

Increased Implementation of Emergent Literacy Screening in Pre-Kindergarten

Marcia Invernizzi · Timothy J. Landrum ·
Avery Teichman · Monika Townsend

Published online: 22 January 2010
© Springer Science+Business Media, LLC 2010

Abstract The recent emphasis on universal prekindergarten programs, coupled with ongoing concern about children's early literacy development in the first years of school, carry significant implications for how preschool educators program for literacy and assess literacy growth. We discuss the evolving role of literacy development in expanded preschool, examine key features of literacy assessments with particular attention to the content that might be assessed in preschool, and provide as an example one preschool literacy assessment tool that is in widespread use in the United States. We discuss the policy and practice implications of the growing emphasis on literacy assessment in preschoolers.

Keywords Emergent literacy · Literacy screening · Preschool literacy · Literacy development

The past 10 years signify a major educational shift toward establishing universal pre-kindergarten (Pre-K) programs for four-year-olds. An important impetus for this policy

shift is the measurable impact Pre-K programs have on early literacy skills and general academic success. Researchers in England have recently reported that the benefits of attending a good preschool program include improved math, reading, and social skills, and that these benefits last for several years, giving children a “leg up” when they enter elementary school (Melhuish et al. 2008). Similarly, analyses in the United States (US) consistently demonstrate that students participating in publically funded preschool programs gain foundational emergent literacy skills essential for kindergarten preparedness and beyond (Virginia Office of Early Childhood Development 2009). Indeed, many states are initiating programs for four-year-olds in efforts to decrease the number of children entering kindergarten at-risk for reading difficulty and delay. As preschool programs become more prevalent, interest in emergent literacy assessment grows. To date, assessment in the US has been the focus of attention mostly from an accountability perspective (e. g., the No Child Left Behind Act). Only recently have educators and policy-makers shifted emphasis away from a high-stakes accountability perspective toward the instructional utility of emergent literacy assessment data, particularly for children as young as 4 years of age.

The new and growing focus on literacy assessment for preschool age children demands a closer examination of several issues related to the nature, purpose, and use of data from early literacy screening. In this paper we (a) discuss the evolving role of early literacy assessment in expanded, if not universal, preschool; (b) examine several of the key features of a preschool literacy assessment, including both general criteria for early literacy assessment instruments, as well as the content that might be assessed in preschoolers; and (c) provide as an example one particular preschool literacy assessment tool that is in widespread use in the US

M. Invernizzi (✉) · A. Teichman
Curry School of Education, University of Virginia, P.O. Box
800785, Charlottesville, VA 22908-8785, USA
e-mail: mai@virginia.edu

T. J. Landrum
College of Education and Human Development,
University of Louisville, Louisville, KY 40292, USA
e-mail: t.landrum@louisville.edu

M. Townsend
Education, Human Development, and The Workforce,
American Institutes for Research, 1000 Thomas Jefferson St.,
NW, Washington, DC 20007, USA
e-mail: mtownsend@air.org

(the Phonological Awareness Literacy Screening for Preschoolers [PALS-PreK], Invernizzi et al. 2004b). We conclude with recommendations for both policy and practice.

The Role of Early Literacy Assessment in Expanded Preschool

One force driving increased attention to early literacy assessment may be the growing recognition of emergent literacy development as a predictor of later school success, but acknowledgement of this relationship has been slow to develop. The very term *emergent literacy* is a relatively new one that evolved in response to evidence that literacy development occurs along a continuum that begins long before children actually start formal schooling and long before they acquire conventional literacy skills such as decoding, oral reading, reading comprehension, spelling, and writing (Lonigan 2006). Recent research in emergent literacy has focused on defining which pre-conventional skills and abilities are the precursors to conventional literacy development that occurs after formal instruction begins. Identifying the foundational precursors that develop between the ages of birth to 5 and prior to the acquisition of conventional literacy skills can help early childhood educators facilitate their development and ensure a smoother transition into formal literacy instruction in the elementary grades. A recent meta-analysis conducted by the National Early Literacy Panel (NELP 2008) in the US identified a wide array of emergent skills and abilities that predict later conventional reading outcomes, even after the influence of other variables such as IQ or socio-economic status (SES) have been accounted for. These emergent skills pertain to three broad areas: (a) *phonological sensitivity*, or children's developing awareness of sound segments within spoken language independent of meaning; (b) *alphabet knowledge*, or children's knowledge of the names and sounds associated with printed letters; and (c) *print knowledge*, or children's ability to use their alphabet knowledge to write their name or other words using phonetic spelling. These broad areas, in addition to personal, social, and emotional development, language, and communication, are also represented in the curriculum for public preschools in other countries, as in England's National Strategies for the Early Years Foundation Stage (EYFS 2009).

Of course emergent literacy skills develop in different children at different rates. Developmental variability is attributable to various aspects of children's environment, inheritance, opportunity, and exposure to language, books, and print (along with the purposes for using them), as well as access to early childhood programs that promote them.

