

ASSESSMENT OF PRESCHOOL EARLY LITERACY SKILLS: LINKING CHILDREN'S EDUCATIONAL NEEDS WITH EMPIRICALLY SUPPORTED INSTRUCTIONAL ACTIVITIES

CHRISTOPHER J. LONIGAN, NICHOLAS P. ALLAN, AND MATTHEW D. LERNER

Florida State University

The importance of the preschool period in becoming a skilled reader is highlighted by a significant body of evidence that preschool children's development in the areas of oral language, phonological awareness, and print knowledge is predictive of how well they will learn to read once they are exposed to formal reading instruction in elementary school. Although there are now a number of empirically supported instructional activities for helping children who are at risk of later reading difficulties to acquire these early literacy skills, limitations in instructional time and opportunities in most preschool settings require the use of valid assessment procedures to ensure that instructional resources are utilized efficiently. In this article, we discuss the degree to which informal, diagnostic, screening, and progress-monitoring assessments of preschool early literacy skills can inform instructional decisions by considering the strengths and weaknesses of each approach to assessment. © 2011 Wiley Periodicals, Inc.

Results of the National Assessment of Educational Progress (NAEP), the Nation's Report Card (National Center for Educational Statistics, 2009), indicated that among fourth-grade children in the United States, only 33% performed at or above the proficient level in reading and 33% performed below the basic level in reading. NAEP results for eighth-grade children were similar. Although it is not uncommon for these NAEP statistics on reading proficiency to be interpreted as indicating that schools are doing an increasingly poor job of teaching children to read, NAEP results across years actually reveal that the percentage of children performing at or above the proficient level has increased and that the percentage of children scoring below the basic level has decreased. The problem is not that schools are worse at teaching children to read now than they used to be; it is that the societal demands for literacy are increasing. These results, however, do indicate that a substantial proportion of children are not acquiring the levels of literacy skills needed to meet the demands for literacy they will face in school, the workplace, and elsewhere.

DEVELOPMENT AND SIGNIFICANCE OF EARLY LITERACY SKILLS

Acquiring well-developed literacy skills is a critically important developmental milestone for children. A large body of research highlights the negative consequences associated with delayed or disordered acquisition of reading skills. Children who are poor readers tend to continue to struggle with reading and writing, and read less than their peers who are more skilled in reading. As a consequence, children who are poor readers tend to receive less practice in reading and less exposure to content knowledge, vocabulary, and other language skills than do children who learn to read early and well (Cunningham & Stanovich, 1998; Echols, West, Stanovich, & Zehr, 1996; Morrison, Smith, & Dow-Ehrensberger, 1995). Reading and reading-related skills tend to be stable from an early point in school (Wagner, Torgesen, & Rashotte, 1994; Wagner et al., 1997). Longitudinal studies indicate

Christopher J. Lonigan, Department of Psychology and the Florida Center for Reading Research, Florida State University; Nicholas P. Allan and Matthew D. Lerner, Department of Psychology, Florida State University.

Preparation of this work was supported, in part, by grants from National Institute of Child Health and Human Development (HD052120) and the Institute of Education Sciences (R305B04074, R305B090021, R324E06086). Views expressed herein are solely those of the authors and have not been reviewed or cleared by the grantors.

Correspondence to: Christopher J. Lonigan, Department of Psychology, Florida State University, 1107 W. Call Street, Tallahassee FL 32306-4301. E-mail: lonigan@psy.fsu.edu.

that children who are poor readers by the end of the first grade almost never acquire average-level reading skills by the end of elementary school without substantial and sustained remediation efforts (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Juel, 1988; Torgesen & Burgess, 1998).

Knowledge about the causes, correlates, and predictors of children's reading success and reading failure in the early elementary grades has expanded greatly in the past several decades (e.g., National Reading Panel, 2000; Snow, Burns, & Griffin, 1998). This knowledge has been incorporated into methods of identifying, monitoring, and helping struggling readers in the elementary school grades. More recently, research has highlighted the significance of the preschool period for the development of skills that contribute to children's acquisition of reading skills (Whitehurst & Lonigan, 1998). The National Early Literacy Panel (NELP; see Lonigan, Schatschneider, & Westberg, 2008a) conducted a meta-analysis of approximately 300 studies that included data about the predictive relation between a skill measured in preschool or kindergarten and reading outcomes (i.e., word decoding, reading comprehension, spelling) for children learning to read in an alphabetic language. The results of this meta-analysis indicated that children's skills related to print knowledge (e.g., alphabet knowledge, print concepts), phonological processing skills (i.e., phonological awareness, phonological access to lexical store, phonological memory), and aspects of oral language (e.g., vocabulary, syntax/grammar, word knowledge) were substantive and independent predictors of children's later reading outcomes.

Similar to results with older children, data from longitudinal studies revealed a high degree of continuity between the levels of reading-related skills displayed by preschool children and the levels of reading-related and reading skills displayed by these children when they are in elementary school (e.g., Lonigan, Burgess, & Anthony, 2000; Storch & Whitehurst, 2002), indicating that the developmental antecedents that underlie the acquisition of reading are found early and prior to the onset of formal schooling. Many children enter kindergarten with well-developed print knowledge, phonological processing skills, and oral language skills, and these children are poised to "crack" the alphabetic code and become skilled readers when provided with effective reading instruction; however, a significant number of children arrive at kindergarten with low levels of these early skills, making it less likely that they will become skilled readers with the typical instruction provided in the early elementary grades.

