Published in final edited form as:

J Sch Psychol. 2014 February ; 52(1): 83–96. doi:10.1016/j.jsp.2013.11.006.

Confirmatory factor analysis of the Early Arithmetic, Reading, and Learning Indicators (EARLI)*

Kate E. Norwalk^{a,*}, James Clyde DiPerna^b, and Pui-Wa Lei^b

^a Department of Educational and School Psychology, The Pennsylvania State University, USA

^b Department of Educational Psychology, School Psychology, and Special Education, The Pennsylvania State University, USA

Abstract

Despite growing interest in early intervention, there are few measures available to monitor the progress of early academic skills in preschoolers. The Early Arithmetic, Reading, and Learning Indicators (EARLI; DiPerna, Morgan, & Lei, 2007) were developed as brief assessments of critical early literacy and numeracy skills. The purpose of the current study was to examine the factor structure of the EARLI probes via confirmatory factor analysis (CFA) in a sample of Head Start preschoolers ($N = 289$). A two-factor model with correlated error terms and a bifactor model provided comparable fit to the data, although there were some structural problems with the latter model. The utility of the bifactor model for explaining the structure of early academic skills as well as the utility of the EARLI probes as measures of literacy and numeracy skills in preschool are discussed.

Keywords

Early literacy; Early numeracy; Preschool; EARLI

1. Introduction

Research regarding the nature, correlates, and predictors of academic achievement has been mounting for decades (National Institute of Child Health & Human Development, 2000). Although the majority of this research has focused on school-aged children, studies within the past decade have begun to highlight the importance of early reading and mathematics skills prior to school entry. Early deficits in literacy and mathematics skills are highly predictive of later negative academic outcomes and remain relatively stable over time if not remediated (Jordan, Glutting, & Ramineni, 2010; Torgesen, 2002). Despite growing interest in early intervention for young children demonstrating skill deficits, there are relatively few measures available to periodically monitor the progress of early academic skills in

*This study was supported by a grant from the Administration for Children and Families, U.S. Department of Health and Human Services (Award No. 90YF0047/01). The opinions expressed in this article do not necessarily reflect the position of the U.S. Department of Health and Human Services, and such endorsements should not be inferred.

* Corresponding author at: School Psychology Program, 137 CEDAR Building, Penn State University, University Park, PA 16802, USA. Tel.: +1 814 865 1881. norwke43@gmail.com (K.E. Norwalk).
jcd12@psu.edu (J.C. DiPerna), puiwa@psu.edu (P.-W. Lei)

preschoolers. The Early Arithmetic, Reading, and Learning Indicators (EARLI; DiPerna, Morgan & Lei, 2007) were developed to provide brief and easily administered assessments of critical early literacy and numeracy skills for preschool students. Initial evidence has suggested that scores from the EARLI probes (i.e., literacy and numeracy tasks) demonstrate good technical adequacy, as demonstrated by moderate to high reliability estimates and concurrent relations with norm-referenced achievement test scores; however, the underlying factor structure of the EARLI probes has not yet been investigated. The purpose of the current study was to examine the underlying factor structure of the EARLI probes via confirmatory factor analysis (CFA).

1.1. Early literacy

Although the National Reading Panel report (National Institute of Child Health & Human Development, 2000) provided an extensive review of research related to reading in grades K-12, it did not include research regarding early literacy in children younger than 5. As a result, the National Early Literacy Panel (NELP) was formed in 2002. This panel distinguished between *conventional* literacy skills, or those that are taught and developed in the primary and secondary school years, and *early* or *emergent* literacy skills. Early or emergent literacy refers to the knowledge and skills developed prior to school entry that are necessary precursors of conventional reading and writing (Missall et al., 2007; Sulzby & Teale, 1991; Teale & Sulzby, 1986). In the NELP report, National Early Literacy Panel (2008), 11 early literacy skills were identified as being predictive of future success in reading. Three of these skill domains, alphabet knowledge, phonological awareness, and oral language are assessed by the EARLI measures featured in the present study.

1.1.1. Alphabet Knowledge—Alphabet Knowledge (AK) broadly consists of both knowledge of the names of letters and knowledge of the sound or sounds that each letter makes. Letter name knowledge (LNK) has been shown to be a powerful predictor of later reading acquisition by aiding in inventive and phonetic spelling and word recognition (see Foulin, 2005 for a review). Knowledge of letter names also significantly aids in the development of letter-sound knowledge. For example, Kim, Petscher, Foorman, and Zhou (2010) found that LNK increased the probability of knowing letter sounds from 4% to 63% in a sample of kindergarten students. The contribution of LNK to knowledge of letter sounds is likely due to the iconic nature of the English language, in that many of the names of letters of the alphabet contain the sound that they represent (Treiman, Kessler, & Bick, 2003). Coupled with LNK, letter-sound knowledge fosters growth in decoding skills (Phillips, Clancy-Menchetti, & Lonigan, 2008).

1.1.2. Phonological awareness—Phonological awareness, or phonological sensitivity, involves the ability to detect and manipulate the sound structure of language. This ability includes skills such as the ability to blend segmented syllables into words, and in turn, the ability to segment words into their constituent parts (Anthony & Lonigan, 2004). Children with strong phonological awareness skills learn to read more quickly than those who are less skilled at performing these tasks (Bryant, MacLean, Bradley, & Crossland, 1990; Lonigan, Burgess, & Anthony, 2000). Although there is general agreement that phonological awareness consists of several different skills varying in terms of linguistic complexity

(Anthony et al., 2002), disagreement exists as to whether these skills are distinct or increasingly complex points on the same developmental continuum. Research within the past decade has shown support for the latter viewpoint, suggesting that phonological awareness is best described as a unitary construct, consisting of several interrelated skills (e.g., Anthony & Lonigan, 2004; Anthony et al., 2002). These skills range from the ability to manipulate larger units of sound, such as words and syllables, to smaller units of sound such as onset-rimes and individual phonemes (Phillips et al., 2008).

1.1.3. Oral language—In addition to code-related skills such as letter knowledge and phonological awareness, oral language skills in preschool are important precursors of reading acquisition. Oral language consists of several interrelated skills, including vocabulary knowledge, discourse and syntax (Dickinson & McCabe, 2001). During preschool, it appears that the development of oral language skills has both direct and indirect effects on decoding skills, and reciprocally, the development of decoding skills fosters growth in oral language skills (Kendeou, van den Broek, White, & Lynch, 2009; Lonigan et al., 2000). As children progress into elementary school, oral language skills become strong predictors of comprehension in the later grades (Storch & Whitehurst, 2002).

Research has demonstrated a high degree of continuity between literacy skills in early childhood and reading achievement in as late as the high school years (Cunningham & Stanovich, 1997). Children who enter school with limited early literacy skills often fail to catch up to their peers by the end of fourth grade and typically continue to have difficulties in reading throughout their school years (Torgesen, 2002). Decoding and oral language skills in preschool are also predictive of comprehension skills in the primary grades (Kendeou et al., 2009). Children who do not develop these skills are subsequently at an increased risk for becoming poor readers when they are older (Juel, 1988) and often lack the skills necessary to be successful in school and work (e.g., Pressley & Rankin, 1994).

1.2. Early numeracy

Researchers also have begun to highlight the importance of early mathematics, or numeracy, skills in predicting later outcomes in mathematics achievement. In an extensive summary of research on early numeracy, Cross, Woods, and Schweingruber (2009) identified two core areas that are important for future success in mathematics: number and geometry/measurement.

1.2.1. Number—According to the National Research Council's report, *Adding it Up* (Kilpatrick, Swafford, & Findell, 2001), *number* is one of the most important concepts necessary for mastery of mathematics in elementary school. At the preschool level, this concept often has been referred to as *number sense* in the empirical literature. Although there is some debate as to the exact definition of number sense, critical skills are thought to include informal skills that underlie the formal mathematical knowledge learned through explicit instruction (Methe et al., 2011). These informal skills include number identification, counting aloud, counting objects, and basic arithmetic skills (e.g., Correa, Nunes, & Bryant, 1997; Kroesbergen & Van Luit, 2003; Stein, Silbert, & Carnine, 1997). Subitizing, or identifying the number of objects in a small group without counting, also develops in

preschool and leads to the ability to count objects in sets in order to make comparisons of less and more (Jordan & Levine, 2009). Counting objects allows children to apply number names to objects and thus develop skills such as one-to-one correspondence and cardinality, or the recognition that the last number counted in a set represents the quantity in that set (Polignano & Hojnoski, 2012).