Nevertheless, the stunning diversity of children's levels of preparation presents a challenge for early childhood educators seeking to provide optimal learning experiences that meet the emergent literacy needs of their diverse learners. The use of assessments that provide information about children's development in the essential areas of emergent literacy can provide teachers with the information they need to plan optimal learning experiences. While assessment can provide information for a variety of purposes, the value of emergent literacy assessment can only be realized when the information is translated into instruction. Developmentally appropriate emergent literacy assessments have the potential to inform developmentally appropriate practices for diverse learners.

Two important policy publications brought national attention to emergent literacy as essential for school preparedness in the US. The first was a joint statement released by the National Association for the Education of Young Children (NAEYC) and the International Reading Association (IRA) entitled, *Learning to Read and Write: Developmentally Appropriate Practices for Young Children* (1998). This statement recognized that emergent literacy is (a) a domain of development in its own right, (b) worthy of the allocation of resources, and (c) capable of supporting its own set of standards. A second important policy publication was *Preventing Reading Difficulties in Young Children*, a synthesis of scientifically based research presented by the National Research Council, which placed emergent reading in a public health model of risk prevention, highlighting specifically the impact of early literacy on school outcomes (Snow et al. 1999).

Building upon this historical policy context, educators began to focus on three primary roles of preschool literacy assessment. First, emergent literacy assessment is essential for planning, implementing, and evaluating instruction. Second, assessment can serve a screening function, identifying children who might benefit from more intense instruction or additional instructional enhancements to prevent reading difficulties later in their educational experience (McAfee et al. 2004). Finally, assessment is a fundamental aspect of preschool program accountability. This latter purpose is particularly relevant in the current educational climate, which demands that preschool programs improve school readiness while promoting positive child outcomes. We consider briefly each of these three roles for preschool literacy assessment.

Improving Instruction

Preschool teachers are increasingly expected to implement more rigorous curricula in language and literacy, and to assess and document children's progress in developing essential skills (Bowman et al. 2001). They want to know

how to meet the needs of their children and best prepare them for successful transitions into kindergarten. The use of assessments that yield specific information about emergent literacy development can help teachers make curricular adjustments to meet individual children's needs and to specify instructional goals and objectives. Preschool teachers need to know the importance of phonological awareness, alphabet knowledge, and concepts about print. They must be able to interpret assessments and monitor the effectiveness of their instruction. In addition, preschool teachers must be able to use a variety of developmentally appropriate instructional strategies and adjust or differentiate these strategies to meet the specific needs of individual children. Emergent literacy assessments can be particularly useful to preschool teachers in meeting these increased demands if the assessments are: (a) broad-based, in that they examine skills across all dimensions of early literacy; and (b) instructionally transparent, in that assessment items and tasks mirror the types of instructional interactions teachers might have with students. Broad-based, instructionally transparent assessments may thus benefit both the teacher and the child, impacting both instruction and learning.

Screening

The term *screening* carries a pejorative tone if it is interpreted to mean "screening out," as in screening for children who don't meet certain benchmarks at expected points in time. However, screening can serve more wholesome purposes, some of which are critical to the iterative process of curricular modification and improvement to better meet children's needs. Preschool teachers might conduct universal, periodic emergent literacy screening to find out what their children know and what they need to learn, as well as to evaluate the developmental match of their own instruction. Routine screening may help identify areas of particular strength and instructional need, without necessarily 'labeling' a child.

While the practice of assessing emergent literacy in preschool is generally recommended, concerns remain about the nature of these assessments and the children's young age. For example, the International Reading Association, the National Association for the Education of Young Children (IRA/NAEYC 1998), and the National Education Goals Panel (NEGP 1998) disavow the use of group-administered, standardized, formal, multiple choice, paper-and-pencil tests before 3rd or 4th grade. Such tests are seen as developmentally inappropriate for young children and they are widely opposed to high-stakes testing, defined as a single measure of student learning used to make important educational decisions prior to the 3rd or 4th grade (AERA et al. 1999; IRA 1999; NAEYC 1987).

The IRA's central concern, as stated in their position statement *High-Stakes Assessment in Reading*, is that testing is a means of controlling and narrowing instruction rather than a way of gathering information to help teachers become better teachers, and children better readers. Meisels (2006) has also argued against high-stakes testing in early childhood because of potentially unintended but negative consequences on teaching and learning. To the extent that these concerns are valid, emergent literacy assessments that are instructionally transparent and use performance ranges for instructional goals, as opposed to hard and fast benchmarks, may be more appropriate. Developmental ranges accommodate the formative nature and variability of a child's early years.

Program Accountability

Emergent literacy assessments may be used to gauge preschool program effectiveness, but only to the extent that such assessments reveal student growth over time or target skills that are the focus of instruction. However, there is understandable concern that teachers will "teach to the test." A logical preemptive solution to this problem is to create developmentally appropriate, broad-based, and ecologically valid assessments that are worth "teaching to." Early childhood educators generally are wise to avoid assessments that may lead to decontextualized teaching of discrete skill sets in the absence of meaningful application and integration. Fragmented, skill-focused assessments may carry the risk that teachers will limit their curriculum, teaching only the decontextualized skills assessed. Emergent literacy assessments, therefore, must be broad-based, addressing all essential literacy fundamentals that lead to reading success, while also complementing and informing classroom instruction. We consider next the specific nature of early literacy assessments, focusing on both criteria for instrumentation and the content to be assessed at this early age.

