if they would like to join you. Say something like “Hello my name is ______. I am playing a math game at the table. Would you like to play with me?” (OR “Hola mi nombre es ______. Estoy jugando un juego de matemáticas en la mesa. ¿Quieres jugar conmigo?”). If the child agrees, seat him/her across from you at the table.

If the child is not already familiar with you, introduce yourself and spend some time talking with the child to establish rapport. Then introduce the assessment by describing the materials and what will come next:

I have two cups of bears here. A cup of blue bears and a cup of red bears. I am going to ask you some questions with the bears. Are you ready?

OR
Tengo dos tazas de osos aquí. Una taza de osos azules y una taza de osos rojos. Voy a hacerle algunas preguntas con los osos. ¿Estás listo?

II. Administering the Assessment

Once the child appears ready, begin the items on the assessment. Place relevant materials on the table in front of the child. Ask the questions as worded, making notations on the score sheet, and scoring each item in the score column (see following section on scoring)

1&2. Operations and Relative Quantities

This first questions relate to both relative quantities (more, less) and operations (1 bear and 1 bear = 2 bears. Put plastic bears on the table according to each prompt.

Same

I am going to put one blue bear here and one red bear here. Does one color have more bears or are they the same?

How many bears are there, all together?

OR

Voy a poner un oso azul aquí y un oso rojo. ¿Hay un color que tiene más o son las mismas?

¿Cuántos osos hay en total?

Less

I have three blue bears and one red bear.

Which color has less?

How many bears are there all together?

OR

Voy a poner tres osos azules y un oso rojo.

¿Cuál color tiene menos?

¿Cuántos osos hay en total?

More

I have three red bears and six blue bears.

Which color has more?

How many bears are there all together?
Voy a poner tres osos rojo y sies osos azules.
¿Cuál color tiene más?
¿Cuántos osos hay en total?

3) One-to-One Correspondence
Take all the bears out of both cups and place them on the table in two rows of 10 bears of the same color.

If I put all of the bears on the table, how many are there?
Will you count how many bears there are by touching each bear on the head?

OR

Si pongo todos los osos en la mesa, ¿cuántos hay?
¿Va a contar cuántos osos hay tocando cada oso en la cabeza?

4) Patterning
Before asking the following questions remove half of the bears from the table leaving 5 blue bears and 5 red bears.

Simple pattern
Now we are finished counting the bears. Can you show me how to make a pattern with the bears?

OR

Ahora que está terminado de contar los osos. ¿Me puede mostrar cómo hacer un patrón con los osos?

If the child is having difficulty creating a pattern, the following prompt can be given to prompt an ABAB pattern, but only if the child is not creating some kind of pattern. If needed, start an ABAB pattern with the bears, and say:

See, red, blue, red, blue

OR

Mira, rojo, azul, rojo, azul
Complex pattern

Can you show me how you can make a different pattern now? Can you finish this pattern blue, blue, red, blue, blue, red?

OR

¿Me puede mostrar cómo hacer un patrón más diferente ahora? ¿Puede terminar este patrón azul, azul, rojo, azul, azul, rojo?

5) Rote Counting

Have the child help you put all of the bears away after making the pattern, and place them out of reach. For the rote counting skill the child counts by memory.

Now that the bears are put away, can you show me how high you can count? Will start with number one and count as high as you can?

OR

Ahora que los osos se ponen lejos, ¿me puede mostrar qué tan alto pueden contar. ¿Se iniciará con el número uno y contar tan alto como puedas?

If the child does not start automatically, the teacher can prompt with:

I will start with you: 1, 2, 3...

OR

Voy a comenzar con usted: 1, 2, 3...

6) Identifying One-Digit Numbers

Take out the laminated sheet of numbers 1 to 12 (see Appendix G). Cover the numbers with a piece of paper - so that one number shows at a time - while asking the following questions.

I know that you have been playing with me for a long time. But will you show me one more thing?

Will you tell me the name of some numbers? I am going to show you some numbers and you can tell what number it is. If you do not know what the number is you can tell me 'I don't know'.