1.2.2. Geometry and measurement—Although number sense is typically considered the most important area of mathematical knowledge in preschool, Sarama and Clements (2009) identified geometric thinking and measurement as equally important to the development of young children's mathematical understanding. Geometric knowledge in preschool involves the ability to identify and label various shapes as well as a general understanding of the characteristics that differentiate shapes from one another (Snow & Van Hemel, 2008). According to the National Council of Teachers of Mathematics (NCTM, 2000), preschool children should be able to recognize, label, draw, and compare both two- and three-dimensional shapes. Measurement bridges the skills of number and geometry by applying quantifiable numbers on spatial concepts such as length and weight (Clements & Stephan, 2004). Key measurement skills in preschool include the ability to recognize, compare, and contrast objects of varying length, weight, and size (NCTM, 2000).

Also of importance to students' mathematical success is the development of algebraic thinking. In preschool, this involves the ability to identify patterns. Preschoolers engage in informal patterning through typical play activities, such as songs, coloring, and shapes (Charlesworth, 2005; Polignano & Hojnoski, 2012). Formal patterning, however, involves the ability to recognize patterns and determine how they can be created and continued (NCTM, 2000). Pattern recognition forms the beginning of algebraic thinking and is related to other skills such as subitizing, counting, and basic numerical operations (Polignano & Hojnoski, 2012).

Acquisition of these skills in kindergarten is a strong predictor of mathematics performance at the end of first grade (Clarke & Shinn, 2004; Jordan, Kaplan, Locuniak, & Ramineni, 2007) and third grade (Jordan, Kaplan, Ramineni, & Locuniak, 2009). These skills are important for later success in mathematics (e.g., Mercer & Miller, 1992) as they help children master the four basic math operations and problem solving (e.g., National Council of Mathematics Educators, 2004; Van Luit & Schopman, 2000).

1.2.3. Measuring early skill development—Given the importance of early literacy and numeracy skills, it is essential to have accurate screening and benchmark assessment tools to periodically monitor acquisition and growth of these skills prior to formal school entry. At the preschool level, such measures can be used to determine whether children are mastering critical early literacy skills (e.g., letter-naming and letter sounds skills) and numeracy skills (e.g., counting and number-naming skills). Results from these measures, in turn, can be used to evaluate response to intervention, inform the pace and content of instruction, and guide modifications based on individual student needs. In addition, measures that can help identify children who are not acquiring early academic skills can help guide early intervention efforts and subsequently remediate skill deficits prior to school entry.

Brief benchmark screening measures of early academic skills have several benefits relative to lengthier, norm-referenced achievement tests. They are inexpensive, require less time and effort to complete, and can be administered by classroom teachers and paraprofessionals with whom children may be more comfortable working (Wilson & Lonigan, 2010). In addition, such measures are designed to be more sensitive to changes in skill level over shorter periods of time, whereas norm-referenced achievement tests are designed to assess skill development over longer periods of time (Marston, Fuchs, & Deno, 1986). Given the rapid developmental changes that take place during the preschool years, periodic progress monitoring may be more helpful to early childhood educators than norm-referenced testing that indicates a child's relative standing compared to same-age peers. Finally, although published norm-referenced achievement measures produce technically adequate scores, these scores typically reflect performance on a small number of items for younger children and often are not highly effective at informing instruction. On the contrary, benchmark screening tools typically measure fundamental skills that are more closely tied to the curriculum, which can assist in instructional planning and decision-making.

Despite their purported utility, studies examining the psychometric properties and decision-making accuracy of periodic screening measures at the elementary level have been somewhat mixed, with some indicating predictive value for the identification of at-risk students (e.g., Methe, Hintze, & Floyd, 2008; Shapiro, Solari, & Petscher, 2008) and others finding inconsistent results (e.g., Goffreda, Diperna, & Pedersen, 2009; VanDerHeyden, Broussard, & Cooley, 2006). Currently, there are a limited number of published studies (e.g., Floyd, Hojnoski, & Key, 2006; Phillips, Lonigan, & Wyatt, 2009) that have examined the psychometric properties of scores from early literacy and numeracy screening measures in preschool children. Given the importance of early intervention, the development of early academic screening measures is crucial. According to Justice, Invernizzi, and Meier (2002), an early literacy screening tool needs to meet three important criteria: (a) yield scores with sound psychometric qualities, (b) measure specific early literacy skills that have been found to be most predictive of later reading achievement, and (c) detect children at-risk for reading difficulties. In addition, screening measures must be developmentally appropriate (i.e., measure skills that are relevant to a particular age, grade, and time point), have a specific target population, and include appropriate cutoff scores that have been cross-validated across multiple groups of students to yield good classification accuracy (Jenkins, Hudson, & Johnson, 2007).

1.3. Preschool measures of early academic skills

The past decade has seen an increase in attempts to develop and validate early literacy and numeracy measures at the preschool level. For example, the Early Literacy Individual Growth and Development Indicators (EL-IGDIs; Early Childhood Research Institute on Measuring Growth and Development, 1998, 2000) were developed as a set of general outcome measures to assess both receptive vocabulary (i.e., Picture Naming) and code-related early literacy skills (i.e., Rhyming, Alliteration, and Segment Blending). The EL-IGDIs have been shown to have adequate psychometric properties (see Missall, Carta, McConnell, Walker, & Greenwood, 2008 for a summary). These measures also have been shown to be sensitive to growth in literacy skills across preschool, and moderate predictors

of literacy skills measured in kindergarten (Missall et al., 2007). Similar to other general outcome measures, the EL-IGDIs were designed to be low cost, easy to administer, and sensitive to growth over time (Missall et al., 2008).

One example of an early numeracy general outcome measure is the Preschool Numeracy Indicators (PNIs; Floyd et al., 2006). The PNIs include assessment tasks such as counting aloud in sequence, counting objects, identifying and stating the names of numbers, and quantity discrimination. Floyd et al. (2006) found adequate psychometric properties for PNI scores including test–retest reliability, moderate to strong intercorrelations among tasks, and sensitivity to age differences. In addition, Hojnoski, Silbergliitt, and Floyd (2009) found the PNIs to be sensitive to growth in numeracy skills across preschool. VanDerHeyden et al. (2004) also developed six experimental CBM early numeracy probes specifically for preschool children. These tasks included selecting numbers, number naming, counting objects, counting aloud, discrimination, and shape identification (VanDerHeyden et al., 2004). In a follow-up study, VanDerHeyden et al. (2006) further revised four of these measures (Choose Number, Count Objects, Free Count, and Discrimination). Although they observed moderate to strong correlations with scores on kindergarten mathematics measures, findings were mixed in terms of the sensitivity of these measures to assessing skill change over time.

Although there is evidence to support the use of the PNIs and the revised probes designed by VanDerHeyden et al. (2006), neither of these measures assess geometric or algebraic thinking. To address this limitation, Polignano and Hojnoski (2012) developed a set of five experimental CBM tasks designed to measure these areas in preschool children. These included Cardinality, Pattern Completion, Shape Naming Fluency, Shape Selection Fluency, and Shape Composition. In their initial study, Polignano and Hojnoski found preliminary evidence to support the use of these measures, particularly the Cardinality and Shape Selection Fluency tasks, with less support for the Pattern Completion task.

1.4. The EARLI probes for preschool skill assessment

Although the number of available preschool progress monitoring tools is growing, few of these measures focus on both the early literacy and early numeracy domains. In addition, many of the measures currently available are only appropriate for 4-year-old children. To address these limitations and meet the need for additional measures of early academic skills for preschool-age children, DiPerna et al. (2007) created the Early Arithmetic, Reading and Learning Indicators (EARLI). Developed specifically for preschool children at-risk for academic difficulty, EARLI includes two sets of brief measures (or probes). The first set assesses critical early literacy skills (e.g., phonological awareness, vocabulary, print knowledge, and letter knowledge), and the second set measures early mathematics skills (e.g., number identification, pattern and shape recognition, and counting).

