LSHSS

Measuring Preschool Attainment of Print-Concept Knowledge: A Study of Typical and At-Risk 3- to 5-Year-Old Children Using Item Response Theory

Laura M. Justice
Ryan P. Bowles
Lori E. Skibbe
University of Virginia, Charlottesville

Emergent literacy skills are children's precursory abilities concerning reading that do not involve reading per se but that prepare the child for the task of learning to read and facilitate progress from beginning to conventional reading (e.g., Whitehurst & Lonigan, 1998). The emergent literacy period not only establishes children's rudimentary awareness of orthography and phonology as alphabetic tools, but also provides a foundation on which other increasingly abstract layers of alphabetic knowledge will build. Current interest in emergent literacy by educational practitioners and policymakers is fostered by a research base showing that children who arrive to beginning reading instruction with well-developed emergent literacy skills progress more rapidly and readily than those who do not have these skills (e.g., Bryant, Maclean, Bradley, & Crossland, 1990; Chaney, 1998; Morris, Bloodgood, Lomax, & Perney, 2003; Storch & Whitehurst, 2002; Walpole, Chow, & Justice, 2004), and that weaknesses in emergent literacy provide an early yet sensitive indicator of a child's vulnerability for experiencing reading disability (e.g., Catts, Fey, Tomblin, & Zhang, 2002; O'Connor & Jenkins, 1999; Stuart, 1995).

The more parsimonious models of emergent literacy development differentiate children's early knowledge bases concerning the alphabetic code into two major areas of accomplishment: print knowledge and phonological awareness (e.g., Justice & Ezell, 2001a; Lonigan, 2004; van Kleeck, 1998; Whitehurst & Lonigan, 1998). *Print knowledge* is an umbrella term that describes children's maturing knowledge about the rule-governed system of orthography and written language. This includes both print-concept knowledge (PCK), defined as knowledge of the rule-governed organizational properties of print (e.g., left-to-right directionality, combinatorial properties of letters to make words), and alphabet knowledge, defined as knowledge of the distinctive features and names of individual

ABSTRACT: **Purpose:** This research determined the psychometric quality of a criterion-referenced measure that was thought to measure preschoolers' print-concept knowledge (PCK).
Method: This measure, titled the Preschool Word and Print Awareness (PWPA), was examined using the partial credit model (PCM) to determine its suitability for use by clinicians, educators, and researchers. The extent to which the PWPA differentiated estimates of PCK for at-risk populations on the basis of socio-economic status (SES) and language ability was also studied. The sample population was one-hundred twenty-eight 3- to 5-year-old children who varied in SES (middle, low) and language ability (typical language, language impairment) as derived from several previous or ongoing studies of emergent literacy intervention.
Results: The PCM fit analyses showed good fit between the overall data and the PCM, indicating that the PWPA provided a valid estimate of the latent PCK trait. SES and language ability were significant predictors of PWPA scores when age was used as a covariate. These results showed the PWPA to be suitable for measuring preschoolers' PCK and to be sensitive to differences among children as a function of risk status.
Clinical Implications: The results show the PWPA to be an appropriate instrument for clinical and educational use with preschool children.

KEY WORDS: emergent literacy, assessment, print awareness, preschool children, specific language impairment

alphabet letters (Justice & Ezell, 2004; Lonigan et al., 1999). *Phonological awareness* describes children's developing sensitivity to the sublexical, segmental structure of the phonological domain of language, including sensitivities to larger units (e.g., words, syllables) and smaller units, particularly phonemes (Anthony, Lonigan, Driscoll, Phillips, & Burgess, 2003; Lonigan, Burgess, & Anthony, 2000; Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999). Well-developed skills in both print knowledge and phonological awareness enable children to more readily profit from beginning reading instruction in kindergarten and beyond (Chaney, 1998; Lonigan, 2004; McBride-Chang & Kail, 2002; Stahl & Murray, 1994; Storch & Whitehurst, 2002).

At present, the developmental literature on children's attainment of print knowledge is relatively underdeveloped relative to the literature on phonological awareness. (For several excellent summaries of the phonological awareness literature, see Burgess & Lonigan, 1998; Gillon, 2004; Lonigan, Burgess, Anthony, & Barker, 1998; Schatschneider et al., 1999; Stahl & Murray, 1994; Stanovich, 2000). This is not to say, however, that children's print knowledge is any less important as a necessary precursor to skilled reading, with current emergent literacy literature consistently showing print knowledge to be one of the better predictors of school-age reading ability and inability. For instance, a meta-analysis conducted by the National Early Literacy Panel (NELP) studied 324 papers that were published in refereed journals that characterized the relationship between prereading skills of preschool and kindergarten children and their later reading abilities in word decoding and comprehension (Lonigan, 2004). Not surprisingly, this panel found the best predictor of decoding skills in first grade and beyond to be the child's ability to decode words before first grade. However, when looking at those skills that are *precursors* to reading, and when controlling for multivariate relationships among variables, the best predictors of children's early reading success included measures of both print knowledge and phonological awareness (Lonigan, 2004).

The findings of the NELP, which underscore the importance of print knowledge to successful school-age reading performance, converge with those of other literacy researchers. Storch and Whitehurst's (2002) structural equation modeling of precursors to reading showed preschool alphabetic skills to predict 38% of the variance in kindergarten reading ability, with children's knowledge of "print principles" comprising the largest factor in the latent variable representing preschool alphabetic skills. Hammill's 2004 meta-analysis of meta-analyses that documented the relationship between prereading skills and later reading abilities identified "graphological" or print skills to be the most robust predictors of later reading competence (also see Scarborough, 1998). These data confirm intuitive knowledge about reading that appreciates that children cannot learn to read without having some knowledge of print, and further, that those children who arrive to reading instruction with well-developed knowledge about print will make relatively better progress than those youngsters with underdeveloped knowledge about print.

Given the importance of early print knowledge to later reading, researchers and practitioners are increasingly concerned with identifying children who show difficulty with the development of print knowledge before kindergarten. Two risk factors that are well documented include language impairment (Bishop & Adams, 1990; Boudreau & Hedberg, 1999; Justice, Chow, Capellini, Flanigan, & Colton, 2003) and low socioeconomic status (SES) (Bowey, 1995; Chaney, 1994; Justice & Ezell, 2001a; Lonigan et al., 1999; Raz & Bryant, 1990; Rush, 1999). Children with underdeveloped language skills, particularly those with clinically depressed language skills or language impairment (LI), show delayed understanding of the symbolic significance of print (Gillam & Johnston, 1985) and are more likely than typically developing children to show low levels of interest toward print (Kaderavek & Sulzby, 1998). Approximately half of kindergarteners with LI will exhibit reading disability in second grade (Catts et al., 2002). Whereas LI may be characterized as a causal variable for emergent (and conventional) literacy difficulties (Scarborough, 1990), being reared in a home that is socioeconomically poor serves as a marker variable to identify an elevated risk status for emergent literacy problems due to a variety of environmental circumstances (e.g., less participation in literacy events, less modeling of adult literacy use; see Burgess, 1999). Lonigan and colleagues (1999) compared middle- and low-SES preschoolers on a variety of emergent literacy measures to include three measures of print knowledge (PCK, alphabet knowledge, environmental print) that were collapsed with a measure of letter-sound knowledge to form a composite "print knowledge" variable. The difference in print knowledge between the two groups was nearly 1 *SD* ($d = 0.83$), a difference similar to that of phonological awareness ($d = 0.75$). These findings converge with a larger research corpus that consistently shows low-SES children to enter school at a significant disadvantage in all areas of emergent literacy (e.g., Bowey, 1995; Chaney, 1994; Raz & Bryant, 1990; Rush, 1999).