OR

Sé que usted ha estado jugando conmigo durante mucho tiempo. ¿Pero vas a mostrarme una cosa más?
¿Va a decirme el nombre de algunos números? Voy a poner los números en la mesa y te puedo decir qué número es. Si usted no sabe cuál es el número que usted me puede decir 'yo no sé'.

III. Finishing the Assessment
Thank the child for participating and excuse the child. Say something like, “Thank you for playing the games with me. You worked hard. Do you want to play somewhere else?” (OR “Gracias por jugar conmigo. ¿Quieres jugar a otro lado?”)

Take a few moments to complete any notes and complete any scoring you were not able to complete during the assessment. Don’t wait or rely on your memory to so at a later time.
Appendix C

Scoring the ENA

As each question is asked, the child's responses should be recorded on the score sheet (shown below and provided in the Appendix F), along with any notes that may be needed to code the item. The following is a description of how to complete the score sheet and how each item is coded.

The top of the score sheet has a place to put the child's id or name, gender, age, test date, and test language (Spanish or English), and the person administering the assessment. These can be completed prior to starting the assessment.

There are four columns on the score sheet. The first column lists the specific numeracy skill being measured: Relative Quantities, Mathematics Operations, One-to One Correspondence, Patterning, Rote Counting, and Identifying Numerals. Notes during the assessment are taken in the second column. The third column shows three boxes for each skill, each representing increasing understanding of the concept. After each task, the child's response should be recorded immediately to avoid errors. Check marks and x's are used for scoring, with check marks indicating correct completion of particular level or task and an x indicating an incorrect item. The final column contains the score for each item, which is the sum of check marks for each skill.
<table>
<thead>
<tr>
<th>Skill</th>
<th>Notes</th>
<th>Tasks/Levels</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Relative Quantities</td>
<td>same</td>
<td>less</td>
<td>more</td>
</tr>
<tr>
<td>2) Operations</td>
<td>1 blue+1 red =2</td>
<td>3 blue +1 red = 4</td>
<td>3 red + 6 blue =9</td>
</tr>
<tr>
<td>3) One-to-one correspondence</td>
<td>1 to 4</td>
<td>1-10</td>
<td>1-20</td>
</tr>
<tr>
<td>4) Patterns</td>
<td>simple pattern</td>
<td>simple, no prompt</td>
<td>2nd pattern</td>
</tr>
<tr>
<td>5) Counting</td>
<td>1-10</td>
<td>1-30</td>
<td>1-50</td>
</tr>
<tr>
<td>6) Number ID</td>
<td>(Circle numerals correctly identified)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Skill</th>
<th>Notes</th>
<th>Tasks/Levels</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Relative Quantities</td>
<td>same</td>
<td>less</td>
<td>more</td>
</tr>
<tr>
<td>2) Operations</td>
<td>1 blue+1 red =2</td>
<td>3 blue +1 red = 4</td>
<td>3 red + 6 blue =9</td>
</tr>
<tr>
<td>3) One-to-one correspondence</td>
<td>1 to 4</td>
<td>1-10</td>
<td>1-20</td>
</tr>
<tr>
<td>4) Patterns</td>
<td>simple pattern</td>
<td>simple, no prompt</td>
<td>2nd pattern</td>
</tr>
<tr>
<td>5) Counting</td>
<td>1-10</td>
<td>1-30</td>
<td>1-50</td>
</tr>
<tr>
<td>6) Number ID</td>
<td>(Circle numerals correctly identified)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Relative Quantities and Mathematics Operations**

Both Relative Quantities and Operations are measured at the same time to streamline the assessment. Relative Quantities refers to the ability to identify whether there are more, fewer, or an equal amount within a set of items. In the case of this assessment, it is whether there are more, fewer, or equal amounts of sets of bears. There is a series of three questions used as part of this assessment: the first asks if the sets are **EQUAL**, the second asks which set has **LESS**, and the third asks which set has **MORE**.