Target skills for the EARLI Literacy and Numeracy measures were selected based on theory and reviews of early childhood literature. Both code-related skills (e.g., phonological awareness and letter knowledge) and oral language skills (e.g., expressive vocabulary) are represented in the literacy tasks. Similarly, the numeracy tasks reflect multiple domains as well, including number (e.g., number naming and counting objects) and geometry/

measurement concepts. In addition, these measures utilize item response theory (IRT) scores to facilitate accurate measurement of change in skills over time. IRT scores allow items from various forms of the measure to be placed on a single ability continuum, which in turn allows for the comparison of scores across these different forms. Initial studies of the reliability and validity of the scores obtained from these measures have suggested the EARLI probes yield psychometrically adequate scores for measuring these early skills in preschool-aged children (DiPerna, Lei, & Morgan, 2009; Lei, Wu, DiPerna, & Morgan, 2009; Reid, Morgan, DiPerna, & Lei, 2006). Further research is needed, however, before they can be used in applied settings to assess the foundational skills necessary for future mathematics and reading success.

1.5. Structure of early academic skills

Examining the underlying factor structure of early academic measures provides insight regarding the relations among these skills in preschool children. Although research on early academic skills and their relation to future success in school has been accumulating, the exact nature of these skills in preschool children is not entirely known. One possible way to explore the nature of this construct is to test various hypothetical models using confirmatory factor analysis (CFA). This method allows multiple relations to be tested simultaneously and to determine which structure best describes the nature of early academic skills. In the present study, four hypothetical models were tested to explain the relations among preschool literacy and numeracy skills.

In the first model, we tested whether early academic skills could be best described by one general factor accounting for relations among all literacy and numeracy indicators. The developmental progression of both literacy and numeracy is similar during the preschool years, in that each involves learning increasingly complex units (i.e., letters to syllables and numbers as unbreakable categories to numbers as composites of smaller numbers) and later the ability to segment these larger units into their constituent parts (Sarama, Lange, Clements, & Wolfe, 2012). Likewise, the ability to identify individual letters and numbers are highly related skills in preschool (Piasta, Purpura, & Wagner, 2010). This common factor may represent early language skills, as the ability to understand many mathematical concepts (e.g., less and more) as well as numbers, are contingent on the understanding of appropriate language terms (Purpura, Hume, Sims, & Lonigan, 2011). It has also been theorized that children learn the names of numbers prior to understanding that each verbal representation has a quantity attached to it (Krajewski & Schneider, 2009).

In contrast to a unitary early academic skills construct, it may be that literacy and numeracy represent separate, albeit highly related, factors in early childhood; therefore, the second model we tested was a two-factor model in which correlated Early Literacy and Early Numeracy factors accounted for their respective tasks on the EARLI measures. In an examination of the structural and predictive validity of an early literacy and numeracy battery, Betts, Pickart, and Heistad (2009) found that a two-factor model better explained the early academic skills of kindergartners than a one-factor model. Although large cross-correlations were found between a literacy and a numeracy factor ($r = .88$), this model provided a substantially better fit than a single-factor model.

Other research has corroborated the finding that reading and mathematics skills are highly interrelated in young children (e.g., McClelland et al., 2007) and that early literacy skills are predictive of the development of future early numeracy skills (Kleemans, Peeters, Segers, & Verhoeven, 2012). There is evidence to suggest, however, that numeracy and literacy are distinguishable constructs at the preschool level. Research from cognitive psychology, for example, has suggested that mathematical ability is present well before verbal ability and that humans may even be born with an innate set of quantitative abilities, including quantity discrimination, basic arithmetic, and counting (Geary, 1995). During the preschool years, these innate abilities become integrated within children's increasing language skills to produce verbal counting ability (Geary, 2000).

For the third hypothesized model, we explored the possibility of residual variation among indicators in the two-factor model that was not accounted for by the Early Literacy and Early Numeracy factors. Betts et al. (2009) discovered significant residual variation among the indicators in their two-factor model of an early literacy and numeracy battery, as evidenced by statistically significant correlations between error terms within their main factors. Although the EARLI measures were developed to have a simple two-factor structure without correlated errors, Bentler and Chou (1987) have stated that forcing all errors to be uncorrelated in structural modeling is often inappropriate when using real data. Correlated errors can indicate item overlap, similar task demands, measurement error, or different response styles to items (see Brown, 2003).

Rather than correlated error terms, Betts, Pickart, and Heistad (2011) proposed that a bifactor model in which relations among individual literacy and numeracy skills were accounted for by a third general factor may better explain this residual variation. Thus, the fourth model we tested was a replication of this model in which two uncorrelated factors, Early Literacy and Early Numeracy, explained their respective sets of indicator variables. In addition, residual variation among literacy and numeracy skills was accounted for in the model by an unobserved general factor. Although Betts et al. did not explicitly examine the nature of this third factor, they speculated it could possibly represent general cognitive ability. Executive functioning and certain cognitive skills have been implicated in the development of both literacy and numeracy skills equally, while others, such as visual-spatial skills, have been shown to be predictive of future mathematics achievement specifically (Bull, Espy, & Wiebe, 2008).

1.6. Purpose

The purpose of this study was to examine the structure of the EARLI probe scores via confirmatory factor analysis (CFA). Examining the underlying factor structure of early academic measures serves two main purposes. The first is to determine that scores meet appropriate validity standards and accurately measure the intended constructs of interest (American Educational Research Association, 1999, American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999). Specifically, the internal structure of an assessment should align with past research and theory regarding the measured constructs (Messick, 1989). Four potential models were tested based on past research on hypothesized relations between literacy and numeracy skills in early childhood.

Factor analytic work can also provide insight regarding the relations among various academic skills in preschool children—both within and across the domains of literacy and numeracy. Using this knowledge to guide the development of measures can help ensure more valid measurement of the constructs of interest. Finally, determining the relations among early literacy and numeracy skills can help guide instruction and intervention efforts.

2. Method

2.1. Participants

The sample was drawn from a 3-year study of children enrolled in a Head Start program. Only data for children who participated in the third year of the study were included for analysis ($N = 289$). The sample consisted of children ranging from 3 ($n = 113$) to 4 ($n = 176$) years of age; their ages ranged from 43 to 67 months ($M = 56.94$ months, $SD = 3.90$ months). The sample was approximately equal in terms of gender, with slightly more boys (54.3%) than girls (45.7%). Participants represented a variety of racial and ethnic backgrounds: 45.0% Hispanic/Latino, 40.8% White, 5.9% Black/African American, 5.9% Biracial/ Multiracial, and 2.4% Asian/Other. The primary language spoken in the majority of homes was English (77.2%) followed by Spanish (17%) and other languages (4.2%). Overall, the sample was representative of the Head Start population from the surrounding community.

2.2. Measures

2.2.1. The Early Arithmetic, Reading, and Learning Indicators (EARLI)—The EARLI measures (DiPerna et al., 2007) consist of 12 probes: six designed to measure early literacy skills (e.g., phonological awareness, vocabulary, and letter knowledge) and six designed to assess early numeracy skills (e.g., number identification and counting aloud). For measures requiring children to look at a visual stimulus (i.e., Expressive Vocabulary, Letter Naming, Counting Objects, Number and Shapes, Grouping, Pattern Recognition, and Measurement), individual items were presented using a laptop and software (Microsoft Powerpoint) to standardize the timing and presentation of each item.

2.2.2. Literacy measures—Alphabet recitation assesses familiarity with the alphabet. Children are asked to recite the letters of the alphabet quickly and in order. If no response is provided within 5 s, examiners prompt children with the first three letters of the alphabet (“A...B...C...”). If an additional 5 s elapse without a response, the examiner provides the child with the fourth letter of the alphabet (“D...”). The prompt “Go on” also is used to encourage the child after the initial two prompts have been provided (or a child reaches “D”). If the child does not respond within 15 s of a prompt, or reaches the 1-minute time limit, then the probe is discontinued. The overall Alphabet Recitation raw score represents the number of letters correctly recited.

Expressive Vocabulary assesses vocabulary knowledge by asking children to state the names of picture of common objects (e.g., blanket, plane, and nose). Using a laptop, children are first presented with a practice item. Testing continues if the child responds correctly. If the child's response is incorrect, the examiner points to the object, provides the name, and

proceeds with the test items. No further feedback is provided. Total administration time for this probe is 1 min, 30 s. The Expressive Vocabulary raw score represents the numbers of pictures correctly named.