As clinicians and educators seek to integrate evidence-based practices into early interventions that accelerate emergent literacy attainment for at-risk preschoolers, approaches to measurement become an essential consideration. Indeed, numerous current public policy initiatives directed at preschool education, such as the U. S. Department of Education's Early Reading First Program, require grantees to implement screening tools that reliably and validly identify children who show delays in emergent literacy development and that can characterize the overall effectiveness of emergent literacy interventions. Whereas the measurement literature for phonological awareness is reasonably well developed (e.g., see Schatschneider et al., 1999), this is not the case for print knowledge, which includes both PCK and alphabet knowledge. A recent review of psychometrically sound and useable instruments for evaluating PCK for elementary-grade children identified only eight instruments available for this purpose; by contrast, 24 instruments were available to assess phonological awareness (see Rathvon, 2004).

As a construct, children's PCK involves a range of skills, to include the functions of print, concept of letter, concept of word, directionality of print, and organization of books as but a few examples (Meisels & Piker, 2001). Researchers have used numerous tools for estimating children's knowledge in this area, including subtests from such standardized norm-referenced instruments as the Test of Early Reading Ability—3 (TERA–3; Reid, Hresko, & Hammill, 2002) and the Developing Skills Checklist (DSC; CTB/McGraw-Hill, 1990). As with most norm-referenced tests, these tests sample a small subset of indicators of PCK and generate standard scores by which to compare a child's scores against those of same-aged peers. Often, researchers desire greater depth in assessment than is provided by such tests to generate theoretical descriptions of print-concept development and to document changes in PCK during intervention (e.g., Chaney, 1994; Justice & Ezell, 2000; McGee, Lomax, & Head, 1988; Neuman & Roskos, 1993).

Thus, a common tradition in research on PCK has been the use of investigator-designed tools (e.g. Bloodgood, 1999; Chaney, 1994; Justice & Ezell, 2000; Lonigan et al., 1999; McGee et al., 1988). For instance, children are asked to "read" presentations of environmental print (e.g., Justice & Ezell, 2000; Lomax & McGee, 1987; Lonigan et al., 1999) or are asked to identify specific print units within the pages of a storybook (e.g., Chaney, 1994; Justice & Ezell, 2001a). On the one hand, such instruments typically exhibit high levels of face validity, produce scores that correlate moderately to highly with later reading performance (predictive validity; see Hammill, 2004), and can be implemented reliably (e.g., Justice & Ezell, 2001a). On the other hand, such instruments have rarely been subjected to rigorous psychometric scrutiny, instrument descriptions (provided in peer-reviewed articles) generally provide too little detail for replication of use, and information on how to interpret scores typically is not included (Meisels & Piker, 2001). To improve practitioners' access to research-based instruments, the developers of such instruments have been asked to undertake "systematic analysis… of the psychometric properties that provide justification for their meaning and use" and to provide guidance for interpreting children's scores (Meisels & Piker, 2001, p. 33).

Of researcher-designed tasks for studying PCK in preschool children, one of the most commonly used has been the adaptation of Clay's Concepts About Print task (CAP; 1979; for examples of use, see Lomax & McGee, 1987; Lonigan et al., 1999; Neuman, 1999; Reutzel, Oda, & Moore, 1989; Tunmer, Herriman, & Nesdale, 1988; Ukrainetz, Cooney, Dyer, Kysar, & Harris, 2000). The CAP task, which was developed for 5- and 6-year-olds, is one of several tasks included in "An Observation Survey of Early Literacy Achievement" (Clay, 1979), a diagnostic criterion-referenced protocol that is used in the Reading Recovery program to study the reading difficulties of children who show delayed reading progress (see Clay, 1993). Children's knowledge of specific print concepts (e.g., the title of a book, the directionality of print) is assessed through 24 items delivered within the context of a shared reading activity between examiner and child. Psychometric analyses for administration with children ages 5 to 7 years have yielded acceptable levels of internal consistency and criterion-related validity, and norms are available for children in this age range (see summary in Meisels & Piker, 2001). Given that Clay's instrument was designed for use with children who *are* reading, many of its tasks are too difficult for prereaders (e.g., reading words). However, a number of researchers have adapted the tool for younger children by eliminating more difficult items (e.g., Boudreau & Hedberg, 1999; Lomax & McGee, 1987; Lonigan et al., 1999; Purcell-Gates, 1996). Justice and Ezell (2001b) published their adaptation of the CAP, which involved expansion of the sample of items that focused on prereading competencies and elimination of more difficult items (e.g., naming punctuation units, reading individual words). Specifically, the number of items was reduced from 24 to 14, the book stimuli was changed to a commercially available children's storybook (rather than the book designed by Clay for the CAP), items were added to look at a broadened range of specific text concepts (e.g., top line of a page), and items that focused on word recognition and punctuation units were deleted. The authors have used this instrument—the Preschool Word and Print Awareness (PWPA)—to document the impact of emergent literacy interventions (e.g., Ezell, Justice, & Parsons, 2000; Justice & Ezell, 2000, 2002) and to describe generally those print concepts that young children exhibit (e.g., Justice & Ezell, 2001a, 2001b). Nonetheless, the extent to which the PWPA exhibits acceptable psychometric standards has not been empirically examined and is the focus of the present research. For practitioners and researchers to use the PWPA, it is necessary to provide justification for its use through empirical examination.

The goals of this study were twofold: (a) to characterize the PWPA's measurement quality for 3- to 5-year-old children using an item response theory (IRT) model, and (b) to examine the extent to which PWPA performance differentiated PCK for three groups of at-risk children relative to typically developing advantaged children. The three at-risk groups included children with LI from middle-class backgrounds, children with typical language skills from low-SES backgrounds, and children with LI from low-SES backgrounds. By studying the utility of the PWPA for children ranging in age from 3 to 5 years, including those exhibiting salient risk factors, the results of this research fill an important measurement need. This age range represents the period during which most children in the United States attend preschool programs, with an estimated 75% of 4-year-olds attending some type of center-based care (National Association for the Education of Young Children, 2000). Given the importance of this period for preparing children for their forthcoming transition to kindergarten, the majority of states currently have educational standards focused on literacy achievement for children in this age range (Neuman, Roskos, Vukelich, & Clements, 2003). PCK is consistently identified as a key domain in such standards and is a particularly important concern for professionals serving children who exhibit vulnerabilities in this developmental area. Yet, there are few instruments available to estimate children's growth in PCK as they participate in preschool interventions. Thus, results are likely to be directly useful to clinicians, educators, and researchers for identifying children whose PCK is lagging behind their peers and for modeling individual children's growth within educational and clinical interventions.

METHOD

Participants

The 128 children in this study were participants in several published and ongoing studies that have used the PWPA as a pre- and posttest measure of PCK to study the impact of preschool literacy interventions (Justice & Ezell, 2000, 2002; Justice et al., 2003). The total sample consisted of one-hundred twenty-eight 3- to 5-year old children from urban, suburban, and rural regions of southeast Ohio (n = 60) and Virginia (n = 68) to include 65 boys and 63 girls, with a mean age of 53 months (SD = 4.8, range 41–62 months). Eighty-two percent of the children were Caucasian (n = 105), 12% were African American (n = 15), and 2% were Asian American (n = 3) (an additional 4% of the sample, or n = 5, was characterized by caregivers as "other"). To participate in any of the studies from which data were drawn, children were required to (a) be native speakers of English and reside in a home in which English was the primary language spoken, both indicated via parental questionnaire; (b) pass a bilateral hearing screening at 25dB or 30dB (depending on ambient noise at time of testing) at 500, 1000, 2000, and 4000 Hz; and (c) have no history of gross motor, hearing, cognitive, or neurological impairment, as indicated by parental questionnaire or formal screening.

Socioeconomic status. Thirty-eight percent of the children (n = 49) resided in low-SES homes as documented by their eligibility to attend preschool programs that used poverty guidelines as eligibility criteria (e.g., Head Start). Thirty-four of the low-SES children were Caucasian, 9 were African American, 2 were Asian American, and 4 were considered "other." All of these children resided in homes in which the household income fell below the federal poverty limits. Household income data, available for 30 of these 48 children, showed a mean household income of $9,477 ($SD$ = 5,239). Approximately two thirds of the low-SES children were from the rural Appalachian region of southeastern Ohio (n = 30), whereas 19 were from urban and rural regions of Virginia.

The remaining 63% of children (n = 79) resided in middle-SES homes in Ohio (n = 30) or Virginia (n = 49). Seventy-one of the middle-SES children were Caucasian, 6 were African American, 1 was Asian American, and 1 was considered "other." None of these children attended need-based preschool programs and they were recruited from among the general community (e.g., private-pay day cares, library story hours) for research participation.

Language status. Whereas the majority of our sample (n = 94) exhibited typically developing language skills (TL) as indicated by norm-referenced testing, approximately one fourth of our sample (27%, n = 34) exhibited LI. We use this term to generally describe children showing a developmental impairment of language that occurs in the absence of gross cognitive, hearing, or neurological deficits. A primary impairment of language affects approximately 7% to 10% of 5-year-old children (Beitchman et al., 1998; Tomblin et al., 1997) and is one of the better documented risk factors for early and later reading difficulties (e.g., Catts et al., 2002; Scarborough, 1990). The data for the 34 children with LI in the present study were drawn from two studies that differed in the norm-referenced language assessment used for identifying presence of LI. In both studies, the diagnostic criteria for identifying LI was substandard performance (i.e., at least 1 SD below the mean) on one or more composites of a valid and reliable norm-referenced measure of oral language (e.g., Preschool Language Scales—3, Zimmerman, Steiner, & Pond, 1992; or Test of Language Development—Primary, Newcomer & Hammill, 1997) and performance better than –1 SD of the mean on a norm-referenced screening of nonverbal abilities (e.g., Matrices subtest of the Kaufman Brief Intelligence Test, Kaufman & Kaufman, 1990).

Of the 34 children with LI, 23 were tested directly by one of the authors or graduate students under their supervision to identify the presence of LI; for the remaining 11 children, their diagnoses were made by speech-language pathologists (SLPs) at their preschool. Of the children with LI, 17 were Caucasian, 13 were African American, 1 was Asian American, and 3 were characterized as "other." Forty-one percent of the 34 children with LI (n = 14) also came from low-SES families, thus exhibiting both risk factors simultaneously.

A comparison of the four groups of preschoolers (middle-SES TL, low-SES TL, middle-SES LI, low-SES LI) showed that they did not differ in chronological age, $F(3, 127) = 1.124$, $p = .342$, but that significant group differences existed for both receptive and expressive language ability, $F(3, 121) = 53.1$, $p < .001$ and $F(3, 121) = 27.67$, $p < .001$, respectively, as would be expected on the basis of how the groups were formed. All four groups were significantly different for receptive and expressive language scores except for the low-SES LI and middle-SES LI children who were similar for expressive language, and the low-SES TL and middle-SES LI children who were similar for receptive language. Mean scores and standard deviations for the four groups are presented in Table 1.

General Procedures

Children were individually tested with the PWPA in their homes, in a laboratory setting, or at their preschool or day care. The measure was administered to children along with other measures of literacy and language pursuant to the goals of the particular study in which the children were enrolled. Regardless of the individual battery used, the assessments were uniform in the following ways. First, the PWPA (and other measures) was administered by a trained examiner, often a doctoral-level research assistant with considerable experience working with young children. Second, during administration of the test, the examiner provided no feedback or praise with the exception of reminders for on-task performance when needed. Third, no single test battery lasted longer than 45 min, and breaks were provided in response to child cues. Often, the test battery was much shorter, lasting perhaps 15 or 20 min. When children had completed testing, they were given a token gift (e.g., stickers, storybook).

Measurement of PCK

The PWPA is an individually administered measure of children's knowledge of 14 concepts about print, as listed in Table 2. These

Table 1. Means (standard deviations) for participant characteristics.

		Typical language		*Language impairment*	
	Total sample (n = *128)*	*Low-SES* (n = *35)*	*Middle-SES* (n = *59)*	*Low-SES* (n = *14)*	*Middle-SES* (n = *20)*
Age	53.5 (4.8)	52.7 (6.0)	53.3 (4.6)	53.6 (3.2)	55.2 (3.4)
Expressive language	100.1 (18.9) (n = 122)	98.9 (12.9)	114.0 (14.0) (n = 54)	76.7 (7.1)	80.7 (9.2) (n = 19)
Receptive language	96.6 (13.3) (n = 98)	96.5 (6.9)	107.0 (10.7) (n = 30)	78.0 (13.1)	94.3 (10.5) (n = 19)

Note. Sample size is indicated when data are not available for all participants. All scores are standard scores with M = 100 and SD = 15. Expressive language = standard score from the One Word Expressive Picture Vocabulary Test—R (Gardner, 1990), Preschool Language Scale—3 Expressive Language subtest (PLS–3; Zimmerman, Steiner, & Pond, 1992), or Test of Language Development-Primary Speaking Quotient (TOLD–P:3; Newcomer & Hammill, 1997). Receptive language = standard score from the Peabody Picture Vocabulary Test—III (Dunn & Dunn, 1997), PLS–3 Receptive Language subtest, or TOLD–P:3 Listening Quotient.

Table 2. Concepts examined in the Preschool Word and Print Awareness (PWPA) measure.

Location of the front of the book
Location of the title of the book
Purpose of the title
Print function: Narrative text
Organization of narrative print: Left to right on the page
Print versus pictures
Organization of book: Left to right page
Organization of print: Top line
Organization of print: Bottom line
Purpose of contextualized print
Letter as a print unit
First letter
Capital letter
Meaning of contextualized print

tasks are administered during an adult–child shared storybook reading using the commercially available picture book, *Nine Ducks Nine* (Hayes, 1990). Similar to Clay's (1979) protocol for the CAP task, the examiner begins the test session by explaining to the child: *We are going to read a book together, and I need you to help me read.* The examiner then embeds a series of tasks into the book-reading interaction, such as *Show me the front of the book, Show me the name of the book,* and *Show me just one letter on this page.* An incorrect or lack of response receives a score of 0. Correct responses typically are assigned 1 point, although for several items, correct responses receive 2 or 3 points and alternate responses receive 1 or 2 points. The order of tasks was administered in the same order to all children, following the protocol provided in the Appendix.

Procedural fidelity and interrater reliability have been reported for a small number of test administrations by Justice and Ezell (2001b) and Justice and Ezell (2000), respectively. For the former, independent scoring via audiotape of test administrators' consistency in providing directions, sequencing tasks, and providing feedback during task administration was studied for a randomly selected 25% of 30 test administrations. The test administrators' fidelity to the administration protocol ranged from 96% to 100%, with an average fidelity score of 99%. For the latter, double-scoring of children's performance on the PWPA was conducted for a randomly selected 25% of 38 test administrations by two trained observers working independently. Point-by-point agreement in scoring was 94%. These data suggest that the measure can be reliably implemented and scored across examiners.

Validation and IRT

The present research examined the validity of the PWPA with an IRT model. IRT is a set of models and associated statistical techniques for analyzing tests, questionnaires, and other instruments containing multiple items with ordered categorical data. The origins of IRT can be traced to the 1960s (Lord & Novick, 1968; Rasch, 1980), and IRT is now considered the dominant form of measurement when an attribute cannot be measured directly (i.e., latent traits; Embretson & Reise, 2000). Unlike directly measurable characteristics, like height and weight, the majority of attributes of interest to SLPs cannot be measured directly (e.g., vocabulary knowledge, cognitive ability), as is the case for children's knowledge of print concepts. When measuring such attributes, researchers and clinicians use derived measurement to estimate an unobservable (latent) trait. The following provides a brief introduction to IRT, and readers are referred to Bond and Fox (2001), Embretson and Reise, and Hambleton, Swaminathan, and Rogers (1991) for a more complete description.

The most common IRT models have several features in common. First, there is an object of measurement (e.g., child) with an unknown level of a single trait of interest (e.g., PCK). Second, there is an instrument of measurement, which in the present study consists of the items of the PWPA. With IRT, the instrument of measurement is described by one or more unknown parameters that reflect aspects of the instrument that affect the observed response, such as item difficulty. Third, there is a model, or functional form, that relates the unknown trait level and instrument parameters to the probabilities of observed responses (Embretson & Reise, 2000). In practice, the observed responses are used in the framework of the model to estimate the unknown parameters; this yields estimates of the level of the trait for each person, as well as information on the functioning of the measurement instrument. An important feature of IRT is that the measurement properties that are obtained are independent of the sample studied. As noted by Schatschneider et al. (1999), this differentiates IRT from other classical validation approaches (e.g., item-to-test correlations) as the model parameters are sample independent.

In the current research, we employed the partial credit model (PCM; Masters, 1982) using the FACETS computer program (Linacre, 2000). The PCM models item responses with one parameter describing the child's trait level, which we term PCK, and a set of item-specific category thresholds. The category thresholds indicate how difficult it is to move up one category in the rating scale. Higher numbers indicate that a category is more difficult to reach. There are one fewer category thresholds than there are categories in the rating scale. Therefore, if the item is dichotomous (i.e., has two rating scale categories), the category threshold is an indicator of item difficulty.

The PCM is an item response model in the Rasch (1960/1980) family of models (Rost, 2001) and has several desirable statistical properties that generally are not achieved with typical raw score summation (Embretson, 1996; Schatschneider et al., 1999). First, Rasch family models are much easier to estimate robustly than other IRT models, and therefore can be applied to relatively small data sets. We considered our sample size of 128 to be sufficient for IRT using the PCM based on previous research on parameter estimate stability (Linacre, 2002; Lord, 1983). Second, if data are found to fit to the PCM, this implies that PCK can be measured on an interval scale (Brogden, 1977; Perline, Wright, & Wainer, 1979), thereby indicating that equal differences between estimates of PCK imply equal differences in the underlying trait levels. For example, the difference between an estimate of 70 and 80 is similar in magnitude to the difference between an estimate of 120 and 130. Other measurement models, especially typical raw score summation, do not have this property, and interval level scaling is assumed rather than tested for. Importantly, lack of fit to the PCM implies that interval level scaling likely is not achieved. Therefore, the assessment of fit is vital.

The present research estimated fit to the PCM using two standard Rasch-based fit statistics, Infit and Outfit, applied to both PWPA items and respondents (i.e., children). Infit and Outfit assess the match between the expected item responses (based on the model predictions) and the actual item responses, with Outfit being relatively more sensitive to large deviations from expectations (e.g., a score of 3 when the expected score is 0), and Infit being relatively more sensitive to patterns of minor deviations (e.g., 5 items with scores of 1 when the expected score is 0). Both statistics are reported in two forms: the mean-square form and the standardized form. The mean-square form assesses the amount or magnitude of misfit, with a value of 1 indicating perfect fit, values greater than 1 indicating misfit due to more randomness (i.e., noise or error) than predicted, and values less than 1 indicating misfit due to less information than expected (i.e., redundancy with the information from other items). Wright and Linacre (1994) suggested that Infit or Outfit mean-square values of less than .6 or greater than 1.4 indicate a practically significant amount of misfit, whereas Smith and Suh (2002) suggested bounds of .9 and 1.1. We employ bounds intermediate to these, with values within .7 to 1.3 indicating acceptable fit. The standardized forms of Infit and Outfit convert the mean square to an approximate standard normal variable so that values less than –2 or more than 2 indicate a statistically significant misfit (with alpha = .05). We consider misfit to be important and of concern if it is both statistically significant, as indicated by the standardized fit statistics, and practically substantial, as indicated by the mean-square fit statistics. Consistent with typical practice, after identifying misfit, we examine potential sources of misfit. If identification of the sources of misfit was possible, an attempt is made to manipulate the data in order to avoid eliminating data while still achieving acceptable fit. If this is not possible, then the misfitting item or respondent is removed from the data.

In addition to trait level estimates, the PCM yields statistical information on the functioning of the rating scales within each item. We examined category usage to identify underused categories (i.e., scores that did not occur at expected rates). In particular, we looked at two statistics: the category thresholds and the percentage of responses in each category. Category thresholds that are similar or that occur in reverse order (i.e., the threshold between two lower scored categories is higher than the threshold between two higher scored categories) indicate that a category or categories are underused and could perhaps be eliminated with little loss of information. We also examined the average PCK within each category (see Linacre, 2004, for the calculation). If two categories are associated with the same level of PCK, then those categories are essentially interchangeable and may possibly be collapsed into a single category without loss of information.

RESULTS

Validation of PWPA

The first research goal was to characterize the PWPA's measurement quality for 3- to 5-year-old children using an IRT model. Fit analyses indicated that the data fit the PCM well at both the respondent and the item level. At the respondent level, although several children had an Infit and Outfit mean square outside the 0.7 to 1.3 bounds, none of this misfit was statistically significant.

At the item level, only one item demonstrated significant misfit. The original version of Item 3 (*What do you think it* [the title] *says?*) had an Infit and Outfit mean square of 1.5 and 2.0 and standardized Infit and Outfit of 4 and 3, respectively, showing there to be more randomness than predicted. Examination of the functioning of the rating scale for Item 3 indicated that the average PCK for a child receiving a score of 0 (no knowledge of the role of the title) was actually higher than the average knowledge of a child receiving a score of 1 (explains the role of the title), although both were lower than a score of 2 (says 1 or more words in the title or relevant title). Thus, the analyses showed no practical difference between responses receiving a 0 and 1, although there were differences between these two responses and those receiving a score of 2. The categories 0 and 1 were collapsed to form a single response, and the PCM analyses were rerun, which eliminated the misfit associated with Item 3.[1] The final version of the PWPA (see Appendix) is identical to the previously available version (Justice & Ezell, 2001b) except that Item 3 was dichotomized into two response scores (0, 1) from three scores (0, 1, 2), and thus the total points possible on the instrument was reduced from 18 to 17. The reliability of the revised version as estimated in FACETS is .74. This analysis also converted PWPA total scores to estimates of PCK scaled for a mean of 100 and a standard deviation of 15, as shown in Table 3.

Examination of the rating scale characteristics indicated that some responses were rarely used. Much of this was associated with item difficulty. For instance, Item 1 (*Show me the front of the book*) was very easy, with 90% of the children responding correctly, so that a score of 0 was rare for this item. Likewise, Item 10 (*Point to the words spoken by the ducks in the illustration*) was very difficult to get correct (13%), so a score of 1 was rare. In addition, the highest scores were difficult to obtain for Items 4 (13%) and 11 (5%). There was one reversal of category thresholds, associated with low probability of use for the middle category (score = 1) for Item 5 (*Then which way do I read?*). It was unusual for a child to indicate that one reads from top to bottom; if the child understands the top-to-bottom direction, he or she is likely also to understand the left-to-right direction. However, the average PCK score increased as expected across score categories, suggesting that a score of 1, although rarely used, does indicate more PCK than a score of 0 and less knowledge than a score of 2.

Table 3. Conversion of PWPA raw scores to print-concept knowledge estimates (M = 100, SD = 15).

PWPA raw score	*Print-concept knowledge estimate*
0	46
1	63
2	74
3	82
4	88
5	92
6	97
7	100
8	104
9	107
10	111
11	115
12	118
13	123
14	128
15	134
16	145
17	161

[1]The misfit associated with children on this item, although not significant before the reanalysis, was also reduced substantially.

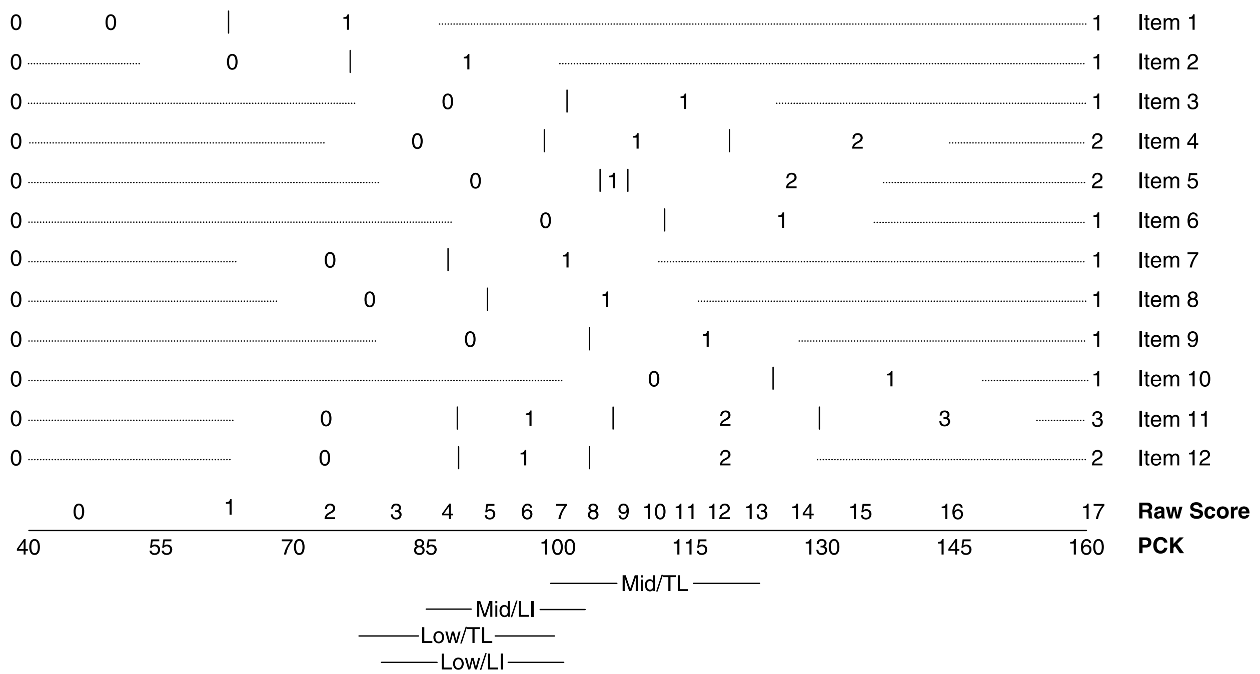
Graphic Representation of Performance

One advantage of using the PCM is that performance on the PWPA can be depicted graphically using a trait map, as presented in Figure 1 (Wright, Mead, & Ludlow, 1980). On Figure 1, the *x*-axis represents PCK estimates (below the axis) on the basis of a mean of 100 and a standard deviation of 15 as they correspond to total raw scores (above the axis). On the *y*-axis, individual PWPA items are listed as well as the possible response categories (e.g., 0, 1 for Item 1). The bar between the response options represents the point at which adjacent response categories are equally probable, with the exception of Item 5, which, because of the category threshold reversal, has misordered equal probability bars. By following the directions presented at the bottom of Figure 1, this graphic representation of performance may be of use to clinicians and educators to document children's overall performance and to study children's responses to individual items. By drawing a vertical line upward from a child's raw score on the trait map, the clinician can compare expected responses (those closest to the vertical line) against observed responses, as illustrated in Figure 2 for a child who received a raw score of 13 and a PCK estimate of 123 (+1.5 *SD* of the mean). By plotting the child's responses to individual items, one notes that three responses (3, 11, 12) were not clustered around or close to the vertical line indicating expected responses. Of these, responses to items 3 and 12 were unexpectedly lower than expected, whereas the response to item 11 was higher than expected. From an applied perspective, modeling individual children's performance allows clinicians and educators to track children's performance over time and to determine the extent to which any idiosyncrasies in PCK may be clinically relevant.

Group Differences in Print Concepts

A second goal of this research was to examine the extent to which PWPA performance differentiated PCK for three groups of at-risk children, namely children with LI, children from low-SES backgrounds, and children with LI from low-SES backgrounds, relative to typically developing advantaged children. We thus examined group differences in PCK as measured by the PWPA for four subgroups of children: middle-SES TL children, low-SES TL children, middle-SES LI children, and low-SES LI children. PCK estimates for the total sample and subgroups are presented in Table 4. Means and standard deviations for the subgroups also are plotted graphically below the *x*-axis on the Figure 1 trait map.

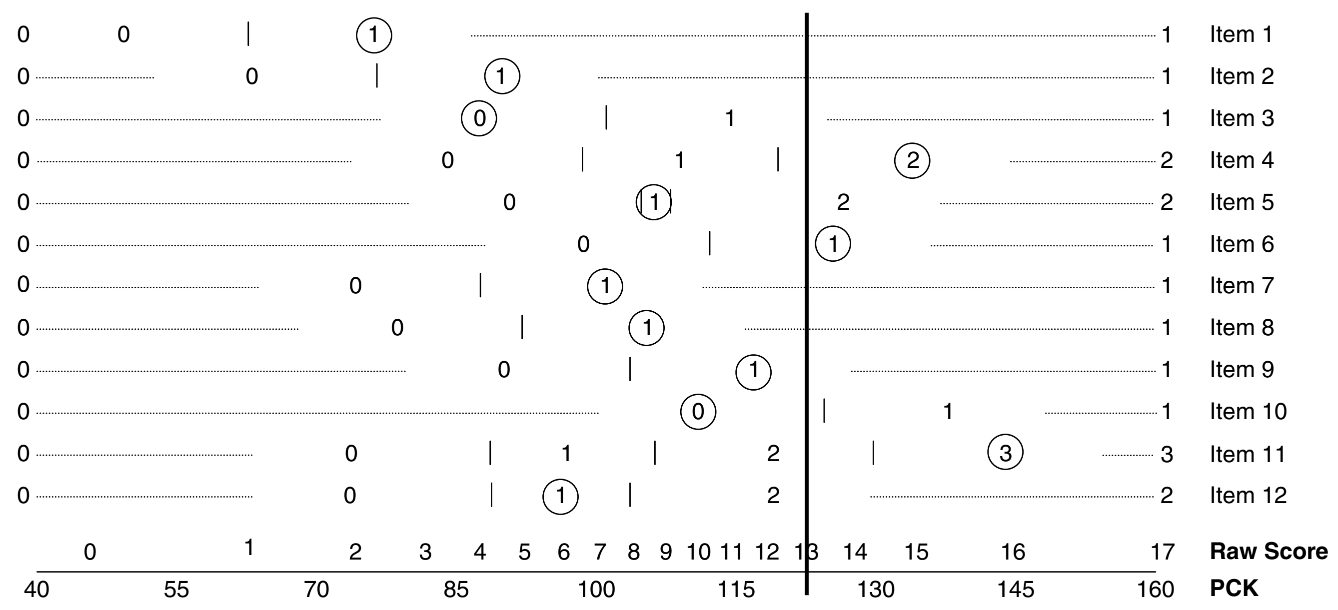
Figure 1. Trait map depicting print-concept knowledge estimates.



Raw Score	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
PCK	46	63	74	82	88	92	97	100	104	107	111	115	118	123	128	134	145	161

Directions: Circle the child's responses to each item; sum and circle the corresponding raw score. Draw a vertical line upwards from the raw score. The responses closest to the vertical line are expected item responses based on the child's raw score. Responses far from the vertical indicate unexpected scores.

Figure 2. Example of use of a print-concept knowledge trait map.



With age serving as a covariate, an analysis of variance (ANOVA) with two predictors, low- versus middle-SES and TL versus LI, was used to predict the PCK estimate from the PWPA. SES, $F(1, 123) = 32, p < .01$; LI, $F(1, 123) = 13, p < .01$; and the interaction between SES and LI, $F(1, 123) = 18, p < .01$, were all significant predictors of PCK ($R^2 = .49$). Low-SES children tended to have PCK estimates that were 22.3 points ($d = 1.5$) lower than those of middle-SES children, whereas children with LI had PCK estimates that were 17.6 points ($d = 1.2$) lower than those of TL children. However, these main effects were tempered by a positive interaction, such that low-SES children with LI had PCK estimates that were 19.1 points ($d = 1.3$) higher than those predicted by the main effects, suggesting that having both risk factors did not present a cumulative risk beyond having either risk factor in isolation. Table 5 provides predicted responses for individual PWPA items to characterize the probability by which children in the different subgroups might be expected to respond to various print-concept items.

Table 4. Print-concept knowledge estimates for all children and by subgroup.

Group	M	SD	*Range*
Total sample (*n* = 128)	100	15	63–134
Middle-SES TL (*n* = 59)	111	12	74–134
Middle-SES LI (*n* = 20)	94	9	74–111
Low-SES TL (*n* = 35)	89	11	63–118
Low-SES LI (*n* = 14)	90	10	63–107

Note. TL = typical language skills, LI = language impairment.

DISCUSSION

An interesting paradigm shift in the reading disability literature of the last decade has been the strong push toward preventive models by researchers, practitioners, and policymakers. Preventive models have become possible because researchers have identified early precursors of both reading success and reading failure (e.g., Catts et al., 2002; Chaney, 1998; Lonigan, 2004; Scarborough, 1998; Storch & Whitehurst, 2002). Preventive models focus on decreasing children's vulnerability for reading difficulties before they show any clinical manifestation of the disorder. The models also provide one avenue for decreasing the staggering rate of reading underachievement among America's schoolchildren, particularly the economically disadvantaged and those with disabilities (e.g., National Assessment of Education Progress, 2003). The cumulative research base intersecting reading processes, child development, and early intervention is making it possible for practitioners to identify children who are vulnerable for reading difficulties well before formal reading instruction and to deliver evidence-based interventions that accelerate early knowledge and mitigate early lags in development.

Of necessity to preventive models is access to validated measures that may be used to identify children whose development of important emergent literacy competencies is lagging behind peers and that model the effectiveness of various intervention approaches. Of the two major dimensions of emergent literacy development—namely, print knowledge and phonological awareness—measurement needs are greatest for the former, specifically the measurement of children's PCK. For this particular dimension of preschool literacy development, the cumulative literature shows that (a) it is an important early predictor of later reading achievement that differentiates children who are likely to be successful readers from those who are likely to be unsuccessful (e.g., Lomax & McGee, 1987; Lonigan, 2004; Storch

Table 5. Predicted responses by proportion of group to individual PWPA items.

Item	*Response category*	*Middle-SES TL*	*Middle-SES LI*	*Low-SES TL*	*Low-SES LI*
1	0	0.03	0.15	0.10	0.13
	1	0.97	0.85	0.90	0.87
2	0	0.08	0.35	0.26	0.31
	1	0.92	0.65	0.74	0.69
3	0	0.39	0.80	0.72	0.77
	1	0.61	0.20	0.28	0.23
4	0	0.28	0.76	0.66	0.72
	1	0.54	0.23	0.31	0.26
	2	0.19	0.01	0.03	0.02
5	0	0.38	0.89	0.83	0.87
	1	0.23	0.08	0.12	0.10
	2	0.40	0.02	0.05	0.03
6	0	0.61	0.91	0.87	0.89
	1	0.39	0.09	0.13	0.11
7	0	0.17	0.57	0.47	0.53
	1	0.83	0.43	0.53	0.47
8	0	0.23	0.66	0.56	0.62
	1	0.77	0.34	0.44	0.38
9	0	0.44	0.83	0.76	0.81
	1	0.56	0.17	0.24	0.19
10	0	0.81	0.96	0.95	0.96
	1	0.19	0.04	0.05	0.04
11	0	0.10	0.55	0.43	0.51
	1	0.42	0.38	0.45	0.41
	2	0.42	0.06	0.11	0.08
	3	0.06	0.00	0.00	0.00
12	0	0.09	0.55	0.42	0.50
	1	0.40	0.37	0.44	0.40
	2	0.51	0.08	0.14	0.10

& Whitehurst 2002; Tunmer et al., 1988), (b) it is an area of clear deficiency among children having little prior experience with print (e.g., children reared in poverty; Lonigan et al., 1999) and children with depressed language skills (e.g., Boudreau & Hedberg, 1999), and (c) it is a developmental competency that is readily amenable to change through intervention (e.g., Justice & Ezell, 2000, 2002). For these reasons, public policy initiatives that guide preschool education, particularly those programs serving large numbers of at-risk children, emphasize the ongoing screening of and intervention for early development of PCK, including the U. S. Department of Education's Early Reading First Program and the U. S. Department of Health and Human Services' Head Start Program (e.g., see Head Start Child Outcomes Framework; Head Start Bureau, 2000).

This study provides evidence that the PWPA is a sound instrument for examining PCK in 3- to 5-year-old children. The PWPA examines young children's attainment of 14 concepts about print in a framework that is similar to Clay's CAP (1979) instrument that was designed for elementary-grade children. Although researchers and practitioners have used Clay's CAP with younger children (e.g., Lomax & McGee, 1985; Lonigan et al., 1999; Neuman, 1999), its appropriateness for children who are of preschool age and not yet reading is questionable. The present study used IRT—namely, the PCM—to show that preschool children's PCK is a single trait that can be estimated reliably using the PWPA. The trait map and PCK estimate conversions from raw scores may be useful to clinicians and educators to identify individual children who show lags in performance or who demonstrate idiosyncratic knowledge of print concepts. Likewise, the PWPA may be used to document pre- and posttest changes during a course of emergent literacy intervention and to compare the relative impact of different intervention approaches. It is important, however, that clinicians recognize that the trait-level estimates provided here were derived from a sample of children varying in age from 3 to 5 years. As we did not disaggregate data by age, and children within the age range show considerable variability in PCK, we caution the use of these data as age-based norms.

The information obtained by studying PWPA performance for subgroups of children who varied as a function of SES and language ability showed that, as a whole, the instrument differentiated children exhibiting specific environmental or developmental risk factors from children without evident risk. Several interesting findings concerning PCK and risk emerged from these analyses. First, the results showed that children from low-SES households performed similarly to children with LI in PCK; compared to their economically advantaged and non-impaired peers, respectively, the impact of either risk factor resulted in a very large effect ($d = -1.5$ for low-SES and $d = -1.2$ for LI). Thus, both environmental and developmental risk exerted a similar level of disadvantage to children in their development of print concepts. Second, the results suggested there to be a lack of a cumulative effect when both risk factors—low SES and LI—were present. That is, the estimates of PCK appeared similar for low-SES TL children and middle-SES LI children, and for the small sample of 14 children exhibiting both low SES and LI, their PCK estimates were similar to middle-SES LI children and low-SES TL children. The lack of a cumulative effect when both socioeconomic disadvantage and LI are present for a child presents an interesting theoretical question for future research.

To sum, this work provides a variety of constituents—including clinicians, educators, and researchers—with an empirically sound method for assessing preschool children's attainment of PCK. The PWPA conceptualizes PCK as a single unitary trait that can be measured reliably as part of a more comprehensive assessment of emergent literacy knowledge. As a clinical tool, this measure may play an important role in the implementation of evidence-based interventions to accelerate PCK for at-risk preschoolers.

ACKNOWLEDGMENTS

This research was supported by Grant DC04933 from the National Institute of Deafness and Other Communication Disorders, Grant HD43204 from the National Institute of Child Health and Human Development, and an Elva Knight Research Award from the International Reading Association. We are grateful to the many children, families, teachers, SLPs, and program administrators who have assisted us with collecting the data presented in this work.

REFERENCES

Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M., & Burgess, S. R. (2003). Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly, 38,* 470–487.

Beitchman, J., Hood, J., Rochon, J., Peterson, M., Mantini, T., & Majumdar, S. (1989). Empirical classification of speech/language impairment in children I. Identification of speech/language categories. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 112–117.

Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, 31,* 1027–1050.

Bloodgood, J. (1999). What's in a name? Children's name writing and literacy acquisition. *Reading Research Quarterly, 34,* 342–367.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.

Boudreau, D. M., & Hedberg, N. L. (1999). A comparison of early literacy skills in children with specific language impairment and their typically developing peers. *American Journal of Speech-Language Pathology, 8,* 249–260.

Bowey, J. (1995). Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. *Journal of Educational Psychology, 87,* 476–487.

Brogden, H. E. (1977). The Rasch model, the law of comparative judgement and additive conjoint measurement. *Psychometrika, 42,* 631–634.

Bryant, P. E., Maclean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology, 26,* 429–438.

Burgess, S. (1999). The influence of speech perception, oral language, the home literacy environment, and prereading knowledge on the growth of phonological sensitivity: A 1-year longitudinal study. *Reading Research Quarterly, 34,* 400–402.

Burgess, S. R., & Lonigan, C. J. (1998). Bidirectional relations between phonological awareness and reading extended to preschool letter knowledge: Evidence from a longitudinal investigation. *Journal of Experimental Child Psychology, 70,* 117–141.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairment. *Journal of Speech, Language, and Hearing Research, 45,* 1142–1157.

Chaney, C. (1994). Language development, metalinguistic awareness, and emergent literacy skills of 3-year-old children in relation to social class. *Applied Psycholinguistics, 15,* 371–394.

Chaney, C. (1998). Preschool language and metalinguistic skills are links to reading success. *Applied Psycholinguistics, 19,* 433–446.

Clay, M. M. (1979). *The early detection of reading difficulties: A diagnostic survey with recovery procedures*. Portsmouth, NH: Heinemann.

Clay, M. M. (1993). *Reading recovery: Guidebook for teachers in training*. Portsmouth, NH: Heinemann.

CTB/McGraw-Hill. (1990). *Developing skills checklist*. Monterey, CA: McGraw-Hill.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341–349.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Ezell, H. K., Justice, L. M., & Parsons, D. (2000). Enhancing the emergent literacy skills of preschoolers with communication disorders: A pilot investigation. *Child Language Teaching and Therapy, 16,* 121–140.

Gardner, M. F. (1990). *Expressive One Word Picture Vocabulary Test-Revised*. Novato, CA: Academic Therapy Publications.

Gillam, R. B., & Johnston, J. R. (1985). Development of print awareness in language-disordered preschoolers. *Journal of Speech and Hearing Research, 28,* 521–526.

Gillon, G. (2004). *Phonological awareness*. New York: Guilford.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory.* Newbury Park, CA: Sage.

Hammill, D. D. (2004). What we know about correlates of reading. *Exceptional Children, 70,* 453–468.

Hayes, S. (1990). *Nine ducks nine*. Cambridge, MA: Candlewick Press.

Head Start Bureau. (2000). *The Head Start child outcomes framework*. Washington, DC: Author.

Justice, L. M., Chow, S. M., Capellini, C., Flanigan, K., & Colton, S. (2003). Emergent literacy intervention for vulnerable preschoolers: Relative effects of two approaches. *American Journal of Speech-Language Pathology, 12,* 320–332.

Justice, L. M., & Ezell, H. K. (2000). Stimulating children's print and word awareness through home-based parent intervention. *American Journal of Speech-Language Pathology, 9,* 257–269.

Justice, L. M., & Ezell, H. K. (2001a). Descriptive analysis of written language awareness in children from low income households. *Communication Disorders Quarterly, 22,* 123–134.

Justice, L. M., & Ezell, H. K. (2001b). Word and print awareness in 4-year-old children. *Child Language Teaching and Therapy, 17,* 207–225.

Justice, L. M., & Ezell, H. K. (2002). Use of storybook reading to increase print awareness in at-risk children. *American Journal of Speech-Language Pathology, 11,* 17–29.

Justice, L. M., & Ezell, H. K. (2004). Print referencing: An emergent literacy enhancement technique and its clinical applications. *Language, Speech, and Hearing Services in Schools, 35,* 185–193.

Kaderavek, J. N., & Sulzby, E. (1998, November). *Low versus high orientation towards literacy in children*. Paper presented at the annual convention of the American Speech-Language-Hearing Association, San Antonio, TX.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Linacre, J. M. (2000). *FACETS* (Version 3) [Computer software]. Chicago: MESA Press.

Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3,* 85–106.

Linacre, J. M. (2004). *A user's guide to Winsteps Ministep* [Computer software manual]. Retrieved October 6, 2004, from http://www.winsteps.com/winman/table3_2.htm

Lomax, R. G., & McGee, L. M. (1987). Young children's concepts about print and reading: Toward a model of word reading acquisition. *Reading Research Quarterly, 22,* 237–256.

Lonigan, C. J. (2004, October). *Development and promotion of emergent literacy: An evidence-based perspective*. Paper presented at the Early Literacy Development Conference, Niagara Falls, Canada.

Lonigan, C. J., Bloomfield, B. G., Anthony, J. L., Bacon, K. D., Phillips, B. M., & Samwel, C. S. (1999). Relations among emergent literacy skills, behavior problems, and social competence in preschool children from low- and middle-income backgrounds. *Topics in Early Childhood Special Education, 19,* 40–53.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology, 36*(5), 596–613.

Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of phonological sensitivity in two- to five-year-old children. *Journal of Educational Psychology, 90,* 294–311.

Lord, F. M. (1983). Small *N* justifies Rasch model. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (pp. 51–61). New York: Academic Press.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 60,* 523–547.

McBride-Chang, C., & Kail, R. V. (2002). Cross-cultural predictors in the predictors of reading acquisition. *Child Development, 73,* 1392–1407.

McGee, L. M., Lomax, R. G., & Head, M. H. (1988). Young children's written language knowledge: What environmental and functional print reading reveals. *Journal of Reading Behavior, 20,* 99–118.

Meisels, S. J., & Piker, R. A. (2001). *An analysis of early literacy assessments used for instruction*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

Morris, D., Bloodgood, J., Lomax, R. G., & Perney, J. (2003). Developmental steps in learning to read: A longitudinal study in kindergarten and first grade. *Reading Research Quarterly, 38,* 302–328.

National Assessment of Education Progress. (2003). *The nation's report card*. Retrieved October 10, 2004, from http://nces.ed.gov/nationsreportcard/

National Association for the Education of Young Children. (2000). *National Center for Education Statistics: Survey*. Retrieved October 1, 2005, from http://www.parenthoodweb.com

Neuman, S. B. (1999). Books make a difference: A study of access to literacy. *Reading Research Quarterly, 34,* 286–311.

Neuman, S. B., & Roskos, K. (1993). Access to print for children of poverty: Differential effects of adult mediation and literacy-enriched play settings on environmental and functional print tasks. *American Educational Research Journal, 30,* 95–122.

Neuman, S. B., Roskos, K., Vukelich, C., & Clements, D. (2003). *The state of state prekindergarten standards in 2003*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

Newcomer, P., & Hammill, D. (1997). *Test of Language Development: Primary* (*3rd ed.*). Austin,TX: Pro-Ed.

O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3,* 159–197.

Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement, 3,* 237–255.

Purcell-Gates, V. (1996). Stories, coupons, and the *TV Guide*: Relationships between home literacy experiences and emergent literacy knowledge. *Reading Research Quarterly, 31,* 406–428.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainments tests* (expanded ed.). Chicago: University of Chicago Press. (Original work published 1960).

Rathvon, N. (2004). *Early reading assessment: A practitioner's handbook*. New York: Guilford.

Raz, I. T., & Bryant, P. (1990). Social background, phonological awareness, and children's reading. *British Journal of Developmental Psychology, 8,* 209–225.

Reid, K. R., Hresko, W., & Hammill, D. (2002). *Test of Early Reading Ability—3*. Austin, TX: Pro-Ed.

Reutzel, D. R., Oda, L. K., & Moore, B. H. (1989). Developing print awareness: The effect of three instructional approaches on kindergartners' print awareness, reading readiness, and word reading. *Journal of Reading Behavior, 21*(3), 197–217.

Rost, J. (2001). The growing family of Rasch models. In A. Boomsma, M. A. J. van Duijn, & T. A. B. Snijders (Eds.), *Essays on item response theory* (pp. 25–42). New York: Springer-Verlag.

Rush, K. L. (1999). Caregiver-child interactions and early literacy development of preschool children from low-income environments. *Topics in Early Childhood Special Education, 19,* 3–14.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61,* 1728–1743.

Scarborough, H. S. (1998). Early identification of children at risk for reading difficulties: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–199). Timonium, MD: York Press.

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91,* 439–449.

Smith, R. M., & Suh, K. K. (2002, April). In *Rasch fit statistics as a direct test of the invariance of parameter estimation*. Paper presented at the 11th International Objective Measurement Workshop, New Orleans, LA.

Stahl, S., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86,* 221–234.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934–947.

Stuart, M. (1995). Prediction and qualitative assessment of five- and six-year old children's reading: A longitudinal study. *British Journal of Educational Psychology, 65,* 287–296.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40,* 1245–1260.

Tunmer, W. E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly, 23,* 134–158.

Ukrainetz, T. A., Cooney, M. H., Dyer, S. K., Kysar, A. J., & Harris, T. J. (2000). An investigation into teaching phonemic awareness through shared reading and writing. *Early Childhood Research Quarterly, 15,* 331–355.

van Kleeck, A. (1998). Preliteracy domains and stages: Laying the foundations for beginning reading. *Journal of Children's Communication Development, 20,* 33–51.

Walpole, S., Chow, S. M., & Justice, L. M. (2004). Literacy achievements during kindergarten: Examining key contributors in an at-risk sample. *Early Education and Development, 15,* 245–264.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848–872.

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

Wright, B. D., Mead, R. J., & Ludlow, L. H. (1980). *KIDMAP: Person-by-item interaction mapping* (MESA Research Memorandum #27). Chicago: MESA Press.

Zimmerman, I., Steiner, V., & Pond, R. (1992). *Preschool Language Scale—3*. San Antonio,TX: The Psychological Corporation.

Received January 3, 2005
Accepted August 22, 2005
DOI: 10.1044/0161-1461(2006/024)

Contact author: Laura M. Justice, Box 400873, Preschool Language and Literacy Lab, University of Virginia, Charlottesville, VA 22904-4873. E-mail: ljustice@virginia.edu

APPENDIX. PRESCHOOL WORD AND PRINT AWARENESS

Print Concepts

DIRECTIONS: Present the following tasks in the order depicted below. Use the book *Nine Ducks Nine* (Hayes, 1990). Read the text presented on the page and then administer the task. Each item may be repeated one time. Do not prompt, reinforce, or provide feedback to the child in any way. Score 0 for items to which the child does not respond or provides an answer that does not meet scoring criteria.

SAY: *We're going to read this book together, and I need you to help me read.*

Score/Item	*Page: Examiner Script*	*Scoring Criteria*
_____1. Front of book	***Before administering task:*** Give book to child with spine facing child. **Cover:** ***Show me the front of the book.***	1 pt: turns book to front or points to front
_____2. Title of book	**Cover:** ***Show me the name of the book.***	1 pt: points to one or more words in title
_____3. Role of title	**Cover:** ***What do you think it says?***	1 pt: says 1 or more words in title or relevant title
_____4. Print vs. pictures	**Page 1–2:** ***Where do I begin to read?*** ***After administering task:*** Put finger on first word in top line and say: ***I begin to read here***	2 pts: points to first word, top line 1 pt: points to any part of narrative text
_____5. Directionality	**Page 1–2:** ***Then which way do I read?***	2 pts: sweeps left to right 1 pt: sweeps top to bottom
_____6. Contextualized print	**Page 3–4:** ***Show me where one of the ducks is talking.***	1 pt: points to print in pictures
_____7. Directionality (left/right)	**Page 5–6:** ***Do I read this page (point to left page) or this page (point to right page) first?***	1 pt: points to left page
_____8. Directionality (top/bottom)	**Page 7–8:** ***There's four lines on this page (point to each). Which one do I read first?*** ***After administering task:*** Put finger on first line and say: ***I read this one first.***	1 pt: points to top line
_____9. Directionality RL(top/bottom)	**Page 7–8:** ***Which one do I read last?***	1 pt: points to bottom line
_____10. Print function	**Page 9–10:** *Point to the words spoken by the ducks in the water, and say:* ***Why are there all these words in the water?***	1 pt: tells that words are what ducks say or similar (e.g., *"because they are talking"*)
_____11a. Letter concept	**Page 11–12:** ***Show me just one letter on this page.***	1 pt: points to one letter
_____11b. First letter	**Page 11–12. Show me the first letter on this page.**	1 pt: points to first letter
_____11c. Capital letter	**Page 11–12.** ***Now show me a capital letter.***	1 pt: points to capital letter
_____12. Print function	**Page 23–24:** ***And the fox says "stupid ducks." Where does it say that?***	2 pt: points to fox's words 1 pt: points to other print on page

OBSERVATIONS

___Difficulty attending

___Asked for repetition

___Timid or reticent

___Difficult to understand

Other:_______________

SCORING INSTRUCTIONS

Total Raw Score:_________

Add the numbers to the left of each item in the Item column.

Print-Concept Knowledge (PCK) Estimate:________

Use the scale provided below to convert total raw scores to PCK Estimates

Raw Score	*PCK Estimate*	*Raw Score*	*PCK Estimate*	*Raw Score*	*PCK Estimate*
0	46	6	97	12	118
1	63	7	100	13	123
2	74	8	104	14	128
3	82	9	107	15	134
4	88	10	111	16	145
5	92	11	115	17	161