Mathematic Operations refers to the ability to complete basic arithmetic problems. In this assessment, after asking for the relative quantity question, the child is asked to show whether she or he can add the bears together to come up with a total number of bears. For example, the first task asks the child to add one blue bear and one red bear ("How many
altogether”). For a child to successfully complete this task, he/she would have to state that there are 2 bears.

**Scoring Relative Quantities and Operations.** Check marks and x’s are used for scoring, with check marks indicating correct completion of item and an x indicating an incorrect item. For example, if in the first question a child states (correctly) that there are the same number of bears, and that there are two bears all together, but did not correctly complete the other items in the task, the assessor should place a check mark in the first box for both operations and relative quantities, but x’s in the other boxes. The score for each skill is obtained by summing the number of check marks for each skill. See the example below.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Notes</th>
<th>Tasks/Levels</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Relative Quantities</td>
<td>same</td>
<td>less</td>
<td>more</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>2) Operations</td>
<td>1 blue+1 red =2 ✓</td>
<td>3 blue +1 red =4 ✓</td>
<td>3 red + 6 blue =9 x</td>
</tr>
</tbody>
</table>

**One-to-One Correspondence**

One-to-one correspondence is the understanding that an object in a set represents a quantity, that is, if there are 3 bears a child can count each bear and say that there are 3 bears. This assessment measures a preschooler’s ability to perform one-to-one correspondence with up to 20 objects by recording how many bears are counted (in order) while pointing or touching individual and separate bears. All of the bears are put on the table and the child is asked “if I put all of the bears on the table, how many bears are there? While you are counting can you touch each bear on the head?” It is crucial that the assessor observe whether the child counts each individual bear as a separate object, an important indicator of one-to-one correspondence. For example, if the child touches 1 bear and says “one”, touches the next bear and says “two”, touches the next bear on the head and says “three”, and touches the same bear on the head and says “four”, then the highest number for one-to-one correspondence is 3. If the child skips a number or says numbers out of order, then the concept is not being demonstrated. Note the highest number correctly identified, in order.

**Scoring One-to-One Correspondence.** One-to-one correspondence is scored using 3 different levels of performance. Check marks indicate that the child has obtained at least that corresponding level of performance. Level 1 is demonstrated by the child counting up to 4 objects correctly. Level 2 understanding is exhibited when a child counts up to 10 objects while touching or pointing at each bear. Level 3 is met when a child counts all 20
bears by touching each individually on the head. Place a check mark in each level correctly achieved.

For example, if the child counts with corresponding touching for 12 bears, then the first two boxes are checked and an x is placed in the third box. The total score for the item is 2

<table>
<thead>
<tr>
<th>Skill</th>
<th>Notes</th>
<th>Tasks/Levels</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>3) One-to-one correspondence</td>
<td></td>
<td>1-20</td>
<td>2</td>
</tr>
</tbody>
</table>

**Patterning**

Patterning is the ability to create a repeating sequence. In the ENA, the child is asked to put the bears in two different sequences, e.g., red, blue, red, blue, red, blue... (ABAB pattern), and one “different” pattern, e.g., red, red, blue, red, blue, red, blue... or red, red, blue, blue... (AAB or AABB). A distinction is made as to whether or not the child can create a simple (ABAB) pattern without adult prompting. Such prompting for the first pattern is done only if the child does not create a pattern on his or her own.

**Scoring Patterning.** As with other skills, check marks are made at each level of completion. Check the first box (level 1), if the child finishes a pattern at all, whether prompted or not. Check the second box (level 2) if the child was able to create the pattern WITHOUT prompting. The third box is checked if the child makes a different pattern (e.g., BBA or BBAA), whether prompted or not. The score is simply the sum of the check marks. Note that despite the hierarchical organization of the items, it is possible for the child to have a check mark in the first and third boxes only, that is, he/she needed a prompt for the first pattern, but was able to create a second pattern.

For example, consider the following scenario: The assessor says to the child, “Now that we are finished counting the bears. Can you show me how to make a pattern with the bears?” The child reaches for the bears and starts making a pattern that is red, blue, red, blue, red, blue... (ABAB). Once the student is finished making this pattern the teacher begins a pattern with the bears and says” Can you show me how you can make another pattern now? Can you finish this pattern, blue, blue, red, blue, blue, red?” The child then finishes the sequence correctly.