Letter Naming assesses letter recognition skills by requiring children to verbally state the names of letters presented in isolation. Using a laptop, children are first presented with a practice item. If the child does not respond, the examiner points to the letter and names it. Children are not provided with corrective feedback after the practice item. If children respond by producing the sound of the letter, the response is scored as incorrect, but the child is reminded to state the letter name. This prompt is provided for a maximum of two items. Total administration time for this probe is 1 min, 25 s. The Letter Naming raw score represents the number of letters correctly named.

Letter Sounds assesses knowledge of letter–sound associations by presenting children with a letter sound and asking them to state the associated letter. Children are presented with a practice item (e.g., “What letter makes the sound /r/, as in *ring*?”). If the child responds correctly, the examiner responds, “That’s right! The letter ‘r’ makes the /r/ sound in ring” and proceeds with the test items. If the child responds incorrectly to the practice item, the examiner responds, “Good try. The letter that makes the /r/ in ring is ‘r.’ Letter ‘r’ makes the /r/ sound.” No further corrective feedback is provided. If the child responds with the letter sound, the response is scored as incorrect, but a verbal prompt is provided reminding the child to state the letter name. This cue is provided for a maximum of two items. Test items are read aloud by the examiner from the administration and scoring manual. Total administration time for this probe is 50 s. The Letter Sounds raw score represents the number of correctly identified letters when given the letter sound.

Segmenting assesses the ability to identify sounds in words by requiring children to listen to a simple word and then state the first sound heard in the word. Children are presented with a practice item. The examiner instructs the child to listen as the examiner says the word *sip*. Then, the examiner provides the prompt, “What’s the first sound you hear in *sip*?” If the child responds correctly, the examiner proceeds to the test items. If the child responds incorrectly, the examiner repeats the word, breaks down the word into sounds, provides the correct response (e.g., “The first sound in *sip*, /s/-/i/-/p/, is /s/.”), and then proceeds with the test items. No correction is provided for test items. Test items are provided in the administration and scoring manual. Total administration time for this probe is 1 min 40 s. The segmenting raw score represents the total number of items in which children provide the correct sound that corresponds with the examiner’s prompt.

Sound Deletion assesses the ability to break down a word into its components by requiring children to listen to a word, mentally delete part of the word, and then state the remaining sound(s) or syllable aloud to the examiner. Children first are presented with a practice item (e.g., “Say *ladybug*. Now say *ladybug* without saying *bug*.”). If the child responds correctly, the examiner proceeds with the next practice item (e.g., “Say *neighbor*. Say *neighbor* without saying /bor/.”). If the child responds incorrectly to either practice item, corrective feedback is provided. No corrective feedback is provided on actual test items. Test items are provided in the administration and scoring manual. Total administration time for this probe

is 1 min 40 s. The Sound Deletion raw score represents the total number of items in which children correctly identify one part of a word when the other part is deleted.

2.2.3. Numeracy measures—Counting Aloud measures familiarity with numerical order by asking children to begin with the number 1 and orally count as high as they can within 1 min. If no response is provided within 10 s, examiners provide a prompt by counting “1...2...3...” and encouraging the child to finish counting. Testing is discontinued if the child hesitates for more than 10 s. The Counting Aloud raw score represents how many numbers children are able to identify in correct order.

Counting Objects measures ability to count using one-to-one correspondence. In this task, children are asked to count different sets of objects as quickly and accurately as they can. Using a laptop, children first are administered a practice item. A set of three identical circles are presented on the screen and the child is asked to count how many circles are on the screen. If the child responds correctly, the examiner proceeds to the test items. If the child responds incorrectly or fails to respond within 10 s, the examiner provides corrective feedback. Corrective feedback is not provided during the test items. Total administration time for this probe is 1 min 20 s. The Counting Objects raw score represents the total number of sets in which children correctly count the number of objects.

Grouping assesses the ability to subitize (quickly and accurately recognize the number of) objects in a small set. Using the laptop, the child is presented with two boxes, each containing one or more black dots. For the practice item, the child is asked to quickly tell the examiner how many dots are in each box. If the child responds correctly, the examiner proceeds to the test items. If the child responds incorrectly, the examiner provides corrective feedback. If the child counts the total number of dots across both boxes, the child is instructed to tell the examiner the number of dots in each box. This prompt is used for the first three test items only. Total administration time for this probe is 1 min 50 s. The Grouping raw score represents the total number of items in which children correctly identify the number of dots in each group.

Measurement is designed to assess knowledge of fundamental measurement concepts (e.g., bigger, smaller, more, and less). Using the laptop, children are presented with a practice item that requires them to identify which of three circles on the screen is bigger than the others. If the child responds correctly, the examiner proceeds to the test items. If the child responds incorrectly or does not respond within 10 s, the examiner provides corrective feedback and proceeds with the test items. (No corrective feedback is provided for the actual test items.) Directions and item questions are provided in the administration and scoring manual. Total administration time for this probe is 2 min. The Measurement raw score represents the total number of items in which children correctly identify bigger, smaller, more, or less.

Numbers & Shapes assesses number and basic shape identification skills. For this task, children are asked to name specific numbers and shapes presented individually (via Powerpoint). Using the laptop, the child is presented with a practice item. If the child responds correctly, the examiner proceeds to the test items. No corrective feedback is

provided on the test items. Total administration time for this probe is 1 min. The Numbers & Shapes raw score represents the total number of correctly identified numbers and shapes.

Pattern Recognition assesses a child's skill in recognizing and completing simple patterns using combinations of three shapes (circle, square, and triangle). Using the laptop, children first are presented a screen with the three shapes and asked to name each. Corrective feedback is provided. If the child correctly identifies all shapes, the practice item is administered. For the practice item, the child is presented with an incomplete pattern (e.g., [blank space], circle, triangle, and triangle) and asked to name which shape finishes the pattern. The examiner provides feedback if the child responds correctly or incorrectly and proceeds to the test items. If the child does not correctly identify all three shapes, alternative directions and procedures are followed. Specifically, the child is provided with a laminated sheet with the circle, square, and triangle response options. The child is asked to point to the shape that completes the pattern in the practice and test items. Corrective feedback only is provided if the examinee fails any of the first 3 items. Total administration time for this probe is 1 min 40 s. The Pattern Recognition raw score represents the total number of items in which children choose the correct shape to complete the pattern provided to them.

2.2.4. Scoring and technical adequacy—One of the original goals for the EARLI probes was to measure growth in early academic skills over time; thus, multiple forms of the EARLI skill probes were developed for use at different time points (i.e., Fall, Winter, and Spring) throughout the preschool year (DiPerna et al., 2009). During the development of the EARLI measures, raw scores across forms were scaled using 1-parameter logistic IRT (Rasch, 1960) concurrent calibration to account for any differences in item difficulty and ability level between age groups (i.e., 3- and 4-year-olds) and time points (Hochstedt, Lei, DiPerna, & Morgan, 2011; Lei et al., 2009). For the present study, however, only scores from a single (Spring) administration were analyzed. The use of IRT scores allows for the examination of the EARLI probes across age groups, as all responses are placed on a common scale based on ability level (Reise, Ainsworth, & Haviland, 2005).

Initial examinations of technical adequacy have indicated that scores from the EARLI literacy and numeracy probes demonstrate good reliability, with Cronbach's alpha coefficients typically exceeding .85. For the literacy measures, stability coefficients ranged from .59 (Segmenting) to .81 (Letter Naming) over a 3-month period and .37 (Alphabet Recitation) to .67 (Letter Naming) over a 6-month period. For the numeracy measures, stability coefficients ranged from .49 (Pattern Recognition) to .83 (Numbers and Shapes) over a 3-month period and .43 (Pattern Recognition) to .74 (Numbers and Shapes) over a 6-month period (DiPerna et al., 2009; Lei et al., 2009; Reid et al., 2006). In addition, the majority of the EARLI literacy and numeracy probe scores have demonstrated moderate to strong correlations with the Woodcock–Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001) Basic Reading Skills and Math Reasoning cluster scores, respectively (median $r = .56$ for Basic Reading Skills and median $r = .41$ for Math Reasoning; DiPerna et al., 2009).

2.3. Procedure

The EARLI measures were administered to participants during a 2-week period at three time points (October, January, and April) across each preschool year. Only data from April of the third year of the EARLI project were used in the present study. All examiners ($N = 6$) were formally trained and provided with manuals containing standardized administration instructions and explicit recording procedures. Training included an overview of each literacy probe, instruction in administration and recording, and opportunities to practice administration with feedback. Completed protocols also were periodically checked during each of the three data collection waves to promote correct administration, recording, and scoring.