INSTRUCTIONAL ACTIVITIES TO PROMOTE EARLY LITERACY SKILLS

Although not as well researched as interventions for older children, there is a growing body of research supporting the use of instructional practices and activities to promote preschool children's early literacy skills (see Lonigan, Shanahan, & Cunningham, 2008; Lonigan, Schatschneider, & Westberg, 2008b; see also the What Works Clearinghouse [WWC] website at www.whatworks.ed.gov). These research reviews identify instructional practices that are supported by causally interpretable evidence (i.e., research studies that rule out alternative explanations for observed gains in children's skills). The types of instructional practices related to reading outcomes that are supported by research can be grouped into two categories: those that promote the skills primarily associated with decoding print (i.e., code-related skills) and those that promote the skills primarily associated with comprehending what is read (i.e., meaning-related skills).

Interventions for Code-Related Skills

Both the NELP and the WWC reported that instructional activities designed to promote young children's phonological awareness (PA) resulted in children acquiring more PA skills, leading to better reading outcomes, particularly when these activities were combined with instructional activities designed to teach children about print. A detailed description of these activities is beyond the scope of this article; however, Phillips, Clancy-Menchetti, and Lonigan (2008) provided an

accessible discussion and guidelines for how these instructional activities could be used in preschool classrooms. Activities to promote PA most often involve some method of requiring children to break words apart into smaller sound units (e.g., syllables, phonemes) or to put smaller sound units together to make words. To date, causally interpretable studies that provide the evidence for the positive impact of PA instruction have included only activities done with individual children or with small groups of children. There is currently no evidence from studies of this type that these activities done with large groups of children or that implicit PA activities (e.g., whole-group syllable-clapping activities, singing word-play songs) result in growth in children's PA.

As mentioned previously, evidence indicates that combining PA activities with activities that focus on print seems to result in the best outcomes for reading-related and reading skills. Surprisingly, there are more studies concerning the effects of PA interventions or combined PA and print interventions than there are concerning the effects of just teaching children about print. Studies of combined PA and print activities demonstrate that children's print knowledge increases because of the interventions. The few available studies of teaching children about letters alone also indicate that children acquire more print knowledge than do children who do not receive such instruction (Lonigan, 2004), and these studies seem to support an advantage of instructional activities that include teaching about both letter names and letter sounds (Piasta, Purpura, & Wagner, 2010).

Interventions for Meaning-Related Skills

For preschoolers, meaning-related skills fall into the oral language skill domain (e.g., vocabulary, syntax) because most preschoolers are not yet reading. Although there is a lot of discussion of how early childhood education professionals should promote children's oral language skills, there is, at present, little of the type of research that allows causal conclusions to guide selection of instructional activities that will achieve this outcome. One instructional technique that is supported by a substantial body of research is a specific type of shared reading called "dialogic reading" (Arnold, Lonigan, Whitehurst, & Epstein, 1994; Whitehurst & Lonigan, 1998). In dialogic reading, picture books are used as an organizing prop for adults (teachers or parents) to assess, teach, and model vocabulary and grammar. For example, instead of reading the book, a teacher asks children to label objects, actions, and attributes of things pictured by responding to "Wh-type" questions (e.g., What is this called? What is she doing? What color is the duck?). Once it is known that children have the vocabulary of the book, teachers encourage children to help with the "reading" of the book by using open-ended prompts (e.g., What's happening on this page? What's going on here?). Teachers use repetition, correction, and expansions of children's answers to model increasingly complex verbal exchanges. Evidence summarized by both the NELP and the WWC indicates that dialogic reading, typically conducted with small groups of three to five children in classroom settings, is an effective means for teachers to promote preschool children's vocabulary skills.

IDENTIFYING CHILDREN NEEDING ADDITIONAL SUPPORT IN THE DEVELOPMENT OF EARLY LITERACY SKILLS

Children arrive in preschool with varying levels of early literacy skills. Depending on where they start, their experiences in the home, and the curriculum being used in their classroom, many children will leave preschool with early literacy skills that put them on a trajectory to transition successfully to learning to read. For some children, however, the support provided by typical classroom practices will not be sufficient for them to acquire these well-developed early literacy skills. Although there are effective instructional practices that can support this development, it is unlikely that early childhood educators in most preschool classrooms have sufficient time to provide this level of instruction for all children in their classrooms. Moreover, depending on the population of children served, a majority

of children do not need this extra level of instructional support. Consequently, a means of identifying those children who are either starting from a low level of skill or are not making sufficient gains in these skills to catch up, or both, is needed. This identification process is where the assessment of children's early literacy skills fits into an integrated system of identification and intervention.

There are three primary forms of assessment that can be used to identify children with specific education needs or to develop education goals, informal assessment, screening/progress monitoring, and diagnostic assessment. The latter two forms of assessment typically involve the use of standardized measures. In the remainder of this article, these three forms of assessment are described, including their typical use, information concerning the technical adequacy of the measures, and evidence of instructional utility.