The child created a pattern (box 1 is checked), did not need a prompt to do so (box 2), and was able to complete the more complex pattern (the third box is checked). The score for Patterning would therefore be 3.
Rote Counting

Rote counting is the ability to count based on memory; it is the abstract version of one-to-one correspondence. This means that instead of counting objects, the student recites number names in order. For the ENA this is demonstrated when a child can continuously count in response to the request, “Now that the bears are put away, can you show me how high you can count? Will start with number one and count as high as you can?”, the child can count in order without skipping numbers. The child may do this spontaneously, but if not, offer the prompt.

When working with the children it is important to remember that they are not counting objects. The children may want to count the bears like they did in the one-to-one correspondence task. Thus, the bears are removed from the workspace for this task.

Scoring Rote Counting. Check the first box (level 1) when the student can count from 1 to 10 in the correct order. That means that the student said at least “one, two, three, four, five, six, seven, eight, nine, ten” without skipping numbers. Check the second box (level 2) if the child reaches at least 30 correctly, Box 3 (level 3) is checked if the student reaches at least 50 correctly in response to the prompt.

For example, consider the following scenario: The teacher/assessor asks the student “can you show me how high you can count? Will you start with number one and count as high as you can?”. The child looks at the teacher and does not say anything. So the teacher says “I will start with you: 1, 2, 3, 4...”. The child begins to count along and continues: “1, 2, 3, 4, 5, 2, 3, 4, 5, 7, 33, 14”.

The student correctly counts from 1 to 5 and then does not demonstrate an understanding that after 5-the sequence is 6, 7, 8, 9, 10, etc. Therefore, only column 1 can be checked on the ENA score sheet. As shown below, the student would receive a score of 1.
Identifying Numerals

Identifying numerals refers to the ability to look at roman numerals and label them. This means that when the child sees the numeral 3, he/she shows that that numeral represents “three”. What is measured in the ENA is the numerals correctly identified as representing the numbers 1 to 12. For this task, it is not necessary for the child to recognize the numerals in any order. Each correct identification is counted.

Scoring Identifying Numerals. The scoring for this section is a little bit different from the other numeracy skills. As each numeral is displayed, the assessor/teacher should circle the numerals correctly identified, then use the total number correctly identified to check the appropriate level boxes. Level 1 is met when a child correctly identifies up to 4 numbers correctly. Level 2 is demonstrated when a child correctly labels 5 to 8 numbers correctly. Level 3 is demonstrated when a child correctly identifies between 8 and 12 numbers correctly.

For example: The child correctly names the numbers 1, 3, 4, 5, 7, 9 correctly and either misnames or says “I don’t know” for the other numbers. The score sheet would show 6 numbers circled as correctly identified. Therefore, boxes (levels) 1 and 2 would be checked by the teacher. With level 1 and 2 checked the teacher/assessor scores the skill as 2.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Notes</th>
<th>Tasks/Levels</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>7) Number ID</td>
<td></td>
<td>(Circle numerals correctly identified)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 10 6 4 5 2 7 11 8 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-4 correct 5-8 correct 9-12 correct</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
</tbody>
</table>
Appendix D

English Script

I have two cups of bears here. A cup of blue bears and a cup of red bears. I am going to ask you some questions with the bears. Are you ready?

1&2. Operations and Relative Quantities
   same
      I am going to put one blue bear here and one red bear here. Does one color have more bears or are they the same?
      How many bears are there all together?
   less
      I have three blue bears and one red bear. Which color has less? How many bears are there all together?
   more
      I have three red bears and six blue bears. Which color has more? How many bears are there all together?

3) One-to-One Correspondence
   If I put all of the bears on the table, how many are there? Will you count how many bears there are by touching each bear on the head?

4) Patterning
   Simple
      Now that we are finished counting the bears. Can you show me how to make a pattern with the bears? If needed, See, red, blue, red, blue....
   Complex
      Can you show me how you can make a different pattern now? Can you finish this pattern blue, blue, red, blue, blue, red?

5) Rote Counting
   Now that the bears are put away, can you show me how high you can count? Will start with number one and count as high as you can?
   can prompt with: I will start with you: 1, 2, 3...

6) Identifying Numbers
   Will you tell me the name of some numbers? I am going to show you some numbers and you can tell what number it is. If you do not know what the number is you can tell me 'I don't know'.
Appendix E

Spanish Script

Tengo dos tazas de osos aquí. Una taza de osos azules y una taza de osos rojos. Voy a hacerle algunas preguntas con los osos. ¿Estás listo?

1&2. Operations and Relative Quantities

*Misma* (same)
- Voy a poner un oso azul aquí y un oso rojo. ¿Hay un color que tiene más o son las mismas?
- ¿Cuántos osos hay en total?

*Menos* (less)
- Voy a poner tres osos azules y un oso rojo.
- ¿Cuál color tiene menos?
- ¿Cuántos osos hay en total?

*Mas* (more)
- Voy a poner tres osos rojos y seis osos azules.
- ¿Cuál color tiene más?
- ¿Cuántos osos hay en total?

3) One-to-One Correspondence

Si pongo todos los osos en la mesa, ¿cuántos hay? ¿Va a contar cuántos osos hay tocando cada oso en la cabeza?

4) Patterning

*Simple*
- Ahora que estas terminado de contar los osos. ¿Me puede mostrar cómo hacer un patrón con los osos?

*If needed, Mira, rojo, azul, rojo, azul*

*Complex*
- ¿Me puede mostrar cómo hacer un patrón más diferente ahora? Puede termina este patrón azul, azul, rojo, azul, azul, rojo?

5) Rote Counting

Ahora que los osos se ponen lejos, ¿me puede mostrar qué tan alto pueden contar.

Se iniciará con el número uno y contar tan alto como puedas?

*can prompt with: Voy a comenzar con usted: 1, 2, 3...*

6) Identifying Numbers

¿Va a decirme el nombre de algunos números? Voy a poner los números en la mesa y te puedo decir qué número es. Si, usted no sabe cuál es el número que usted me puede decir 'yo no sé'.
Appendix F
ENA Score Sheet

CHILD ID ___________________________ CHILD AGE ___________________________
YEARS ___________ MONTHS ___________ SEX: M F
TEST DATE ___________________________ TEST LANGUAGE: SP ENG
ADMINISTRATOR ___________________________

<table>
<thead>
<tr>
<th>Skill</th>
<th>Notes</th>
<th>Tasks/Levels</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Relative Quantities</td>
<td>same</td>
<td>less</td>
<td>more</td>
</tr>
<tr>
<td>2) Operations</td>
<td>1 blue+1 red =2</td>
<td>3-blue+1 red =4</td>
<td>3 red + 6 blue =9</td>
</tr>
<tr>
<td>3) One-to-one correspondence</td>
<td>1 to 4</td>
<td>1-10</td>
<td>1-20</td>
</tr>
<tr>
<td>4) Patterns</td>
<td>simple pattern</td>
<td>simple, no prompt</td>
<td>2nd pattern</td>
</tr>
<tr>
<td>5) Counting</td>
<td>1-10</td>
<td>1-30</td>
<td>1-50</td>
</tr>
<tr>
<td>6) Number ID</td>
<td>(Circle numerals correctly identified)</td>
<td>3 10 6 9 4 5 2 7 11 8 1 12</td>
<td>1-4 correct</td>
</tr>
</tbody>
</table>

Additional Comments:

88
Appendix G

Identifying Numerals Template
References


Bagliçi, S., Coddin, R., & Tryon, G. (2010). Extending the research on the tests of early numeracy: Longitudinal analyses over two school years. Assessment for Effective Intervention, 35(2), 89-102.


Early, Diane M., Maxwell, Kelly L., Burchinal, Margaret, Alva, Soumya, Bender, Randall H., Bryant, Donna, ... Zill, Nicholas. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development, 78*(2), 558-580.


doi:10.1037/0012-1649.33.4.711