Two Head Start preschool centers consented to participate, and letters were sent to parents of each inviting them to participate in the study. Consent was received for approximately 97% of the children across the two centers. Children with parental consent were brought by examiners to a quiet location outside of the classroom and asked if they were willing to work with the examiner. After children assented to their participation, they were tested individually in a quiet location outside of the classroom. Examiners then explained to the children that they would be asked some questions about letters and sounds, and children were offered stickers as a reward for trying their best on each task. Typical administration time was approximately 2 min per probe (12 minutes total) at each time point. All child responses were recorded and scored by hand as per guidelines outlined in a detailed manual provided to examiners. Scored protocols were periodically checked during each wave of data collection to ensure correct administration and scoring.

2.4. Analyses

Based on past research and the theoretical factor structure of the EARLI probes, a competing model approach was used to test four hypothetical models. The first of these models was a one-factor model in which the 12 EARLI probes served as indicators for a single Early Academic Skills factor. The second was a two-factor model in which the six EARLI literacy probe scores loaded on an Early Literacy factor, and the six numeracy probe scores loaded on an Early Numeracy factor. The correlation between the Early Literacy and Early Numeracy factors in this model was freely estimated. The correlation matrix of error terms was examined to determine which correlations, if any, were statistically significant ($p < .05$). The Lagrange Multiplier (LM) test was then used to determine which of these error correlations, if freely estimated, would result in better model fit. Based on this information, the two-factor model was again tested (Model 3), with correlations between select error terms freely estimated. Finally, the fourth model was a replication of the bifactor model tested in Betts et al. (2011) in which all 12 indicator variables (i.e., EARLI probe scores) were loaded on a single general factor, and the EARLI literacy and numeracy probe scores also loaded on uncorrelated Early Literacy and Early Numeracy factors, respectively. The chi-square difference test was used to compare the fit of the one- and two-factor models, the two-factor models with and without correlated error terms, and the two-factor model with correlated error terms with the bifactor model ($p < .05$).

CFAs were conducted using AMOS for Windows version 18.0. For purposes of scaling the factors, the factor loading of the probe score that had previously generated the strongest internal consistency reliability estimate per factor (determined via reliability analyses) was fixed to 1.0 for each model (Numbers & Shapes for the one-factor model; Letter Sounds and Numbers & Shapes for the two-factor and bifactor models). In addition, the loading from the general factor to the Letter Sounds probe score was fixed to 1.0 in the bifactor model. Path coefficients for the 12 associated error terms in all models were fixed to values of 1.0. All other parameters were estimated including factor variances and covariances as well as error terms.

The covariance matrix for the sample was factor-analyzed using maximum likelihood estimation. Several criteria were used to assess the goodness of fit of the hypothesized models. The first was the chi-square statistic; however, due to its sensitivity to sample size (Bentler & Bonett, 1980), the following criteria were also used to support good model fit, including (a) the ratio of chi-square to degrees of freedom less than 2 (Hair et al., 1995); (b) the comparative fit index (CFI $\geq .95$; Hu & Bentler, 1999); (c) the standardized root mean square residual (SRMR $\leq .06$; Hu & Bentler, 1999), and (d) the root mean square error of approximation (RMSEA $\leq .06$; Browne & Cudeck, 1993). The CFI measures the improvement of a hypothesized model over a null model, which assumes that there are no relations among the observed variables. The SRMR represents the average of the standardized residuals between the actual sample covariance matrix and the model-implied covariance matrix (Kline, 2005). The RMSEA is used to assess lack of fit for a hypothesized model while accounting for sample size and model complexity.

3. Results

3.1. Test of competing models

Table 1 provides the means, standard deviations, correlation matrix, and reliability estimates of the literacy and numeracy probe scores. Skewness values ranged from -0.89 to 0.94 , and kurtosis values ranged from -1.40 to 1.99 , indicating that the probe scores were approximately normally distributed. In addition, the majority of correlations between the EARLI probe scores were moderate (i.e., between $.30$ and $.50$). Cronbach's alpha coefficients for the Spring EARLI probe scores ranged from $.79$ to $.98$.

Table 2 presents a summary of the goodness of fit indices for the tested models. The one-factor model failed to meet any of the a priori fit criteria, $\chi^2(54) = 213.99$, $\chi^2/df = 3.96$, CFI = $.91$, SRMR = $.06$, RMSEA = $.10$, suggesting that early academic skills are not best described as a unitary construct, as measured in the present study. The two-factor model provided statistically significantly better fit to the data than the original one-factor model, $\chi^2(1) = 20.67$, $p < .01$, $\chi^2/df = 3.65$, CFI = $.92$, SRMR = $.05$, RMSEA = $.10$, although falling short of a priori fit criteria. The correlation between the Literacy and Numeracy factors was very large ($r = .91$), suggesting that while separate, these constructs are highly related in preschool.

An examination of the correlation matrix between error terms revealed several statistically significant ($p < .05$) correlations, suggesting that there was residual variation among

indicators in the two-factor model that was not accounted for by the Early Literacy and Early Numeracy factors. Results from the LM test suggested that model fit would improve by allowing the following error terms to covary: Alphabet Recitation and Expressive Vocabulary; Expressive Vocabulary and Letter Naming; Letter Sounds, Segmenting, and Sound Deletion; and Pattern Recognition and Grouping. This model (see Fig. 1) provided a statistically significantly better fit than the two-factor model without correlated error terms, $\chi^2(6) = 50.70, p < .01$. With the exception of the SRMR, all fit indices fell just short of a priori criteria, $\chi^2(47) = 142.62, \chi^2/df = 3.03, CFI = .94, SRMR = .04, RMSEA = .08$. Standardized factor loadings ranged from .44 to .85, and all were statistically significant ($p < .05$). The correlation between the Early Literacy and Early Numeracy factors was again large ($r = .97$).

The bifactor model provided the best fit of all four models, $\chi^2(42) = 139.60, \chi^2/df = 3.32, CFI = .94, SRMR = .04, RMSEA = .09$. This improvement in fit from the two-factor model with correlated errors was statistically significant, $\chi^2(6) = 13.87, p < .05$; however, there were some problems with this model. Specifically, three of the indicators (Alphabet Recitation, Letter Naming, and Expressive Vocabulary) had negative loadings on the Early Literacy factor, and the factor loading for the Measurement indicator on the Early Numeracy factor was not statistically significant. In an effort to improve this model, these four parameters were dropped from the model. The remaining literacy indicators (i.e., Letter Sounds, Segmenting, and Sound Deletion) were hypothesized to load on a Phonological Awareness factor, as each of these tasks require children to detect or manipulate the sound structure of letters and words. The final model is presented in Fig. 2. Making modifications to this model resulted in similar fit indices, but a statistically significant increase in the chi-square value, $\chi^2(4) = 12.51, p < .05, \chi^2/df = 3.31, CFI = .94, SRMR = .05, RMSEA = .09$. Likewise, although these fit indices were comparable to the two-factor model with correlated errors, the former model provided a statistically significantly better fit, $\chi^2(1) = 9.49, p < .01$. Nonetheless, making these modifications resulted in all factor loadings being positive and statistically significant ($p < .05$). The majority of loadings on the general factor were larger than the loadings on the Phonological Awareness and Early Numeracy factors, ranging from .44 to .82. Factor loadings on the Phonological Awareness factor ranged from .21 to .57, and loadings on the Early Numeracy factor ranged from .20 to .37.

4. Discussion

The purpose of the current study was to examine the underlying factor structure of the EARLI Literacy and Numeracy probes via confirmatory factor analysis (CFA). Four hypothetical models based on theory and prior research were tested: (a) a one-factor model with all literacy and numeracy task scores loading on an Early Academics factor, (b) a simple structure two-factor model with Early Literacy and Early Numeracy factors, (c) a two-factor model with correlated error terms, and (d) a bifactor model based on the work of Betts et al. (2011) that included an Early Literacy factor and an Early Numeracy factor along with a third, general factor. Although CFA results indicated that a two-factor model modified with correlated restraints and the bifactor model provided similar fit to the data, the former model better explained the data as evidenced by a statistically significantly lower chi-square value.