Technical Adequacy of Assessments

The accuracy of assessment results is one of the key issues when determining the value of any particular test. Obviously, decisions made based on inaccurate information will be of little benefit to a child and may prevent more beneficial decisions from being made. The accuracy of an assessment is typically referred to as the psychometrics of the measure. The two key psychometric characteristics of an assessment are its reliability (i.e., consistency of measurement) and validity (i.e., the degree to which an assessment measures what it is intended to measure). In terms of reliability, measures with low internal consistency (i.e., $< .70$), indicating that the items on the measure are not good indicators of a single construct, or low test-retest reliability, indicating that the measure captures significant variance unrelated to the construct of interest, are unlikely to provide sufficiently accurate information on which to base educational decisions. In terms of validity, measures with low concurrent or predictive validity coefficients or high relative discriminative validity coefficients (i.e., a print knowledge assessment that correlated as highly with a measure of general intelligence as it did with an assessment of letter knowledge) are unlikely to provide teachers with the information on which children would benefit from additional instruction in specific skill domains.

Informal Assessment

Traditional methods of determining the skills of preschool children often involve observations of children by the classroom teacher. Such informal assessments of children's skills can be useful for teachers as they teach new skills because they can be used to scaffold instructional activities with a child (i.e., using a child's responses to an instructional interaction to increase or decrease the difficulty level of the interaction). Generally, this type of informal assessment is neither standardized nor highly structured; this type of assessment simply reflects a teacher's judgment based on casual observation of the child. Teachers' observational assessments can include the use of checklists, rating scales, or portfolios of children's work products. Some of these methods provide a means of comparison against a normative expectation; however, they often do not. When no comparison metric is provided, it is difficult to determine whether a child has made or is making adequate progress in key skill areas.

Several studies have investigated the accuracy of teachers' ratings of children's academic skills. In these studies, teachers typically complete some form of rating scale on children's skills, and the children are administered a standardized measure of the relevant skills. Much of this research has involved children in kindergarten or later grades. For instance, Hecht and Greenfield (2001) reported that teachers' ratings of first-grade students' academic competence using the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) were significantly and moderately correlated with both concurrent and third-grade direct assessments of reading and reading-related skills. Similarly, Beswick, Willms, and Sloat (2005) reported a moderate and significant correlation between teachers'

ratings of kindergarten students' reading-related skills and a concurrent direct assessment of reading. In contrast to these positive findings, Graney (2008) found that second-grade teachers were likely to overestimate the rate of growth of children's reading skills, and they reported correlations of near-zero between teachers' ratings and direct assessments of growth in reading skills.

A small number of studies have examined the accuracy of teachers' ratings of preschool children's early literacy skills. Cabell, Justice, Zucker, and Kilday (2009) had preschool teachers complete the 26-item Pre-Literacy Rating Scale from the Clinical Evaluation of Language Fundamentals Preschool-Second Edition (Wiig, Secord, & Semel, 2004), and they administered direct assessments of print concepts, letter knowledge, and name writing to the children. Cabell et al. reported correlations that ranged from .43 to .60 among the three dimensions of the teacher-rating scale and the three direct assessments. In our own work (Farrington & Lonigan, 2010), we have found that preschool teachers' ratings of children's early literacy skills on the SSRS were correlated with direct assessments of children's vocabulary, print knowledge, and PA (correlations ranged from .32 to .65). However, examination of the correlations between the specific skill domains rated by teachers and scores on the different direct assessments suggested that teacher ratings had only limited discriminant validity. Although teachers' ratings of children's print knowledge was more strongly correlated with direct assessments of print knowledge than with direct assessments of vocabulary or PA, teachers' ratings of children's oral language skills were correlated equally with direct assessments of vocabulary, PA, and print knowledge. Consequently, although these measures can identify children with low levels of skills in general, they do not accurately identify the specific skill domains in which children have strengths and weaknesses.

Other commonly used indirect assessments of preschool children are related to children's behaviors or behavior problems. Most of these assessments involve teachers completing rating scales of specific child behaviors (e.g., inattention, oppositionality, hyperactivity, anxiety/withdrawal, anger/aggression), and there are varying degrees of evidence for the reliability and validity of these measures. Ratings of problem behaviors—with the exception of ratings of inattention, however, are not particularly good indicators of how children might perform on later reading-related tasks (e.g., see Duncan et al., 2007).

The Creative Curriculum Developmental Continuum Assessment is a teacher-report indirect assessment that includes ratings of both behaviors and academic skills. This measure is widely used in preschools that use the Creative Curriculum as well as other curricula, and it includes 50 items that are rated by teachers using a 4-point scale. To our knowledge, there is only a single unpublished technical report concerning the psychometric properties of the measure. Lambert (2004) conducted a factor analysis of the 50 items with a sample of 1,590 3- to 5-year-old children and reported that there were four interpretable factors (with 3 items not associated with any factor), Social Development (13 items), Physical Development (7 items), Cognitive Development (23 items), and Expressive Language Development (4 items). The Cognitive Development factor included all items related to reading-related skills (e.g., demonstrates knowledge of the alphabet, writes letters and words, comprehends and interprets meaning from books and other texts), items related to math skills (e.g., uses numbers and counting, uses one-to-one correspondence), items related to language (hears and discriminates the sounds of language), and other content (e.g., observes objects and events with curiosity, classifies objects, makes believe with objects). Lambert reported high internal consistencies for the scales created from the factors (α s > .92), which is exactly what would be expected when a scale is created based on a factor analysis and its reliability is determined in the same sample. No evidence for the convergent, predictive, or discriminant validity of the measure or its subscales was reported.

Informal assessments are relatively easy to create and use and can serve as broad screening measures to identify children in need of more rigorous assessments such as diagnostic assessments.