Emergent Literacy Screening Criteria and Content

Criteria for Instrumentation

As is true for virtually any assessment used for educational purposes, emergent literacy assessments must meet minimally acceptable criteria in at least two broad areas: (a) the technical adequacy of the instrument, and (b) the instructional utility of the instrument.

First and foremost, emergent literacy assessments must comply with guidelines for establishing technical adequacy as set forth by the *Standards for Educational and Psychological Testing* (AERA et al. 1999). The technical

adequacy of an instrument refers to evidence supporting its reliability and validity. Reliability refers to the consistency of an instrument. For example, the consistency of the assessment results should be demonstrated between administrators (inter-rater reliability), between two halves of the same set of items (split-half reliability), and across time (test–retest reliability); moreover, items within the same instrument should perform similarly (internal consistency). Validity refers to the degree to which the assessment measures what it purports to measure. For example, assessment results can be compared to results obtained using other measures of the same construct (construct validity), or to results on another measure administered at the same time (concurrent validity) or at a future point in time (predictive validity). However, as Invernizzi et al. (2005) noted, tensions can exist between the practical, instructional priorities of teachers and standards for scientifically based research in the arena of educational assessments. That is, an instrument may be quite sound on technical grounds, but provide little instructional utility or benefit to teachers. Quality instruments must find ways to satisfy the needs of all stakeholders.

Justice et al. (2002) proposed another set of criteria against which early childhood assessments might be measured. Justice et al., and others, have suggested that the usefulness of early literacy protocols should be judged according to pragmatic criteria that affect the consequential validity of the assessment results (Invernizzi et al. 2004a). Consequential validity refers to the valid uses of assessment information to improve instruction. Specifically, Justice et al. proposed that to be useful to the field, early literacy assessments must be broad-based, sensitive, easily and efficiently administered, and must yield results that are easily interpreted.

The first criterion, broad-based, refers to the content of the assessment. Current theories suggest that emergent literacy consists of several domains, and a useful assessment system must strike a balance between covering as many essential content domains as possible without becoming cumbersome and compromising valuable instructional time. The second criterion, sensitivity, refers to the instruments' ability to differentiate students' individual literacy needs and to distinguish between those who are responding to the curriculum and instruction and those who are not. The third criterion, ease and efficiency of use, means simply that assessments should be easy enough for a range of early childhood professionals and paraprofessionals to conduct the assessments with minimal training, and to assess many children without significant disruptions to instructional routines. The last criterion, ease of interpretation, suggests simply that assessment results should have immediate and understandable instructional significance for users. Findings from The Center for the

Improvement of Early Reading Achievement (CIERA) “beat the odds” study support the importance of using instructionally transparent assessments. The CIERA study showed that student achievement improved as teachers began to understand assessment data and its appropriate use (Taylor et al. 1999).

Both sets of criteria, those related to technical adequacy and those related to the use of assessment tools, are consistent with Early Reading First (ERF) requirements in the US. ERF was funded with \$75 million in 2002 to transform existing preschools into centers of excellence by enhancing teacher practices, instructional content, and classroom environments to ensure young children start school with the skills needed for academic success (e.g., U. S. Department of Education 2007). ERF requires that emergent literacy assessments be: “(1) valid, reliable, and based on scientifically based reading research; and (2) a brief procedure designed as a first step in identifying children who may be at high risk” (No Child Left Behind Act of 2001) for later reading failure. This language exemplifies policy guidance for instrument development and implementation consistent with current practices and research standards.

The Content of Emergent Literacy Assessment

Research consistently supports several emergent literacy content domains and abilities as predictors of later reading success. These content areas are summarized in a number of reports and incorporated into policy guidelines as targets for which assessments should be developed and toward which instruction should be directed.

In 2000, the National Reading Panel's report, *Teaching Children to Read*, identified five content-area “pillars” for reading: fluency, phonological awareness, phonics, comprehension, and vocabulary. Although the report analyzed reading research for all ages, it did not focus specifically on emergent literacy for children birth to age 5. Fluency and reading comprehension are less relevant domains during emergent stages of literacy development, when greater emphasis is placed on developing oral language competencies, alphabet knowledge, and concepts about print. Early Reading First (ERF) helped narrow the national focus on reading in the US, requiring preschool programs to provide instructional activities in four areas: oral language, phonological awareness, print awareness, and alphabet knowledge (U. S. Department of Education 2003).

Given this backdrop of (a) established standards for assessments on technical grounds; (b) growing emphasis on making early literacy assessments broad-based, easy to administer and interpret, and instructionally useful; and (c) increased policy and practice guidance from both governmental initiatives (e.g., ERF) and a growing body of

research on emergent literacy (e.g., NELP 2008), we consider one example of an early literacy assessment tool that was developed specifically for preschoolers with an eye toward informing early literacy instruction. The Phonological Awareness Literacy Screening for Preschoolers (PALS-PreK; Invernizzi et al. 2004b) is now in widespread use in the US, and growing in use in other English-speaking countries. We briefly describe the instrument in the following sections, with particular attention to the ways that it might meet the growing professional consensus that is emerging regarding standards, purposes, developmental appropriateness, and uses for early literacy assessment tools.

The Phonological Awareness Literacy Screening for Preschoolers (PALS-PreK)

PALS-PreK was designed to measure children's developing literacy skills and to help teachers plan targeted literacy instruction. PALS-PreK was designed to be broad based, covering three of the four content domains specified by the National Early Literacy Panel (National Institute of Literacy 2008): alphabet knowledge, phonological awareness, print concepts, and writing. Given that the two best predictors of later reading success are (a) accurate recognition of the letters of the alphabet, and (b) an awareness of speech sounds (e.g., NELP), PALS-PreK was designed to assess both. PALS-PreK assesses the first of these using an Alphabet Knowledge task, and the second through tasks in Rhyme and Beginning Sound Awareness. Three additional tasks assess the ecological application of the two domains; these are Name Writing, Print and Word Awareness, and Nursery Rhyme Awareness.

Designed to meet criteria for being ecologically valid, developmentally appropriate, sensitive, and easy to administer and interpret, PALS-PreK consists of tasks with administration procedures that offer teachers examples of instructional practices in literacy competencies that are shown to be associated with later reading success. For example, during the Beginning Sound task, after the child provides his/her initial response, the teacher provides the correct answer by modeling *an/s/./m/*, and *b/* picture sort. The emphasis on modeling is intentional, and serves to reinforce the skill and assess children's learning; that is, the task assesses the child's response to teaching and provides teachers with an example of scaffolded instruction. Finally, PALS-PreK scores are interpreted against developmental ranges that are associated with later reading success. Developmental ranges help teachers gauge the need for additional enhancements. By comparing the scores of children in their classes to these developmental ranges, teachers can easily see what their

children know, what they still need to learn, and what the teacher needs to teach.

PALS-PreK is administered in a one-on-one setting, preferably by the classroom teacher or aide. There is no time limit; thus, the amount of time required to administer the assessment will vary. For most children, the entire assessment can be completed in 25 min. It is recommended that teachers break up the assessment, so that different tasks are administered on different days, but complete the entire assessment within a contiguous 2 week period. In the following sections we describe (a) the tasks included in PALS-PreK, (b) evidence of the technical adequacy of the instrument, and (c) the results of an external review of PALS-PreK.

Alphabet Knowledge

The alphabet knowledge task consists of three separate, conditionally sequenced tasks: upper-case alphabet, lower-case alphabet, and letter sounds. First, the teacher asks the child to name the 26 upper-case letters of the alphabet presented on a single card in random order. Children who know 16 or more upper-case letters also take the lower-case alphabet recognition task. The lower-case alphabet task is similar, and involves showing the child 26 lower-case letters also presented in random order on a single card. Children who know 9 or more of these lower-case letters are then asked to produce the sounds associated with 23 letters and 3 consonant digraphs (*ch*, *sh*, *th*). The stimuli for the Letter Sounds task are upper-case letters because children know more of these than lower-case letters. This method of measuring alphabet knowledge is designed to avoid frustration for children who know few letters and to avoid ceiling effects for children who demonstrate more advanced alphabet knowledge. Thresholds between tasks were based on initial analyses of pilot data, which indicated that if children knew at least 16 upper-case letters, they would also know some lower-case letters. Inter-rater reliabilities for alphabet knowledge are consistently high, and significant, ranging from .96 to .99. Test-retest reliability was estimated to be .98 for upper-case alphabet. Test-retest reliability was lower for lower-case alphabet ($r = .85$) and letter sounds ($r = .84$), both of which had lower sample sizes due to administration protocol.

Phonological Awareness

Tasks measuring phonological awareness and related skills in PALS-PreK are Beginning Sound Awareness, Rhyme Awareness, and Nursery Rhyme Awareness. The Beginning Sound Awareness task is an oral production task. In this task, the teacher says the name of a picture and asks the child to produce the beginning sound for words that start

with/s/,/m/and/b/. Each item is scored based on the child's first oral response; then the correct answer is modeled by the teacher who then sorts the picture by its beginning sound to reinforce the concept. In the Rhyme Awareness task, the teacher names each of four pictures on a page, one of which is the stimulus and three of which are the response options. The teacher asks the child to point to the picture that rhymes with the stimulus. In the Nursery Rhyme Awareness task, the teacher recites a familiar nursery rhyme, but stops before the end for the child to supply the final rhyming word. For example, the child supplies the missing word in the phrase, "Jack and Jill went up the ____."

Both inter-rater and test–retest reliability estimates for the phonological awareness tasks are included in the PALS-PreK Teacher's Manual. For Beginning Sound, inter-rater reliability was .99 and test–retest reliability at 2 weeks was found to be .84. Rhyme Awareness results were similar, with inter-rater reliability estimated at .96 and test–retest reliability estimated at .88. Nursery Rhyme Awareness inter-rater reliability was estimated at .99 while test–retest reliability was .87.

Name Writing and Print and Word Awareness

In the Name Writing task, the child is asked to draw a self-portrait and to write his or her name. The child's name writing effort is scored on a seven-point scale. Zero points are awarded for scribbles that represent both writing and picture combined, and increasing numbers of points are awarded for (a) clear separation of writing from the picture, (b) inclusion of letters, symbols, and letter-like forms, (c) inclusion of increasing numbers of correct letters, and (d) correct direction/orientation for letters and writing. Reliability estimates are high for both inter-rater reliability ($r = .97$) and test–retest reliability ($r = .82$).

The Print and Word Awareness task provides an ecologically valid means of assessing children's developing knowledge of the form and function of book parts (title and print conventions) as well as skill with pointing accurately to letters and words. In this task, the teacher reads a familiar nursery rhyme printed in a book format and asks the child to point to different print components. The task is administered in a shared book reading context suggestive of instruction that might occur during a read aloud. Estimates of reliability for this task are comparable to other tasks, with inter-rater reliability estimated at .97 and test–retest reliability estimated at .80.

Developmental Ranges

For each task, developmental ranges are provided in the Teacher's Manual to give educators a general sense of

Table 1 Spring developmental ranges and maximum scores

PALS-PreK tasks	Spring developmental range	Maximum score
Name writing	5–7	7
Upper-case alphabet	12–21	26
Lower-case alphabet	9–17	26
Letter sounds	4–8	26
Beginning sound	5–8	10
Print and word awareness	7–9	10
Rhyme awareness	5–7	10
Nursery rhyme awareness	6–10	10

developmental trajectories for four-year-old children on a path of literacy development associated with later successful reading. Developmental ranges are based on analysis of several data sources, including longitudinal analyses of student performance on the state-provided first grade literacy assessment 2 years later. Score ranges for each task were established by examining the PALS-PreK scores of children whose instructional oral reading level was found to be on or above grade level at the end of first grade. These ranges are applicable during the spring of the child's four-year-old preschool year, and are designed to provide evidence of children's literacy growth during the preschool year, and to provide some indication of kindergarten preparedness. Spring developmental ranges and maximum scores for PALS-PreK tasks are shown in Table 1.

Technical Adequacy

The technical adequacy of PALS-PreK is described extensively in the PALS-PreK Teacher's Manual (Invernizzi et al. 2004a), which includes discussion of evidence of the reliability and validity of the instrument. Reliability is discussed in terms of internal consistency and inter-rater reliability. Evidence of construct validity is discussed in terms of task intercorrelations, and predictive validity is discussed in terms of several analyses of the predictive power of PALS-PreK tasks in predicting students' future PALS identification status (i.e., above or below grade level benchmarks).

In addition to the inter-rater and test–retest reliability estimates mentioned previously, the internal consistency of items within PALS-PreK tasks was also assessed in pilot samples yielding internal consistency estimates (Cronbach's alpha) ranging from .75 to .93 across tasks.

The construct validity of PALS-PreK is supported by consistent patterns of intercorrelation among tasks. With the exception of obvious examples involving the alphabet, which would be expected to be high (e.g., the correlation

between upper case alphabet recognition and lower case alphabet recognition was .91), task intercorrelations are generally moderate, suggesting that these tasks do share significant variance with one another. Perhaps more importantly, we examine the extent to which this pattern of intercorrelation holds across subgroups of children (e.g., males vs. females; children of different race or ethnicity). The consistency of patterns observed across these groups supports that PALS-PreK tasks and items behave similarly across subgroups. The opportunity to analyze intercorrelation across socioeconomic subgroups does not present itself, given that the data are drawn from publicly funded preschools serving primarily low SES children.

The predictive validity of PALS-PreK was determined by examining the relationship between children's PALS-PreK scores and later performance in subsequent assessments using PALS-K in kindergarten and PALS 1-3 in first grade. However, it is important to note two cautions about such analyses. First, PALS-PreK was designed as a diagnostic tool to help guide teachers' literacy instruction. This stands in contrast to PALS-K and PALS 1-3, which were designed to serve a screening function in identifying children who need additional instruction beyond that which is typically provided in classroom literacy instruction. Second, given the particularly uneven nature of children's development at earlier ages, caution is necessary in interpreting the relationship between a formative, emergent literacy assessment like PALS-PreK and later literacy achievement (Bowman et al. 2001). For example, some children who initially achieve relatively low scores on PALS-PreK tasks will improve relatively quickly in response to instruction, as they progress through preschool and kindergarten. Despite these caveats, some degree of consistency between PALS-PreK and later literacy success is evident across time. For example, the correlations between summed scores on PALS-PreK, PALS-K and PALS 1-3 are moderate to high across time, ranging from .44 to .79 ($p < .001$).

The predictive power of PALS-PreK scores in relation to children's risk for reading difficulty as identified at later screenings using PALS-K and PALS 1-3 has been assessed using discriminant analyses. Again, extreme caution is in order in interpreting these analyses, as the functions for which the instruments were designed are different. Specifically, four separate discriminant function analyses were conducted to evaluate the extent to which PALS-PreK scores would predict Kindergarten risk status, defined as above or below PALS-K benchmarks 1 year later, both from fall PreK to fall Kindergarten and spring PreK to spring Kindergarten; and the extent to which PALS-PreK scores would predict first-grade risk status, defined as above or below PALS 1-3 first grade benchmarks 2 years

later, from fall PreK to fall first grade, and spring PreK to spring first grade.

The combination of children's PALS-PreK scores resulted in discriminant functions that accurately classified the risk status of more than 90% of children screened 1 year later in kindergarten ($\lambda = .875, p < .001$). Further, the classification of students' risk status 2 years later in first grade was only slightly less accurate; 86 and 84% of students were accurately classified in first grade based on their PALS-PreK scores from 2 years earlier ($\lambda = .843, p < .001$ and $\lambda = .901, p < .001$). In three of the four discriminant analyses, the upper-case alphabet recognition task once again contributed the most to the discriminant function. In the first analyses, predicting fall kindergarten performance from fall PALS-PreK performance, the Name Writing task contributed most, but even in this case, upper-case alphabet recognition was the second largest contributor.

Preschool Teacher Survey

In order to assess the perceptions of users with regard to ease of administration and interpretation, a survey was mailed to all preschool teachers in Virginia who entered PALS-PreK scores in the PALS Online Score Entry and Reporting System in 2003 ($n = 951$). Forty-one percent ($n = 387$) who entered scores returned completed surveys. Although this response rate limits generalizability, the majority of teachers who responded perceived PALS-PreK to be a sensitive, accurate measure that provides a comprehensive analysis of early literacy skills. Over 90% of respondents agreed that PALS-PreK instructions are easy to understand, that the instrument is easy to score, and the scores easy to interpret. The majority of teachers commented favorably on the sensitivity of PALS-PreK in identifying students' literacy strengths (84% of the respondents) as well as areas of literacy need (86% of respondents). Approximately 87% of teachers reported that PALS-PreK data assists their classroom literacy instruction. One Program Director commented: "[PALS-PreK] provides useful information because it lets you see how each child is doing [and] if they need more help in certain areas."

Results of Implementation

Although developed in Virginia, in large part to extend the reach of statewide literacy screening efforts from early elementary years into preschool, PALS-PreK is also in use in many states across the US and in three different countries. We describe briefly the expanding use of PALS-PreK in Virginia and beyond.

PALS-PreK Implementation

First implemented in Virginia in 2003, the PALS-PreK instrument is now among the most widely used state preschool assessments in the US In 2008–09; more than 1,400 preschool teachers used PALS-PreK to assess the emergent literacy development of more than 21,000 preschool children prior to kindergarten entry. Moreover, in the US, Early Reading First grant recipients, subject to the Government Performance and Reporting Act (GPRA 1993), must report programmatic performance targets and whether or not those targets are met on an annual basis to evaluate overall program effectiveness. Under the GPRA, ERF recipients are required to administer the PALS-PreK upper case alphabetic knowledge task (<http://www.ed.gov/programs/earlyreading/performance.html>). ERF provides competitive grants to US preschool programs, including public school districts and Head Start centers, increasing PALS-PreK use across the nation.

In Spring 2009, the eighth year of widespread implementation of PALS-PreK in Virginia, preschool teachers in 118 school divisions (89% of all school divisions) in 694 schools across the Commonwealth participated in PALS-PreK. Participating classrooms include those funded through various public sources, including Head Start, Title I, and Early Childhood Special Education, as well as those participating in the Virginia Preschool Initiative. Teachers using PALS-PreK enter screening results into the PALS website, which provides a secure, password-protected, encrypted Internet database. Teachers and preschool program directors have immediate access to a variety of interpretive reports designed to help teachers plan, deliver, and monitor instruction (c.f., Partridge et al. 2003).

Participant characteristics for the eighth preschool cohort are presented in Table 2. Males comprised slightly more than half the sample (51%, $n = 10,376$). Because

Table 2 Participants in eighth cohort (spring 2009) of PALS-PreK implementation ($N = 20,545$)

Category	<i>N</i>	Percentage of sample (%)
Gender		
Male	10,376	50.5
Female	10,169	49.5
Ethnicity		
Caucasian	7,915	39
African American	8,671	42
Hispanic	2,632	13
Asian & Pacific Islanders	519	3
Native American	38	<1
Other/mixed	770	4

PALS-PreK tasks and developmental ranges were designed with typically developing 4-year-olds in mind, we describe here only scores for children who had reached the age of 4 by September 30 of their preschool year. It is worth noting that teachers nonetheless sometimes administered PALS to children who were younger than this minimum age. Children screened in Spring 2009 (and who had reached the age of 4 by September 30) ranged in age from 4 years, 6 months, to 7 years, 3 months, with a mean age of 5 years, 1 month ($SD = 4$ months). The cohort represented a diverse group of students who mirrored the race and ethnicity of the preschool enrollment in Virginia: 39% of students ($n = 7,915$) were Caucasian, 42% ($n = 8,617$) were African American, 13% ($n = 2,632$) were Hispanic, 3% ($n = 519$) were Asian or Pacific Islanders, 4% ($n = 770$) were of other or mixed race and ethnicity, and fewer than 1% ($n = 38$) were Native American. A number (10%) of the participating children ($n = 2,111$) received additional school services based on identified special needs, including special education services for speech/language impairment (6%), developmental delay (4%), and autism (<1%). Four percent were designated English language learners. It is important to note that the Virginia Preschool Initiative, similar to Head Start and other publicly funded preschool programs, primarily enrolls children of low socioeconomic status. PALS-PreK data in Virginia therefore, disproportionately represent the emergent literacy skills of low-income preschoolers.

Descriptive Data on Use in Virginia

Descriptive statistics for PALS-PreK tasks from spring 2009 are presented in Table 3. In the statewide sample, preschool children achieved a mean of greater than 6 on the 7-point Name Writing task. In the spring of their 4-year-old program year, preschoolers were able to identify, on average, approximately 21 upper-case letters. Those who knew 16 or more upper-case letters (89% of those screened) were administered the lower-case alphabet recognition task, and students progressing to this task knew

Table 3 Statewide performance on PALS-PreK (spring 2009)

Section (total possible score)	<i>M</i>	<i>SD</i>	<i>N</i>
Name writing (7)	6.41	1.18	20,545
Upper case alphabet (26)	21.36	7.15	20,545
Lower case alphabet (26)	20.85	6.89	18,306
Letter sounds (26)	15.52	8.02	18,092
Beginning sound awareness (10)	8.24	2.78	20,469
Print and word awareness (10)	8.18	1.97	20,523
Rhyme awareness (10)	7.77	2.66	20,514
Nursery rhyme awareness (10)	8.17	2.18	20,531

Table 4 Statewide performance on PALS-PreK: group comparisons (spring 2009)

Task (developmental range)	Below range <i>M</i> (SD)	Meeting range <i>M</i> (SD)	Percent Meeting range (%)
Name writing (5–7)	3.10 (1.23)	6.70 (0.57)	92
Upper case alphabet (12–21)	5.85 (3.31)	23.82 (3.64)	86
Beginning sound awareness (5–8)	2.05 (1.56)	9.14 (1.46)	87
Print and word awareness (7–9)	4.59 (1.63)	8.91 (1.01)	83
Rhyme awareness (5–7)	2.88 (1.23)	8.65 (1.73)	85
Nursery rhyme awareness (6–10)	3.64 (1.47)	8.82 (1.34)	88

more than 20 lower-case letters. Further, children who knew 9 or more lower-case letters were administered the Letter Sounds task (99%), and these students identified the sounds that correspond to upper-case letters for more than 15 letters. In the spring of their 4-year-old year, preschool children identified, on average, more than 8 of the 10 beginning sounds presented on the Beginning Sound Awareness task, and correctly identified nearly 8 of the 10 rhyming words on the Rhyme Awareness task.

Spring developmental ranges for PALS-PreK are based on results associated with reading success in First grade, defined as reading on or above grade level. Table 4 contrasts the performance of children who met the spring developmental range on each task with those who did not. The data in Table 4 indicate that the greatest performance discrepancies between the two groups are observed in upper-case alphabet recognition (recognition of about 6 letters vs. nearly 24 letters) and Beginning Sound Awareness (accurate identification of beginning sounds on about 2 vs. 9 of the 10 Beginning sounds items).

Summary and Recommendations for Policy and Practice

Political emphasis on the use of research based, scientifically validated assessments and instructional techniques in all grades, including preschool, has increased dramatically in recent years. In the US, federal funding is currently offered through programs such as Early Reading First to support preschool efforts to provide scientifically based reading assessment and instruction. However, differences between the expectations for scientifically rigorous assessment practice and teachers' needs for a broad-based, sensitive assessment with tangible value for everyday classroom use contribute to a frequently cited disconnect between assessment and instruction (Taylor et al. 1999). In this paper we reviewed some of the issues that contribute to the discord that may naturally occur when scientific rigor in assessment and instruction, coupled with demands for accountability, intersect with the instructional needs of

teachers of children of preschool age. Indeed, it is at the preschool level that the potential for disconnect may be greatest.

In discussing criteria for early literacy assessment in preschoolers, we suggested that in addition to matters of technical adequacy, assessments must be broad-based, easy and efficient to administer and interpret, and must provide immediate instructional benefit to teachers. We described PALS-PreK as one example of an effort to meet these criteria. The widespread use of PALS-PreK and its inclusion in a number of federal policy initiatives in the US suggests that it may indeed meet both demands for instructional utility from the field, as well as demands for scientific rigor. External reviews have been especially favorable concerning the organization and ease of use of the instrument (e.g., Kaderavek 2006). Administrators and reading professionals may find this information particularly useful in their efforts to meet emergent literacy standards.

PALS-PreK is not the only emergent literacy assessment available, however, and many other assessment tools are in widespread use (c.f. Lonigan 2006). Researchers and practitioners are urged to be mindful of differences in assessment packages' purpose, design, and results interpretation before selecting assessments for use in preschool classrooms. Perhaps foremost is the fundamental question of why children need to be assessed in the first place. For example, is assessment intended to inform teachers' planning and delivery of instruction? Is the assessment a screener, intended to identify students who may need some additional intervention or service? Are assessment data to be used to evaluate the impact of a specific program or curriculum on children's literacy performance? Are assessment data to be used more broadly to evaluate preschool programs at a division or even state level? Important too is recognition that more than one assessment may be necessary if multiple assessment purposes must be met, but we caution again that the burden on teachers and children can become great if assessment takes on so great a role that it compromises instructional time. Clearly, different assessment packages provide different returns, and practitioners and administrators of preschool programs must be

wise in choosing assessments that best meet their programmatic needs.

References

- AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*. New York: AERA.
- Bowman, B. T., Donovan, M. S., & Burns, M. S. (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Research Council.
- Government Performance & Reporting Act. (1993). U. S. C. § 1101. Public Law 103-62.
- International Reading Association. (1999). High-stakes assessment in reading: A position statement of the International Reading Association. Retrieved from http://www.reading.org/Libraries/Position_Statements_and_Resolutions/ps1035_high_stakes.sflb.ashx.
- International Reading Association and National Association for the Education of Young Children. (1998). Learning to read and write: Developmentally appropriate practices for young children. *Young Children*, 53, 3–46.
- Invernizzi, M. A., Justice, L. M., Landrum, T. J., & Booker, K. C. (2004a). Early literacy screening in kindergarten: Widespread implementation in Virginia. *Journal of Literacy Research*, 36, 479–500.
- Invernizzi, M., Sullivan, A., Meier, J., & Swank, L. (2004b). *Phonological awareness literacy screening: Preschool (PALS-PreK)*. Charlottesville, VA: University of Virginia.
- Invernizzi, M. A., Landrum, T. J., Howell, J. L., & Warley, H. P. (2005). Toward the peaceful coexistence of test developers, policymakers, and teachers in an era of accountability. *The Reading Teacher*, 58(7), 610–618.
- Justice, L. M., Invernizzi, M. A., & Meier, J. D. (2002). Designing and implementing an early literacy screening protocol: Suggestions for the speech-language pathologist. *Language, Speech, and Hearing Services in Schools*, 33, 84–101.
- Kaderavek, J. (2006). *External review of PALS-PreK*, March 2, 2006.
- Lonigan, C. J. (2006). Development, assessment, and promotion of literacy skills. *Early Education and Development*, 17(1), 91–114.
- McAfee, O., Leong, D., & Bodrova, E. (2004). *Basics of assessment: A primer for early childhood professionals*. Washington D.C.: National Association for the Education of Young Children.
- Meisels, S. (2006). *Accountability in early childhood: No easy answers*. Occasional paper number 6, Erikson Institute: Herr Research Center for Children and Social Policy. [Online]. Available <http://www.erikson.edu/default/research/researchpubs.aspx>.
- Melhuish, E. C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. B., et al. (2008). Preschool influences on mathematics achievement. *Science*, 321(5893), 1161–1162.
- National Association for the Education of Young Children. (1987). *Developmentally appropriate practices in programs for four and five year olds* (Rev. Ed.). Washington, DC: National Association for the Education of Young Children.
- National Education Goals Panel. (1998). *National education goals panel recommendations regarding the implementation of standards*. [Online]. Available <http://www.negp.gov/page1-13-9.htm>.
- National Institute of Literacy. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. [Online]. Available <http://www.nifl.gov/earlychildhood/NELP/NELPShanahan.html>.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- No Child Left Behind Act of 2001, 20 U. S. C. § 6301. (2002).
- Partridge, H., Invernizzi, M., Meier, J. & Sullivan, A. (2003, Nov/Dec). Linking assessment and instruction via web-based technology: A case study of a statewide early literacy initiative. *Reading Online*, 7 (3).
- Snow, C. E., Burns, M. S., & Griffin, P. (1999). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Taylor, B. M., Pearson, P. D., Clark, K. F., & Walpole, S. (1999). *Beating the odds in teaching all children to read. CIERA Report #2–006*. Ann Arbor, MI: University of Michigan Center for the Improvement of Early Reading.
- The National Strategies. (2009). *Learning, playing, and interacting: Good practice in the Early Years Foundation Stage*. www.nationalstrategies.standards.dcsf.gov.uk.
- U. S. Department of Education. (2003). *Guidance for the Early Reading First program*. Subpart B, Part B, Title I of the Elementary and Secondary Education Act of 1965 (ESEA) as amended by No Child Left Behind Act of 2001. Pub. L. No. 107–110, 107th Congress (2001) Retrieved from www.ed.gov.
- U. S. Department of Education. (2007). *National evaluation of early reading first: Final report to congress*. Washington, DC: U. S. Department of Education.
- Virginia Office of Early Childhood Development. (2009). *School Readiness Report*. http://www.earlychildhood.virginia.gov/school_readiness.shtml.

Copyright of Early Childhood Education Journal is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.