The present study extends the literature on early academic skill measurement with preschool children. Few studies have examined the construct validity of a literacy and numeracy battery in early childhood, as most have focused solely on one skill area or the other (e.g., Floyd et al., 2006; Townsend & Konold, 2010). Consistent with a study by Betts et al. (2009), which examined the factor structure of a comprehensive CBM battery with a sample of kindergarten children, a one-factor model provided poor fit to the EARLI data, whereas a model including correlated, but separate, Early Literacy and Early Numeracy factors yielded improved fit. The results of the current study also corroborate those of Betts et al. (2011) in that a two-factor model with correlated error terms provided an improved fit over a two-factor model where error correlations were constrained to be zero. Collectively, these results suggest that performance on early literacy and numeracy tasks is highly related in young children. The correlation between the Early Literacy and Early Numeracy factors in both of the two-factor models was very large. There was also residual variation among the EARLI indicators not accounted for by the Early Literacy and Early Numeracy factors, as evidenced by statistically significant correlations between error terms.

Initially, the bifactor model provided an improved fit over the two-factor model with correlated error terms. Introduction of the general factor eliminated the strong relation among the latent literacy and numeracy factors, which in turn, allows for more accurate estimation of pathways between unique (i.e., independent) latent constructs and their respective indicators. Likewise, consistent with the findings of Betts et al. (2011), the introduction of a general factor into the original two-factor model greatly reduced the magnitude of the pathways from the individual literacy and numeracy indicators to their respective latent factors. It is not surprising that a general factor accounted for significant variance within each of the EARLI academic skill indicators given the strong relation between the Early Literacy and Early Numeracy factors in the 2-factor model. However, for two literacy indicators (Alphabet Recitation and Letter Naming) and one numeracy indicator (Measurement), the addition of the general factor resulted in nonsignificant pathways with their respective latent factor. These results suggest that the individual Early Literacy or Early Numeracy factors do not account for unique variance in these indicators beyond the variance explained by the general factor. In addition, though still statistically significant, the direction of the loadings of three literacy (Alphabet Recitation, Expressive Vocabulary, and Letter Naming) indicator switched from positive to negative with the addition of the general factor. Deleting these parameters from the model resulted in positive and statistically significant factor loadings, but a small, albeit statistically significant degradation in model fit from the two-factor model with correlated errors.

In the final bifactor model, the general factor accounted for variation across all the EARLI measures, whereas a Phonological Awareness factor and an Early Numeracy factor accounted for variability not explained by this general factor. Given the negative factor loadings of the Alphabet Recitation, Expressive Vocabulary, and Letter Naming indicators in the original bifactor model, these tasks may be measuring skills that are more closely tied to the general factor and less involved in early literacy skills specifically. One possibility is that they measure verbal ability, as each of these tasks requires verbal memory or expressive language skills. Conversely, phonological awareness may represent an important early literacy skill independent of the general factor. Research findings have demonstrated that

phonological awareness predicts reading acquisition and achievement even when controlling for such factors as intelligence, language skills, and level of maternal education (Bryant et al., 1990).

Given the significant relations exhibited between the general factor and all of the EARLI measures, it is plausible that this factor represents the general factor of mental ability (*g*), which has been shown to correlate highly with academic performance (e.g., Jensen, 1998). On average, measures of the general factor have been shown to account for 25 percent of the variance in academic performance (Neisser et al., 1996). Thus, there are other factors that account for variability in performance in academic performance measures that are not accounted for by the general factor. Although beyond the scope of this study to examine the nature of these factors, some possibilities include socioeconomic status, opportunities for learning at home, attention, and motivation.

4.1. Limitations

As with any research endeavor, there are a few limitations of this study that should be addressed in future studies. First, participants were drawn from a single geographic location and from homes with fewer economic resources, and therefore, results may not generalize to preschool children from different backgrounds. Second, although the present sample included both 3- and 4-year-old children, the size of the sample was insufficient to directly test for developmental differences across these age groups. Finally, further research is necessary to clarify if the general factor specified in bifactor model employed in this study reflects the general factor (Jensen, 1998) or some other latent ability.

4.2. Implications

The results of the present study also have implications for the measurement of early academic skills in preschool. Given multiple independent studies with young children have found that a general factor accounts for significant variance in measures of early academic skills, future studies of such measures should test their structure via a bifactor approach. This approach, in turn, will enable users to better understand the constructs that are assessed by each measure. In addition, it will allow researchers to further refine their measurement tasks to more accurately assess specific constructs of interest beyond a general factor. In the case of the EARLI measures, a few of the indicators did not share latent variance with the other literacy or numeracy indicators beyond that explained by the general factor (i.e., Alphabet Recitation, Letter Naming, and Measurement). As such, additional studies are necessary to determine if the variance specific to these measures is useful in further understanding the early academic skills of young children.

Given the EARLI measures are intended to assist with periodic progress monitoring and the early identification of children at-risk for academic difficulty, further study also is necessary to determine the utility of the EARLI measures in predicting future skill development. Research examining sensitivity of the EARLI measures to growth throughout the preschool year also will provide further insight regarding the effectiveness of these measures for progress monitoring purposes. Replication of the current findings with preschool children

from various ethnicities, socioeconomic statuses, and geographic locations also will provide insight regarding generalizability of the current findings.

References

- American Educational Research Association. Standards for educational and psychological testing. Author; Washington, DC: 1999.
- Anthony JL, Lonigan CJ. The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology*. 2004; 96:43–55. <http://dx.doi.org/10.1037/0022-0663.96.1.43>.
- Anthony JL, Lonigan CJ, Burgess SR, Driscoll K, Phillips BM, Cantor BG. Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology*. 2002; 82:65–92. <http://dx.doi.org/10.1006/jecp.2002.2677>. [PubMed: 12081460]
- Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*. 1980; 88:588–606. <http://dx.doi.org/10.1037/0033-2909.88.3.588>.
- Bentler PM, Chou CP. Practical issues in structural modeling. *Sociological Methods & Research*. 1987; 16:78–117. <http://dx.doi.org/10.1177/0049124187016001004>.
- Betts J, Pickart M, Heistad D. Construct and predictive validity evidence for curriculum-based measures of early literacy and numeracy skills in kindergarten. *Journal of Psychoeducational Assessment*. 2009; 27:83–95. <http://dx.doi.org/10.1177/0734282908323398>.
- Betts J, Pickart M, Heistad D. Investigating early literacy and numeracy: Exploring the utility of the bifactor model. *School Psychology Quarterly*. 2011; 26:97–107. <http://dx.doi.org/10.1037/a0022987>.
- Brown TA. Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy*. 2003; 41:1411–1426. [http://dx.doi.org/10.1016/S0005-7967\(03\)00059-7](http://dx.doi.org/10.1016/S0005-7967(03)00059-7). [PubMed: 14583411]
- Browne, MW.; Cudeck, R. Alternative ways of assessing model fit.. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Sage; Newbury Park, CA: 1993. p. 445-455.
- Bryant PE, MacLean M, Bradley LL, Crossland J. Rhyme and alliteration, phoneme deletion, and learning to read. *Developmental Psychology*. 1990; 26:429–438. <http://dx.doi.org/10.1037/00121649.26.3.429>.
- Bull R, Espy KA, Wiebe SA. Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*. 2008; 33:205–228. [PubMed: 18473197]
- Charlesworth R. Prekindergarten mathematics: Connecting with national standards. *Early Childhood Education Journal*. 2005; 32:229–236. <http://dx.doi.org/10.1007/s10643-004-1423-7>.
- Clarke B, Shinn MR. A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*. 2004; 33:234–248.
- Clements, DH.; Stephan, M. Measurement in pre-K to grade 2 mathematics.. In: Clements, DH.; Sarama, J.; DiBiase, A-M., editors. *Engaging young children in mathematics: Standards for early childhood mathematics education*. Lawrence Erlbaum Associates; Mahwah, NJ: 2004. p. 299-317.
- Correa J, Nunes T, Bryant P. Young children's understanding of division: The relation between division terms in a noncomputational task. *Journal of Educational Psychology*. 1997; 90:321–329. <http://dx.doi.org/10.1037/0022-0663.90.2.321>.
- Cross, CT.; Woods, TA.; Schweingruber, H. *Mathematics learning in early childhood: Paths toward excellence and equity*. National Academies Press; Washington DC: 2009.
- Cunningham AE, Stanovich KE. Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*. 1997; 33:934–945. <http://dx.doi.org/10.1037/0012-1649.33.6.934>. [PubMed: 9383616]
- Dickinson DK, McCabe A. Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research & Practice*. 2001; 16:186–202.