There are several considerations to be aware of when using these informal assessments. Because informal assessments typically do not use a standardized procedure, the conditions of elicitation of children's skills are not uniform across children. Therefore, whether or not a child exhibits a particular skill may be due to the child's skill or the eliciting context, or both. In cases where norms are not available, the meaning of the level of skills observed is unknown, and there is a general dearth of information concerning the reliability and validity of these measures. Some of these measures do have evidence for reliability and validity—at least at a broad level. They identify children who have generally good or generally poor early literacy skills. However, there is limited evidence that these measures provide specific valid information for the different domains of early literacy (e.g., weakness in PA versus weakness in oral language). The lack of precision concerning specific areas of strengths or weaknesses in the different early literacy domains (i.e., oral language, PA, print knowledge) does not allow the teacher to match accurately the available empirically supported interventions, which are domain specific, with the child's need. Therefore, these informal assessments can be potentially useful initial methods of identifying children in need of more in-depth assessment measures.

Standardized Assessments

A standardized measure is one in which a common set of stimulus materials and questions, a consistent set of administration procedures, and conventional scoring procedures are used, and the scores are often based on a norming sample. Standardized measures allow meaningful comparisons among children (or between assessments of a single child over time) because they (a) have clear and consistent administration and scoring criteria (i.e., the measure is always given and scored in the same manner), (b) demonstrate generally good reliability and validity, and (c) raw scores are converted into scores that reflect a child's performance relative to the performance of a normative group. Because the assessment is the same for each child administered the measure, resultant scores have consistent meaning across children, examiners, and assessment sites. Standardized measures are usually normed within large and representative samples. Two types of standardized measures have potential utility for identifying children who need additional instructional support for early literacy skill development: screening/progress monitoring assessment and diagnostic assessment.

Screening Assessments

Screening assessments are typically brief measures that allow a snapshot of children's current skills. These measures are designed so that individuals who have minimal training in assessment can administer them. Results from screening assessments are often interpreted around one or more cut-scores that indicate a child's relative likelihood of needing additional assessment, more careful monitoring, or additional instruction.

There are two currently available screening tools of children's early literacy skills, the Get Ready to Read! Revised Screening Tool (GRTR-R; Whitehurst & Lonigan, 2001; Lonigan & Wilson, 2008) and the Individual Growth and Development Indicators (IGDIs; McConnell, 2002). Teachers can administer either of these screening assessments, and each usually takes less than 10 minutes to complete. The GRTR-R is a 25-item task that measures print knowledge and PA. Children are shown a page with four pictures for each item. The test administrator reads the question aloud and the child is asked to point to the correct picture. The IGDIs contains a number of tasks designed to measure an array of developmental domains from birth to approximately 8 years of age. The tasks relevant to early literacy are Alliteration and Rhyming (measures of PA), as well as Picture Naming (a measure of oral language). Each IGDIs task includes many items, and the score for a child is the number of items completed correctly within a 1- or 2-minute administration period. Other measures are described as screening assessments (e.g., Phonological Awareness Literacy Screening

Pre-K version; Invernizzi, Sullivan, & Meier, 2001); however, whereas teachers can administer this measure, its length (i.e., 121 items) is more like a diagnostic assessment than a screening assessment.

Studies of the psychometric properties of the GRTR-R support its internal consistency (e.g., $\alpha = .88$), as well as its concurrent (Molfese, Molfese, Modglin, Walker, & Neamon, 2004) and predictive validity (Phillips, Lonigan, & Wyatt, 2009; Wilson & Lonigan, 2009). Similarly, studies of the psychometric properties of the relevant IGDIs tasks support their reliability (e.g., 3-week test-retest $r_s = .44-.89$), concurrent validity, and predictive validity (Missall & McConnell, 2004; Wilson & Lonigan, 2009).

A central question concerning the utility of screening assessments is the accuracy with which they classify individuals into categories (e.g., at-risk vs. not at-risk). A good screening assessment will have a low number of false positives (e.g., children predicted by the screener to be at risk who actually are not at risk) and a low number of false negatives (e.g., children predicted by the screener to not be at risk who actually are at risk). Wilson and Lonigan (2010) compared the short-term predictive accuracy of the GRTR-R and IGDIs to a diagnostic assessment of early literacy skills in the same sample of preschool children. Whereas both measures performed similarly in identifying true positives (i.e., children with low levels of early literacy skills), the IGDIs produced a higher level of false positives (i.e., children predicted to have low levels of early literacy skills who did not) across outcomes than did the GRTR-R. Overall classification accuracy for both screening assessments was highest for the composite early literacy skill score (i.e., GRTR-R = .88, IGDIs = .48). Neither measure was particularly accurate in identifying low scores in specific domains of early literacy skills (i.e., overall classification accuracy ranged from .16 to .64).

Similar to some informal teacher-report assessments reviewed earlier, these results indicate that currently available screening assessments provide valid information to identify children who have generally good or generally poor early literacy skills, but they do not provide specific valid information for the different domains of early literacy. Importantly, teacher-report assessments have not been subjected to analyses of classification accuracy. Correlations between screening assessments and diagnostic assessments (i.e., GRTR-R $r_s = .74-.82$) are higher than those between teacher-report assessments and diagnostic measures (i.e., $r_s = .32-.65$), suggesting that teacher-report assessments have lower classification accuracy than does the GRTR-R. Overall, screening measures provide a time-effective, cost-effective, and valid means of identifying children with less well developed early literacy skills; however, at present, they do not provide the specific information needed to match instructional support to children's specific needs. Once identified as having low early literacy skills, children will need additional assessment (e.g., diagnostic assessment or other assessment) to determine their specific patterns of strengths and weaknesses to allow effective application of instructional support.

Progress-Monitoring Assessments

Like screening assessments, progress-monitoring assessments are brief measures that allow a snapshot of children's skills within a specific area. Instead of consisting of a single version of a measure, however, progress-monitoring assessments often have parallel forms (i.e., versions of the assessment that are equated so that scores have the same meaning across versions) that allow repeated assessments to determine children's growth in a skill area over time or as a result of instruction. Progress-monitoring assessments can be used as screening assessments (i.e., the first administration of a progress-monitoring assessment can determine which children will be monitored frequently or will receive additional instruction). To our knowledge, there are currently no widely available progress-monitoring assessments for preschoolers' early literacy skills. Most such assessments are developed for specific applications, such as use in a curriculum or for a specific program.

One example is the assessment system we created for a comprehensive preschool research curriculum that we developed (see Lonigan, Farver, Phillips, & Clancy-Menchetti, 2011). A recent analysis of the initial assessment measures in this system—essentially a screening assessment—yielded good evidence of reliability and validity for these teacher-administered assessments (Goodrich & Lonigan, 2010). Indices of internal consistency ranged from just adequate to high for the measures of vocabulary ($\alpha = .68$), PA ($\alpha = .82$), and print knowledge ($\alpha = .95$) skills. Correlations between the teacher-administered measures and diagnostic measures administered by trained research staff to 400 preschoolers were moderate to high for diagnostic assessments conducted at the beginning of the year (i.e., concurrent validity; $r_s = .33$ for PA, $.62$ for vocabulary, and $.71$ for print knowledge) and at the end of the year (i.e., predictive validity; $r_s = .24$ for PA, $.58$ for vocabulary, and $.66$ for print knowledge). At both assessment points, there was strong evidence of discriminant validity for the teacher-administered vocabulary and print knowledge measures.

A second example is an assessment system we are helping the State of Florida to develop for its universal pre-kindergarten program. All of the measures for this system were developed using an item-response theory framework. Three parallel forms of measures for oral language, PA, and print knowledge were created. Preliminary data from an ongoing field test provide strong evidence of reliability and validity. Across forms, indices of internal consistency were adequate to high ($\alpha_s = .75$ to $.86$) as were test-retest reliability coefficients ($r_s = .61$ to $.80$). Concurrent validity coefficients also were high for oral language ($r_s = .63$ – $.68$), PA ($r_s = .57$ – $.72$), and print knowledge ($r_s = .86$ – $.87$). Moreover, each progress-monitoring measure was more highly correlated with its target domain on the diagnostic assessment than with its nontarget domain (e.g., progress-monitoring print knowledge measures were more highly correlated with the diagnostic assessment of print knowledge than with the diagnostic assessments of PA or oral language), providing evidence of discriminant validity of the progress-monitoring measures.

Diagnostic Assessments

In general, diagnostic assessments are those standardized measures that provide highly detailed information about an individual skill area. Multiple items within the measure are intended to probe and explore different levels of competence or achievement of the skill. Some diagnostic assessments include multiple subtests, each intended to provide an index of strengths or weaknesses in a particular skill domain. Diagnostic measures tend to have very prescribed administration procedures, multiple items tapping a single construct, and basal and ceiling rules that allow optimal usage of assessment time in the region of a child's developmental level. The key advantages of diagnostic assessments include in-depth examination of specific skill areas or facets of a skill area, generally high reliability, established validity of the measure, and the ability to compare a specific child's performance with a known reference group. Hence, these standardized measures allow a meaningful, accurate, and in-depth determination of the early literacy skill areas in which a child has strengths or weaknesses relative to a developmental norm.

Exemplars of some commonly used standardized diagnostic measures for the assessment of preschool children's oral language, phonological processing, and print knowledge skills are listed in Table 1. As can be seen in the table, these measures have high levels of internal consistency and test-retest reliability, indicating that these diagnostic measures generally assess a single dimension and that they are relatively free of extraneous variance (i.e., error) that would result in inconsistent scores. Some opponents of using direct assessments of preschool children's skills argue that these skills cannot be reliably measured in young children; however, the available data suggest that the reliability of diagnostic measures used with preschool children is comparable to the reliability of similar measures used with older children. The validity data summarized in Table 1 also indicate

Table 1
 Domain, Administration, and Psychometric Information for Exemplars of Diagnostic Assessments for Oral Language, Phonological Processing, and Print Knowledge Skills

Name of Measure	Specific Skill Area Assessed	Age or Grade Range for Use	Time to Administer (Minutes)	Reliability	Validity
Measures of Oral Language Skills					
Clinical Evaluation of Language Fundamentals-Preschool-2	Expressive Vocabulary & Syntax, Receptive Vocabulary & Syntax	3 to 6 years, 11 months	15 to 20 (core tests)	.75 to .95 ^a .75 to .94 ^b	.69 to .88 ^c
Expressive One-Word Picture Vocabulary Test (2000 ed.)	Expressive Vocabulary	2 to 18 years, 11 months	10 to 15	.93 to .95 ^a .88 to .89 ^b	.64 to .71 ^c
Expressive Vocabulary Test-2	Expressive Vocabulary	2 years, 6 months to ≥90 years	10 to 20	.94 ^a .94 to .97 ^b	.51 to .82 ^c
Oral and Written Language Scales	Listening Comprehension, Oral Expression	3 to 21 years	40 to 60	.84 to .91 ^a .80 to .89 ^b	.46 to .91 ^c
Peabody Picture Vocabulary Test-4	Receptive Vocabulary	2 years, 6 months to ≥90 years	10 to 15	.94 ^a .93 ^b	.41 to .84 ^c .50 to .75 ^d
Preschool Language Scale-4	Auditory Comprehension, Expressive Communication	2 weeks to 7 years	15 to 40	.85 to .94 ^a .82 to .94 ^b	.66 to .88 ^c
Receptive One-Word Picture Vocabulary Test	Receptive Vocabulary	2 to 18 years, 11 months	10 to 15	.95 to .96 ^a .80 to .89 ^b	.65 to .71 ^c
Test of Language Development-Primary:3	Vocabulary, Syntax, Verbal Memory	4 to 8 years	30 to 60	.81 to .96 ^a .77 to .92 ^b	.52 to .97 ^c
Test of Preschool Early Literacy	Definitional Vocabulary	3 to 5 years, 11 months	10 to 15	.94 to .95 ^a .81 ^b	.71 ^c
Measures of Phonological Processing Skills					
Developing Skills Checklist	Phonological Awareness, Phonological Memory, Letter Knowledge	Pre-K to K	20 to 30	.81 to .92 ^a	.41 to .57 ^c

(Continued)

Table 1
Continued

Name of Measure	Specific Skill Area Assessed	Age or Grade Range for Use	Time to Administer (Minutes)	Reliability	Validity
Phonological Awareness and Literacy Screenings-Pre-K	Phonological Awareness	Pre-K	–	.77 to .93 ^a	.41 to .67 ^c .53 to .56 ^d
Phonological Awareness Test	Phonological Awareness	5 to 9 years, 11 months	40	.72 to .96 ^a .45 to .98 ^b	Contrasted Groups ^c
Test of Language Development-Primary:3	Phonological Awareness	4 to 8 years, 11 months	30	.89 to .94 ^a .77 to .87 ^b	.65 to .78 ^c
Test of Preschool Early Literacy	Phonological Awareness	3 to 5 years, 11 months	10	.86 to .88 ^a .83 ^b	.59 to .65 ^c .40 to .62 ^d
Woodcock Johnson III Tests of Cognitive Abilities and Achievement	Phonological Awareness, Phonological Memory, Phonological Access	2 to 90 years	2 to 10 per subtest	.82 to .98 ^a .57 to .77 ^b	.46 to .49 ^c
Developing Skills Checklist	Print Conventions, Writing	Measures of Print Knowledge Skills Pre-K to K	20 to 30		
Phonological Awareness and Literacy Screenings-Pre-K	Print Knowledge, Writing	Pre-K	–	.75 ^a	.41 to .67 ^c .53 to .56 ^d
Test of Early Reading Achievement-Third Edition	Alphabet Knowledge, Print Conventions, Print Meaning	3 years, 6 months to 8 years, 6 months	15 to 45	.82 to .95 ^a .86 to .99 ^b	.34 to .98 ^c .65 to .81 ^d
Test of Preschool Early Literacy	Print Knowledge	3 years to 5 years, 11 months	10	.93 to .96 ^a .89 ^b	.59 to .77 ^c .40 to .62 ^d
Woodcock Johnson III Tests of Cognitive Abilities and Achievement	Letter-Word Identification, Word Attack	2 to 90 years	2 to 10 per subtest	.93 to .99 ^a .79 to .91 ^b	.36 to .40 ^c
Woodcock Reading Mastery Test-R	Letter Knowledge, Decoding	5 years to ≥75	10 to 25	.94 to .99 ^a	

Note. ^aInternal consistency reliability; ^btest-retest reliability; ^cconcurrent validity; ^dpredictive validity. Pre-K = pre-kindergarten; K = kindergarten.

that these diagnostic measures provide valid information for the construct that they are intended to measure. Most measures listed in Table 1 yield standard scores that reflect a child's skill level on the measure relative to a nationally representative normative sample, allowing for comparisons of a child's score to scores of children of a similar age.

Although the strengths of diagnostic measures are their high reliability, good evidence of validity, and the ability to compare a child's scores to developmental expectations, these strengths often come at a cost. Diagnostic measures are typically fairly lengthy assessments. From Table 1, it can be seen that a comprehensive diagnostic assessment for the three primary domains of early literacy skills would take at least 30 minutes per child to complete, and such an assessment could take much longer depending on the particular tests used. Because of the expense of development, norming, and psychometric examination typically associated with a well-developed diagnostic measures, their purchase costs are relatively high. In addition to time and cost demands, most diagnostic measures require a degree of specialized training in their administration, scoring, and interpretation. In fact, many diagnostic measures are not available for purchase to those without evidence of a specific level of training in assessment.

These three characteristics associated with diagnostic measures most likely preclude their usage as a general means of identifying children who would benefit from additional instructional support in one or more key areas of early literacy skills. Moreover, most of these measures only have a single form, which precludes their use as a method of identifying the amount of progress children are making as a result of exposure to instructional activities. Whereas diagnostic assessments provide the "gold standard" for identifying individual children's patterns of strengths and weakness, as well as their absolute level of skill development relative to developmental expectations, some means of identifying the children who would most benefit from this type of in-depth assessment is a practical necessity. Informal assessments by teachers or measures developed as screeners or progress-monitoring tools can serve such a purpose.

INSTRUCTIONAL UTILITY OF EARLY LITERACY ASSESSMENTS

Glover and Albers (2007) noted that research on assessments must move beyond demonstrations of technical proficiency, such as reliability and validity. Screening and progress-monitoring assessments are developed to inform instructional decisions. Ultimately, the instructional utility of an assessment needs to be demonstrated by evidence that using screening or progress-monitoring assessments improves children's educational outcomes, either by use of the measure alone or by use of the measure in conjunction with an intervention. Currently, there is no strong evidence from preschool or elementary school that just the use of screening or progress-monitoring assessments leads to improved educational outcomes for children. There is a limited amount of evidence from elementary school studies that the use of screening or progress-monitoring assessments in conjunction with teacher-guided intervention efforts does result in improved educational outcomes. Results of a number of studies, however, indicate that teachers often fail to use the information from these assessments to guide instructional activities (Stecker, Fuchs, & Fuchs, 2005).

To our knowledge, there has been a single study of the instructional utility of screening or progress-monitoring assessments in a preschool setting. Landry, Anthony, Swank, and Monseque-Bailey (2009) reported a study in which 262 teachers across 158 schools were randomly assigned to one of five conditions: a control condition or one of four conditions crossing type of progress-monitoring assessment (paper-and-pencil vs. handheld computer) and in-class mentoring (present vs. absent). Overall, the study produced a complex pattern of results that did not yield evidence of an unconditional advantage of progress-monitoring assessment, in-class mentoring, or their combination. More research is needed to understand the utility of assessment methods for guiding interventions.

SUMMARY AND CONCLUSIONS

A growing body of research highlights the importance of the preschool period for becoming a skilled reader. Children's development in the areas of oral language, PA, and print knowledge is predictive of how well they will learn to read once they are exposed to formal reading instruction in elementary school. Even by the start of preschool, children vary considerably in their level of skill development in these three areas, and in the absence of strong instructional support, there is significant continuity between these early skills and reading outcomes. Children who enter preschool with low levels of early literacy skills are likely to be the children who will have difficulty learning to read when they are in kindergarten and first grade. At present, there are a number of instructional activities that have strong evidence for their positive effects on children's early literacy and reading skills; however, the effects of these interventions are specific to the domains they are intended to address.

The use of assessments that provide information on children's developmental achievements in key areas of early literacy can provide teachers with the information they need to provide optimal learning experiences for children. Of course, assessment is not an end in and of itself. It is one part of an identification, intervention, and evaluation sequence. Whereas accurate assessment can be a powerful tool for acquiring information, its value can only be realized in the context of a well-developed intervention program that translates the information obtained from assessments into curriculum modifications and specific instructional tactics and goals that are matched to the individual needs of a child. These assessments can more clearly focus educational activities on building key early literacy skills, enabling the targeting of skill areas in which a child needs the most help, and providing a means for determining whether instructional goals have been achieved.

The evidence reviewed in this article indicates that different methods and forms of assessments have specific strengths and weaknesses with respect to the kinds of information that is needed to provide children with optimal instructional support for the development of early literacy skills. Most currently available screening measures do not provide valid information about the specific domains of early literacy in which a child may have a particular strength or weakness. Consequently, these assessments are insufficient to allow the type of matching of instructional activities to children's educational needs that is required to improve children's educational outcomes. Diagnostic assessments provide the type of in-depth assessment needed to identify children's specific educational needs. Because of time and money costs, these assessments should be utilized following identification of children in need of this more targeted assessment. Preliminary evidence concerning newer progress-monitoring assessments that have been developed to assess specifically each of the key domains of early literacy is promising. Such assessments are likely to provide initial identification of children with slower development of early literacy skills, information concerning the specific skill areas that require targeted instruction, and a means to measure growth in these skills as a function of instruction.

Given currently available measures, a useful strategy would be for teachers to use a two-step process. Because the strengths of informal assessments and widely available screening measures are that they are brief and easy to use, some forms of these assessments can be used to identify children who have generally good or generally poor early literacy skills and to indicate which children are in need of further diagnostic assessment. For example, a screening assessment like the GRTR-R can be used to identify which children would most benefit from diagnostic assessment. Once these children are identified, a diagnostic assessment can be used to provide information about a child's specific strengths and weaknesses in key early literacy domains. With this information, teachers can make valid education decisions about what instructional activities or instructional adaptations are most likely to benefit the child.

REFERENCES

- Arnold, D. S., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture-book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86*, 235–243.
- Beswick, J. F., Willms, D., & Sloat, E. A. (2005). A comparative study of teacher ratings of emergent literacy skills and student performance on a standardized measure. *Education, 126*, 116–137.
- Cabell, S. Q., Justice, L. M., Zucker, T. A., & Kilday, C. R. (2009). Validity of teacher report for assessing the emergent literacy skills of at-risk preschoolers. *Language, Speech, and Hearing Services in Schools, 40*, 161–173.
- Cunningham, A. E., & Stanovich, K. E. (1998). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934–945.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., et al. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428–1446.
- Echols, L. D., West, R. F., Stanovich, K. E., & Zehr, K. S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88*, 296–304.
- Farrington, A. L., & Lonigan, C. J. (2010, June). Validity and domain specificity of parent and teacher reports of preschool children's academic competence. Paper presented at the Fifth Annual Institute of Education Sciences Research Conference, Washington, DC.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal individual growth curve analysis. *Journal of Educational Psychology, 88*, 3–17.
- Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117–135.
- Goodrich, J. M., & Lonigan, C. J. (2010, June). Accuracy of teacher-administered assessments of preschool children's early language and literacy skills. Paper presented at the Fifth Annual Institute of Education Sciences Research Conference, Washington, DC.
- Graney, S. B. (2008). General education teacher judgments of their low-performing students' short-term reading progress. *Psychology in the Schools, 45*, 537–549.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Hecht, S. A., & Greenfield, D. B. (2001). Comparing the predictive validity of first grade teacher ratings and reading-related tests on third grade levels of reading skills in young children exposed to poverty. *School Psychology Review, 30*, 50–69.
- Invernizzi, M., Sullivan, A., & Meier, J. (2001). *Phonological Awareness Literacy Screening-Pre-Kindergarten*. Charlottesville, VA: University of Virginia.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437–447.
- Lambert, R. G. (2004). *The Developmental Continuum Assessment System for ages 3 to 5: The assessment component of The Creative Curriculum® for Preschool; Technical Report*. Retrieved May 11, 2010 from www.teachingstrategies.com/content.pageDocs/Dev/Dev_Continuum.Technical_Report.pdf
- Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology, 101*, 448–465.
- Lonigan, C. J. (2004, August). Evaluation of components of a pre-literacy intervention with children at-risk. Paper presented at the Annual Convention of the American Psychological Association, Honolulu, HI.
- 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*, 596–613.
- Lonigan, C. J., Farver, J. M., Phillips, B. M., & Clancy-Menchetti, J. (2011). Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. *Reading and Writing, 24*, 305–337.
- Lonigan, C. J., Schatschnieder, C., & Westberg, L. (2008a). Impact of code-focused interventions on young children's early literacy skills. In *Developing early literacy: Report of the National Early Literacy Panel* (pp. 107–151). Washington, DC: National Institute for Literacy.
- Lonigan, C. J., Schatschnieder, C., & Westberg, L. (2008b). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. In *Developing Early Literacy: Report of the National Early Literacy Panel* (pp. 55–106). Washington, DC: National Institute for Literacy.
- Lonigan, C. J., Shanahan, T., & Cunningham, A. (2008). Impact of shared-reading interventions on young children's early literacy skills. In *Developing Early Literacy: Report of the National Early Literacy Panel* (pp. 153–171). Washington, DC: National Institute for Literacy.
- Lonigan, C. J., & Wilson, S. B. (2008). Report on the Revised Get Ready to Read! Screening Tool: Psychometrics and normative information. Technical Report. New York: National Center for Learning Disabilities.

- McConnell, S. R. (2002). *Individual Growth and Development Indicators*. Minneapolis, Minnesota: University of Minnesota.
- Missall, K. N., & McConnell, S. R. (2004). Psychometric characteristics of Individual Growth and Development Indicators: Picture naming, rhyme, and alliteration (Tech. Rep. No. 8). Minneapolis, MN: University of Minnesota, Center for Early Education and Development.
- Molfese, V. J., Molfese, D. L., Modglin, A. T., Walker, J., & Neamon, J. (2004). Screening early reading skills in preschool children: Get Ready to Read. *Journal of Psychoeducational Assessment*, 22, 136–150.
- Morrison, F. J., Smith, L., & Dow-Ehrensberger, M. (1995). Education and cognitive development: A natural experiment. *Developmental Psychology*, 31, 789–799.
- National Center for Education Statistics. (2009). The nation's report card: Reading 2009. Retrieved May 20, 2010 from <http://nces.ed.gov/nationsreportcard/pdf/main2009/2010458.pdf>
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read*. Washington, DC: U.S. Department of Health and Human Services.
- Phillips, B. M., Clancy-Menchetti, J., & Lonigan, C. J. (2008). Successful phonological awareness instruction with preschool children: Lessons from the classroom. *Topics in Early Childhood Special Education*, 28, 3–17.
- Phillips, B. M., Lonigan, C. J., & Wyatt, M. (2009). Predictive validity of the Get Ready to Read! screener: Concurrent and long-term relations with reading-related skills. *Journal of Learning Disabilities*, 42, 133–147.
- Piasta, S. B., Purpura, D. J., & Wagner, R. K. (2010). Fostering alphabet knowledge development: A comparison of two instructional approaches. *Reading and Writing*, 23, 607–626.
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, D.C.: National Academy Press.
- Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 42, 795–819.
- 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.
- Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 161–188). Mahwah, NJ: Lawrence Erlbaum.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, 30, 73–87.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, 33, 468–479.
- Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 69, 848–872.
- Whitehurst, G. J., & Lonigan, C. J. (2001). *Get Ready to Read! screening tool*. New York, NY: National Center for Learning Disabilities.
- Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals Preschool* (2nd ed.). San Antonio, TX: Harcourt Assessment.
- Wilson, S. B., & Lonigan, C. J. (2009). Emergent literacy screeners for preschool children: An evaluation of Get Ready To Read! and Individual Growth and Development Indicators. *Annals of Dyslexia*, 59, 115–131.
- Wilson, S. B., & Lonigan, C. J. (2010). Identifying preschool children at-risk of later reading difficulties: Evaluation of two emergent literacy screening tools. *Journal of Learning Disabilities*, 43, 62–76.