- DiPerna JC, Lei P-W, Morgan PL. Development of Early Arithmetic, Reading, and Learning Indicators for Head Start (The EARLI Project). Semi-Annual Performance Report to the U.S. Department of Health and Human Services Administration for Children and Families (Award No. 90YF0047/01). 2009
- DiPerna JC, Morgan PL, Lei P-W. Development of Early Arithmetic, Reading, and Learning Indicators for Head Start (The EARLI Project). Semi-annual performance report to the U.S. Department of Health and Human Services Administration for Children and Families (Award No. 90YF0047/01). 2007
- Early Childhood Research Institute on Measuring Growth and Development. Individual growth and development indicators for preschool children: Picture naming/expressing meaning, rhyming/early literacy, alliteration/early literacy, picture naming in Spanish/expressive meaning. Center for Early Education and Development, University of Minnesota; Minneapolis, MN: 2000.
- Floyd RG, Hojnoski R, Key J. Preliminary evidence of the technical adequacy of the preschool numeracy indicators. *School Psychology Review*. 2006; 35:627–644.
- Foulon J. Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing*. 2005; 18:129–155. <http://dx.doi.org/10.1007/s11145-004-5892-2>.
- Geary DC. Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist*. 1995; 50:24–37. [PubMed: 7872578]
- Geary DC. From infancy to adulthood: The development of numerical abilities. *European Child & Adolescent Psychiatry*. 2000; 9:11–16. [PubMed: 10795851]
- Goffreda CT, DiPerna JC, Pedersen JA. Preventive screening for early readers: Predictive validity of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). *Psychology in the Schools*. 2009; 46:539–552. <http://dx.doi.org/10.1002/pits.20396>.
- Hair, JF., Jr.; Anderson, RE.; Tatham, RL.; Black, WC. *Multivariate data analysis with readings*. 4th ed.. Prentice Hall; Englewood Cliffs, NJ: 1995.
- Hochstedt KS, Lei PW, DiPerna JC, Morgan PL. Examining the dimensionality of EARLI literacy skill scores using nonlinear factor analysis. *Journal of Psychoeducational Assessment*. 2011; 29:395–406. <http://dx.doi.org/10.1177/0734282910384930>.
- Hojnoski RL, Silberglitt B, Floyd RG. Sensitivity to growth over time of the preschool numeracy with a sample of preschoolers in head start. *School Psychology Review*. 2009; 38:402–418.
- Hu LT, Bentler PM. Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- Jenkins JR, Hudson RF, Johnson ES. Screening for at-risk readers in a response to intervention framework. *School Psychology Review*. 2007; 36:582–600.
- Jensen, AR. *The g factor: The science of mental ability*. Praeger; Westport, CT: 1998.
- Jordan NC, Glutting J, Ramineni C. The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*. 2010; 20:82–88. <http://dx.doi.org/10.1016/j.lindif.2009.07.004>. [PubMed: 20401327]
- Jordan NC, Kaplan D, Locuniak MN, Ramineni C. Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*. 2007; 22:36–46. <http://dx.doi.org/10.1111/j.1540-5826.2007.00229.x>.
- Jordan NC, Kaplan D, Ramineni C, Locuniak MN. Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*. 2009; 45:850–867. <http://dx.doi.org/10.1037/a0014939>. [PubMed: 19413436]
- Jordan NC, Levine SC. Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews*. 2009; 15:60–68. <http://dx.doi.org/10.1002/ddrr.46>. [PubMed: 19213011]
- Juel C. Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*. 1988; 80:437–447. <http://dx.doi.org/10.1037/0022-0663.80.4.437>.
- Justice LM, Invernizzi MA, Meier JD. Designing and implementing an early literacy screening protocol: Suggestions for the speech–language pathologist. *Language, Speech, and Hearing Services in Schools*. 2002; 33:84–101. [http://dx.doi.org/10.1044/0161-1461\(2002/007\)](http://dx.doi.org/10.1044/0161-1461(2002/007)).

- Kendeou P, van den Broek P, White MJ, Lynch JS. Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*. 2009; 101:765–778. <http://dx.doi.org/10.1037/a0015956>.
- Kilpatrick, J.; Swafford, J.; Findell, B., editors. Adding it up: Helping children learn mathematics. National Academy Press; Washington, DC: 2001.
- Kim YS, Petscher Y, Foorman BR, Zhou C. The contributions of phonological awareness and letter-name knowledge to letter-sound acquisition—A cross-classified multilevel model approach. *Journal of Educational Psychology*. 2010; 102:313–326. <http://dx.doi.org/10.1037/a0018449>.
- Kleemans T, Peeters M, Segers E, Verhoeven L. Child and home predictors of early numeracy skills in kindergarten. *Early Childhood Research Quarterly*. 2012; 27:471–477. <http://dx.doi.org/10.1016/j.ecresq.2011.12.004>.
- Kline, RB. Principles and practice of structural equation modeling. 2nd ed.. New York, NY: Guilford Press: 2005.
- Krajewski K, Schneider W. Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction*. 2009; 19:513–526. <http://dx.doi.org/10.1016/j.learninstruc.2008.10.002>.
- Kroesbergen EH, Van Luit JE. Mathematics interventions for children with special educational needs a meta-analysis. *Remedial and Special Education*. 2003; 24:97–114. <http://dx.doi.org/10.1177/07419325030240020501>.
- Lei P-W, Wu Q, DiPerna JC, Morgan PL. Developing short forms of the EARLI numeracy measures: Comparison of item selection methods. *Educational and Psychological Measurement*. 2009; 69:825–842. <http://dx.doi.org/10.1177/0013164409332215>.
- Lonigan CJ, Burgess SR, Anthony JL. Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*. 2000; 36:596–613. <http://dx.doi.org/10.1037/0012-1649.36.5.596>. [PubMed: 10976600]
- Marston D, Fuchs LS, Deno SL. Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Assessment for Effective Intervention*. 1986; 11:77–90. <http://dx.doi.org/10.1177/073724778601100203>.
- McClelland MM, Cameron CE, Connor CM, Farris CL, Jewkes AM, Morrison FJ. Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*. 2007; 43:947–959. <http://dx.doi.org/10.1037/0012-1649.43.4.947>. [PubMed: 17605527]
- Mercer CD, Miller SP. Teaching students with learning problems in math to acquire, understand, and apply basic math facts. *Remedial & Special Education*. 1992; 13:19–35. <http://dx.doi.org/10.1177/074193259201300303>.
- Messick, S. Validity.. In: Linn, RL., editor. Educational measurement. 3rd ed.. Macmillan; New York: 1989.
- Methe SA, Hintze JM, Floyd RG. Validation and decision accuracy of early numeracy skill indicators. *School Psychology Review*. 2008; 37:359–373.
- Methe SA, Hojnoski R, Clarke B, Owens BB, Lilley PK, Politylo BC, et al. Innovations and future directions for early numeracy curriculum-based measurement: Commentary on the special series. *Assessment for Effective Intervention*. 2011; 36:200–209. <http://dx.doi.org/10.1177/1534508411414154>.
- Missall KN, Carta JJ, McConnell SR, Walker D, Greenwood CR. Using individual growth and development indicators to measure early language and literacy. *Infants and Young Children*. 2008; 21:241–253. <http://dx.doi.org/10.1097/01.IYC.0000324553.85187.dc>.
- Missall KN, Reschly A, Betts J, McConnell SR, Heistad D, Pickart M, et al. Examination of the predictive validity of preschool early literacy skills. *School Psychology Review*. 2007; 36:433452.
- National Council of Mathematics Educators. [September 29, 2010] Principles and standards for school mathematics. 2004. from. <http://www.nctm.org/standards>
- National Council of Teachers of Mathematics. Principles and standards for school mathematics. NCTM; Reston, VA: 2000.

- National Early Literacy Panel. Developing early literacy: Report of the National Early Literacy Panel. National Institute for Literacy; Washington, DC: 2008.
- National Institute of Child Health and Human Development. Report of the National Reading Panel. U.S. Government Printing Office; Washington, DC: 2000. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)..
- Neisser U, Boodoo G, Bouchard TJ Jr, Boykin AW, Brody N, Ceci SJ, et al. Intelligence: Knowns and unknowns. *American Psychologist*. 1996; 51:77. <http://dx.doi.org/10.1037/0003-066X.51.2.77>.
- Phillips BM, Clancy-Menchetti J, Lonigan CJ. Successful phonological awareness instruction with preschool children: Lessons from the classroom. *Topics in Early Childhood Special Education*. 2008; 28:3–17. <http://dx.doi.org/10.1177/0271121407313813>. [PubMed: 26038613]
- Phillips BM, Lonigan CJ, Wyatt MA. Predictive validity of the Get Ready to Read! screener: Concurrent and long-term relations with reading-related skills. *Journal of Learning Disabilities*. 2009; 42:133–147. <http://dx.doi.org/10.1177/0022219408326209>. [PubMed: 19074622]
- Piasta SB, Purpura DJ, Wagner RK. Fostering alphabet knowledge development: A comparison of two instructional approaches. *Reading and Writing*. 2010; 23:607–626. <http://dx.doi.org/10.1007/s11145-009-9174-x>. [PubMed: 20563245]
- Polignano JC, Hojnoski RL. Preliminary evidence of the technical adequacy of additional curriculum-based measures for preschool mathematics. *Assessment for Effective Intervention*. 2012; 37:70–83. <http://dx.doi.org/10.1177/1534508411430323>.
- Pressley M, Rankin JL. More about whole language methods of reading instruction for students at risk for early reading failure. *Learning Disabilities Research & Practice*. 1994; 9:157–168.
- Purpura DJ, Hume LE, Sims DM, Lonigan CJ. Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of Experimental Child Psychology*. 2011; 110:647–658. <http://dx.doi.org/10.1016/j.jecp.2011.07.004>. [PubMed: 21831396]
- Rasch, G. Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research; Copenhagen, Denmark: 1960.
- Reid EE, Morgan PL, DiPerna JC, Lei PW. Development of measures to assess young children's early academic skills: Preliminary findings from a Head Start university partnership. *Insights on Learning Disabilities*. 2006; 3:25–38.
- Reise SP, Ainsworth AT, Haviland MG. Item response theory fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science*. 2005; 14:95–101. <http://dx.doi.org/10.1111/j.0963-7214.2005.00342.x>.
- Sarama, J.; Clements, DH. Early childhood mathematics education research: Learning trajectories for young children. Routledge; New York, NY: 2009.
- Sarama J, Lange AA, Clements DH, Wolfe CB. The impacts of an early mathematics curriculum on oral language and literacy. *Early Childhood Research Quarterly*. 2012; 27:489–502. <http://dx.doi.org/10.1016/j.ecresq.2011.12.002>.
- Shapiro ES, Solari E, Petscher Y. Use of a measure of reading comprehension to enhance prediction on the state high stakes assessment. *Learning and Individual Differences*. 2008; 18:316–328. <http://dx.doi.org/10.1016/j.lindif.2008.03.002>. [PubMed: 26347390]
- Snow, CE.; Van Hemel, SB., editors. Early childhood assessment: Why, what, and how. National Academies Press; Washington DC: 2008.
- Stein, M.; Silbert, J.; Carnine, D. Designing effective mathematics instruction: A direct instruction approach. Merrill/Prentice Hall; Upper Saddle River, NJ: 1997.
- Storch SA, Whitehurst GJ. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*. 2002; 38:934–947. <http://dx.doi.org/10.1037/0012-1649.38.6.934>. [PubMed: 12428705]
- Sulzby, E.; Teale, W. Emergent literacy.. In: Barr, R.; Kamil, M.; Mosenthal, P.; Pearson, PD., editors. *Handbook of reading research*. Vol. 2. Longman; New York, NY: 1991. p. 727-758.
- Teale, WH.; Sulzby, E., editors. *Emergent literacy: Writing and reading*. Ablex; Norwood, NJ: 1986.
- Torgesen JK. The prevention of reading difficulties. *Journal of School Psychology*. 2002; 40:7–26. [http://dx.doi.org/10.1016/S0022-4405\(01\)00092-9](http://dx.doi.org/10.1016/S0022-4405(01)00092-9).

- Townsend M, Konold TR. Measuring early literacy skills: A latent variable investigation of the phonological awareness literacy screening for preschool. *Journal of Psychoeducational Assessment*. 2010; 28:115–128. <http://dx.doi.org/10.1177/0734282909336277>.
- Treiman R, Kessler B, Bick S. Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. *Cognition*. 2003; 88:49–78. [PubMed: 12711153]
- Van Luit JEH, Schopman EAM. Improving early numeracy of young children with special educational needs. *Remedial and Special Education*. 2000; 21:27–40. <http://dx.doi.org/10.1177/074193250002100105>.
- VanDerHeyden AM, Broussard C, Cooley A. Further development of measures of early math performance for preschoolers. *Journal of School Psychology*. 2006; 44:533–553. <http://dx.doi.org/10.1016/j.jsp.2006.07.003>.
- VanDerHeyden AM, Broussard C, Fabre M, Stanley M, Legendre J, Creppell R. Development and validation of curriculum-based measures of math performance for preschool children. *Journal of Early Intervention*. 2004; 27:27–41. <http://dx.doi.org/10.1177/105381510402700103>.
- Wilson SB, Lonigan CJ. Identifying preschool children at risk of later reading difficulties: Evaluation of two emergent literacy screening tools. *Journal of Learning Disabilities*. 2010; 43:62–76. <http://dx.doi.org/10.1177/0022219409345007>. [PubMed: 19822699]
- Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock–Johnson III tests of achievement. Riverside Publishing; Itasca, IL: 2001.

Fig. 1.

Confirmatory factor analysis two-factor model with correlated error terms ($N = 289$), standardized estimates. ABC = Alphabet Recitation; EV = Expressive Vocabulary; LN = Letter Naming; LS = Letter Sounds; S = Segmenting; SD = Sound Deletion; CA = Counting Aloud; M = Measurement; CO = Counting Objects; NN = Number Naming; PR = Pattern Recognition; G = Grouping.

Fig. 2.

Confirmatory factor analysis modified bifactor model ($N = 289$), standardized estimates. Error terms associated with each EARLI probe score are not pictured. ABC = Alphabet Recitation; EV = Expressive Vocabulary; LN = Letter Naming; LS = Letter Sounds; S = Segmenting; SD = Sound Deletion; CA = Counting Aloud; M = Measurement; CO = Counting Objects; NN = Number Naming; PR = Pattern Recognition; G = Grouping.

Correlation Matrix, Means, Standard Deviations, and Reliability Estimates of Early Arithmetic, Reading, and Learning Indicators Probe Item Response Theory Scores

Table 1

EARLI score	1	2	3	4	5	6	7	8	9	10	11	12	M	SD	Coefficient alpha
1. Alphabet Recitation													-0.11	1.06	.98
2. Expressive Vocabulary	.35												-0.21	0.99	.79
3. Letter Naming	.39	**											-0.33	1.13	.94
4. Letter Sounds	.31	**	**										-0.37	1.02	.87
5. Segmenting	.27	**	**	**									-0.31	0.96	.96
6. Sound Deletion	.29	**	**	.51	.63	.59							-0.39	1.04	.94
7. Counting Aloud	.39	**	**	**	.54	.48	.50						-0.34	0.98	.97
8. Measurement	.22	**	**	.42	.52	.46	.57	.40					-0.33	1.07	.81
9. Counting Objects	.41	**	**	.47	.48	.43	.47	.53	**				-0.45	1.19	.87
10. Number Naming	.34	**	**	.72	.63	.52	.64	.50	**	**			-0.41	1.05	.94
11. Pattern Recognition	.22	**	**	.37	.44	.40	.41	.36	**	**	**		-0.24	0.94	.90
12. Grouping	.13	*	**	.41	.55	.52	.48	.46	**	**	**	**	-0.17	0.98	.96

Note. $N = 289$. IRT scores have a mean of zero and a standard deviation of one.

*
 $p < .05$

**
 $p < .01$

Table 2

Summary of Fit Indices from Confirmatory Factor Analysis of the Early Arithmetic, Reading, and Writing Indicators Probe Item Response Theory Scores

Model	χ^2	df	χ^2/df	CFI	SRMR	RMSEA
One-Factor	213.99	54	3.96	.91	.06	.10
Two-Factor	193.32	53	3.65	.92	.05	.10
Two-Factor with Correlated Error Terms	142.62	47	3.03	.94	.04	.08
Bifactor (with Modifications)	152.11	46	3.31	.94	.05	.09

Note. $N = 289$. CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation.