

# 3

## Identifying Early Literacy Learning Needs

### Implications for Child Outcome Standards and Assessment Systems

Elizabeth J. Spencer, Trina D. Spencer, Howard Goldstein, and Naomi Schneider

**T**he origins of conventional literacy skills are evident in early childhood development. Emergent literacy skills, as measured in preschool and kindergarten, are strong predictors of later literacy achievement. Current educational research and policy (e.g., No Child Left Behind Act of 2001, PL 107-110, 115 Stat. 1425, 2001) emphasize assessment of preschool children to inform identification and instruction. The assessment of emergent literacy skills can serve to identify those children who may be at risk for later reading difficulties. Furthermore, assessment can guide the content and delivery of early literacy instruction. Failure to identify children early and provide appropriate intervention to promote emergent literacy skills is likely to have serious repercussions for later development of conventional reading skills.

---

*Acknowledgment:* Preparation of this chapter was supported by Cooperative Agreement R324C080011 from the U.S. Department of Education, Institute of Education Sciences, awarded to The Ohio State University.

Effective early literacy assessment can provide valuable information. However, decisions about what to assess must be guided by evidence. The National Early Literacy Panel (NELP; 2008) report identifies those emergent literacy skills that are reliable predictors of later reading skill. In this chapter we review the predictors identified by NELP and make recommendations for assessment relative to those predictors. We discuss the different purposes of early literacy assessment and provide guidelines for the implementation of a measurement framework that encompasses these purposes in light of the availability of well-developed assessments. We also discuss the alignment of assessment with state early childhood education standards.

## **WHAT INFORMATION DOES THE NELP REPORT PROVIDE TO INFORM EARLY LITERACY ASSESSMENT?**

One purpose of the NELP report was to identify preschool and kindergarten predictors of conventional literacy skills (i.e., later reading, writing, and spelling outcomes). To address this research question, the panel conducted a meta-analysis of empirical studies published in refereed journals (see Chapter 1). The selected studies provided information to allow for the calculation of average correlations to identify predictors that were interpreted as strong (average correlations of .50 or larger), moderate (between .30–.49), and small (less than .30); a minimum of three studies examining a predictor variable were required to compute an effect size. In addition, the panel analyzed information provided by multivariate studies; these studies provided information about the strength of predictors when additional variables (e.g., IQ) were controlled.

The panel identified a set of skills that are precursors to later literacy achievement in decoding, reading comprehension, and spelling. A review of correlational evidence identified skills with high predictive validity, including *alphabet knowledge* (knowledge of letter names and sounds), *phonological awareness* (the ability to detect, manipulate, or analyze spoken words independent of meaning, including syllable and phoneme-level tasks), *rapid automatized naming* (the ability to rapidly name a repeating sequence of random sets of letters, numbers, colors, or pictures), *early writing or name writing* (the ability to write letters in isolation or write one's own name), and *phonological memory* (the ability to remember spoken information for a short time). Skills with moderate predictive validity included *concepts about print* (knowledge of print conventions and concepts, such as reading from left to right), *print knowledge* (combination of alphabet knowledge, concepts about print, and early decoding ability), *oral language* (the ability to produce and comprehend spoken language, including semantics

and syntax), *visual processing* (the ability to match or discriminate symbols), and *reading readiness* (combination of alphabet knowledge, concepts about print, vocabulary, memory, and phonological awareness).

As the authors of the NELP report point out, this approach is limited by the research available. If there is a lack of research on a particular emergent literacy skill, it is not possible to examine that skill as a predictor. Moreover, different statistical methods for identifying predictors of later literacy development may reveal different information (Paris & Luo, 2010).

The panel is careful to caution against drawing causal conclusions about the relationships between predictors and outcomes. We encourage this caution as well. Experimental research is needed to determine causality. The first step is to establish the efficacy of intervention approaches on what are thought to be important predictors. When there are robust intervention approaches, it is easier to determine their immediate and long-term effects on conventional literacy development (e.g., decoding, fluency, comprehension, writing). Identifying predictors is helpful but does not always fully inform the goal of promoting the full complement of literacy skills. For example, within the important domain of phonological awareness, phoneme awareness may be a stronger predictor of decoding than rhyme awareness (Macmillan, 2002).

At least three domains with predictive validity remain poor candidates for emergent literacy instruction given the current knowledge base: rapid automatized naming, phonological memory, and visual processing tasks. Interventions that target these skills and produce robust short- and long-term learning effects have yet to appear in the literature. Moreover, it is difficult to imagine practical tasks that could be taught to improve these skills in ways that relate to literacy development. Therefore, practitioners would be wise to focus on teaching skills that appear to have functional or causal relationships to later reading acquisition (e.g., phonological awareness).

Two additional domains with predictive validity warrant further explanation. Alphabet knowledge involves the naming of letters and their associated sounds. Identifying the names of letters, as an isolated skill, does not have a direct influence on learning to read. Learning letter names is a strong predictor of learning to read because it facilitates learning letter sounds (Ehri & Wilce, 1979), but naming letters without phonological awareness and letter-sound association has little effect on reading development. Concepts about print play a similar role in the development of reading. Knowing the directionality of print, differences between print and pictures, and other print conventions are indicators of children's familiarity with books and can help in learning other more critical literacy skills, but they do not have a direct causal link to read-

ing development (Neuman & Roskos, 2005). Although research indicates that concepts about print and letter names can be taught successfully to young children, in this chapter we focus on those predictors that have the greatest relevance for early literacy instruction.

The domains of phonological awareness, alphabet knowledge (with emphasis on letter–sound correspondence), oral language, and early writing are the focus of this chapter because the NELP report identifies them as moderate or strong predictors of later reading performance and because there is substantial evidence supporting their causal role in literacy development, thus highlighting their relevance to early literacy instruction. As we discuss the predictors of later reading in the context of early childhood assessment, the potential value of assessing the less practical predictors of reading ability (e.g., rapid automatized naming, concepts about print) should not be ignored. They may be especially good in discriminating among children who should be eligible for special education services, for example. On the other hand, practitioners might be expected to give little priority to tracking those skills. This is discussed further in the context of the various purposes of assessment described next.

## **EMERGENT LITERACY STANDARDS**

Recognition of critical emergent literacy skills has substantial implications for early childhood educational practices. One way the NELP report is likely to influence practice is through the development or revision of states' early learning guidelines (also called child outcome standards). With encouragement from federal initiatives to improve early childhood education such as *Good Start, Grow Smart* (White House, 2002), states began developing early childhood standards that resemble those mandated for K–12 education. Initially, though, many standards were crafted by a consensus of content experts instead of referencing research evidence (Neuman & Roskos, 2005). This is understandable for content domains lacking well-developed literature bases, but the release of the NELP report eliminates a lack of scientific evidence as a feasible excuse for neglecting key emergent literacy skills in state early childhood standards.

At present, all 50 states have developed, are developing, or are revising their early childhood guidelines (Barnett, Carolan, Fitzgerald, & Squires, 2011). Of those states that are implementing early learning standards, many have not included guidelines for all key emergent literacy skills. Some states, however, paid close attention to early literacy research and either created a new domain called emergent literacy (e.g., Florida) or expanded their language and literacy domain to include pho-

nological awareness, alphabet knowledge, print recognition, and writing strategies (e.g., California). California and Florida's early learning standards, for example, reflect rather comprehensive coverage of the essential precursors of conventional literacy skills.

The importance of including all of the critical emergent literacy skills is evident when considering the purpose of state standards. Child outcome standards describe the development and learning expectations for young children. Standards guide curriculum, assessment, and professional development (Bodrova, Leong, & Shore, 2004). Thus, standards prescribe what should happen in classrooms. In the standards-based education reform movement, standards are believed to lead to higher student achievement. In K–12 education, states that implemented standards-aligned instruction have shown improved student achievement (Education Commission of the States, 2000). Moreover, students taught by teachers whose professional development matched state standards and reform plans demonstrated impressive gains in reading (U.S. Department of Education, 2001). Early learning guidelines have the potential to have a similar impact for preschool children. However, if critical emergent literacy skills are neglected from the standards as key child outcomes, they are likely to be neglected in preschool classrooms as key instructional objectives. Because of the foundation that emergent literacy provides for later reading achievement, educational programs incur substantial risk of poor outcomes if they fail to teach skills identified in the NELP report as moderate or strong predictors of later reading performance (i.e., alphabet knowledge, phonological awareness, oral language, and early writing skills).

The identification of early predictors of later reading and writing achievement also amplifies the need to assess them. In general, educators and policy makers agree that assessment is an integral component of an effective early childhood educational program, but there is little agreement on how assessment should be carried out. For example, there is no consensus on how assessment data should be collected, who should collect assessment data, or how assessment information should be interpreted and reported. Several books and policy papers have addressed the challenges to effective early childhood assessments (Bagnato, Neisworth, & Pretti-Frontczak, 2010; Epstein, Schweinhart, DeBruin-Parecki, & Robin, 2004; National Association for the Education of Young Children [NAEYC] & National Association of Early Childhood Specialists in State Departments of Education [NAECS/SDE], 2009; National Research Council, 2008; Shepard, Kagan, & Wurtz, 1998). Concerns regarding the resources necessary to properly assess young children, the appropriateness of norm-referenced, standardized tests, and questionable reliability and validity associated with assessment alternatives for young children

are paramount (Bagnato et al., 2010; Burns, Midgette, Leong, & Bodrova, 2003; Epstein et al., 2004).

Perhaps the most fundamental and consistent recommendation is that assessment instruments should be used for their intended purposes. Using tests for reasons other than their intended purpose is an unfortunate and common misuse of assessment instruments in early childhood. To reduce this risk and help prepare all children for kindergarten, practitioners need psychometrically sound and socially valid instruments to accomplish each educationally relevant purpose.

In this climate of accountability and increasing calls for scientifically based education, effective applications of assessment will be essential for early identification and instruction of emergent literacy. The NELP report suggests which early literacy skills are worthy of thoughtful assessment. As a next step, this chapter describes three major purposes of assessment and focus on how those purposes relate to identifying early literacy needs for children in early childhood educational programs. In Table 3.1, we overlay the purposes of assessment onto the alphabet knowledge, phonological awareness, oral language, and early writing domains to help guide practitioners in the responsible use of available early literacy assessment instruments. Although it is not an exhaustive list, the table includes many of the assessment tools used in preschool for early literacy assessment. In addition, the following analysis of the intersection of assessment purposes and key emergent literacy skills helps identify areas of need in research and development.

## **PURPOSEFUL ASSESSMENT OF EMERGENT LITERACY SKILLS**

This section discusses three primary purposes of early childhood assessment: 1) informing instructional decisions, 2) identifying children who require intensified intervention, and 3) helping educational programs make systematic improvements (NAEYC & NAECS/SDE, 2009). In the context of early learning guidelines, the focus of assessment remains on promoting successful outcomes for children and facilitating positive programmatic changes (Bodrova et al., 2004).

### **Informing Instructional Decisions**

Early childhood educators use assessment data to inform two types of instructional decisions. Before delivering instruction, teachers first assess children's strengths and needs with respect to the classroom curriculum. This type of information helps identify what to teach and informs how to teach it (i.e., instructional planning). Once instruction begins, teachers use assessment data to monitor the effect their instruction has on

Table 3.1. Standardized early literacy instruments for preschoolers

Purpose	Phonological awareness	Alphabet knowledge	Oral language	Early writing	Composite of early literacy
Instructional planning	GGG PALS-PreK	PALS-PreK	GGG NLM:P	CELF-P2 supplemental subtest PALS-PreK	CELF-P2 PALS-PreK
Progress monitoring	GGG		GGG NLM:P		
Screening	GGG GRTR-R PALS-PreK	GRTR-R PALS-PreK	DELV-Screening Test Fluharty-2 GGG LAP-D PLS-4 Screen SPELT-3	PALS-PreK	GRTR-R PALS-PreK
Eligibility	CELF-P2 TOPEL	TERA-3 TOPEL	CELF-P2 DELV-Screening Test EVT LAP-D OWLS PLS-4 PPVT-4 TOLD-P:4 TOPEL		TERA-3 TOPEL
Program improvement	Program-level instruments: CLASS Pre-K, ECERS-R, ELLCO Pre-K				

Key: CELF-P2, Clinical Evaluation of Language Fundamentals, Preschool—Second Edition (Wiig, Secord, & Semel, 2004); CLASS Pre-K, Classroom Assessment Scoring System™ Pre-K (Pianta, La Paro, & Hamre, 2008); DELV-Screening Test, Diagnostic Evaluation of Language Variation-Screening Test (Seymour, Roeper, & de Villiers, 2003); ECERS-R, Early Childhood Environment Rating Scale—Revised Edition (Harms, Clifford, & Cryer, 2005); ELLCO Pre-K, Early Language and Literacy Classroom Observation Tool, Pre-K (Smith, Brady, & Anastasopoulos, 2008); EVT, Expressive Vocabulary Test—Second Edition (Williams, 2007); Fluharty-2, Fluharty Preschool Speech and Language Screening Test—Second Edition (Fluharty, 2000); GGG, Get it, Got it, Got it, Got it, Got it (Early Childhood Research Institute on Measuring Growth and Development, 1998); GRTR-R, Get Ready to Read! Revised (National Center for Learning Disabilities, 2009); LAP-D, Learning Accomplishment Profile—Diagnostic—Third Edition (Hardin, Peisner-Feinberg, & Weeks, 2005); NLM:P, Narrative Language Measures: Preschool (Spencer & Petersen, 2010); OWLS, Oral and Written Language Scales (Carow-Woolfolk, 1995); PALS-PreK, Phonological Awareness Literacy Screening for Preschool (Invernizzi, Sullivan, Meier, & Swank, 2004); PLS-4, Preschool Language Scale—Fourth Edition (Zimmerman, Steiner, & Pond, 2002); PLS-4 Screen, Preschool Language Scale-4 Screening Test—Fourth Edition (Zimmerman, Steiner, & Pond, 2005); PPVT-4, Peabody Picture Vocabulary Test—Fourth Edition (Dunn & Dunn, 2007); SPELT-3, Structured Photographic Expressive Language Test 3 (Dawson, Stout, & Eyer, 2003); TERA-3, Test of Early Reading Ability—Third Edition (Reid, Hresko, & Hammill, 2001); TOLD-P:4, Test of Language Development—Primary: Fourth Edition (Newcomer & Hammill, 2008); TOPEL, Test of Preschool Early Literacy (Lonigan, Wagner, Torgesen, & Rashotte, 2007).

student learning. Progress monitoring via repeated probes of student performance further informs decisions regarding when and how to make instructional adjustments.

In early childhood education, criterion-referenced tests (CRTs) that compare student performance to a set of preestablished learning objectives or criteria (Sattler, 2000) often are used to document students' progress and identify targets that need to be taught. Many comprehensive preschool curricula include companion CRTs that Bagnato et al. (2010) call curriculum-embedded assessments. Examples of curriculum-embedded CRTs include *The Creative Curriculum Developmental Continuum for Ages 3–5* (Dodge, Colker, & Heroman, 2006) and *The Carolina Curriculum for Preschoolers with Special Needs* (Johnson-Martin, Attermeier, & Hacker, 2004). Often, curriculum-embedded assessments lack standardized administration procedures, as well as evidence of reliability and validity. In contrast, early childhood developmental CRTs that are not companions to specific curricula are more likely to have standardized administration and scoring procedures and established norms so that they can also be used to determine eligibility for intensive services (e.g., the Battelle Developmental Inventory–Second Edition (BDI-2) [Newborg, 2005] and the BRIGANCE Inventory of Early Development II [Brigance & Glascoe, 2010]). Even though developmental CRTs and comprehensive curriculum-embedded CRTs can be used to plan instruction and monitor progress, the extent to which CRTs capture children's performance in key emergent literacy domains is restricted. Comprehensive curriculum-embedded CRTs and CRTs with norms typically cover multiple developmental domains such as personal-social, cognition, communication, fine motor, and gross motor domains. If included, emergent literacy skills may be buried among many other equally weighted skills. Thus, although general developmental inventories and curriculum-embedded tests are useful in early childhood assessment, they may not be sufficient for informing emergent literacy instruction.

As of 2012, there are only a few publicly available assessment tools designed to inform instructional decisions that target emergent literacy skills specifically and have adequate psychometric properties (see Table 3.1). The Phonological Awareness Literacy Screening for Preschool (PALS-PreK; Invernizzi, Sullivan, Meier, & Swank, 2004) has brief, simple, and standardized administration and scoring procedures and includes early writing, alphabet knowledge, print knowledge, and phonological awareness tasks. Another assessment tool, Get it, Got it, Go! (GGG; Early Childhood Research Institute on Measuring Growth and Development, 1998), also has brief and standardized administration and scoring procedures and assesses two of the domains identified by NELP: phonological awareness and oral language. GGG includes three



Individual Growth and Development Indicators (IGDIs): rhyming, alliteration (both of which are measures of phonological awareness), and picture naming (a measure of expressive vocabulary). A third assessment tool, Narrative Language Measures: Preschool (NLM:P; Spencer & Petersen, 2010) includes personal narrative, narrative retell, and story comprehension subtests, all of which fall into the oral language domain. The NLM:P administration and scoring procedures also are brief and standardized.

The assessment schedule or frequency of test administration affects the extent to which results can be used to inform instructional decisions. Assessment schedules vary according to decisions made at local levels and the measurement tools available. In early childhood education, an assessment schedule of three times per year (e.g., fall, winter, spring) is common because the majority of available tools have lengthy administration times, making them impractical for more frequent measurement (e.g., curriculum-embedded CRTs, developmental CRTs). The time necessary to administer PALS-PreK and GGG is sufficiently brief to be practical for more frequent monitoring of early literacy skills; however, neither assessment tool includes multiple equivalent forms to be used weekly. Repeated administrations of GGG are allowed, but the developers recommend that repetition be limited to once per month.

An infrequent assessment schedule may not be sufficient for all early literacy progress-monitoring needs or to inform strategic planning of emergent literacy instruction. Research has shown that frequent monitoring of students' progress enhances teachers' ability to plan instruction and make timely instructional changes that have positive effects on student achievement (Connor et al., 2009; Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, & Hamlett, 1993; VanDerHeyden, Snyder, Broussard, & Ramsdell, 2008). NLM:P was specifically designed for frequent monitoring of language growth over time and has 40 equivalent forms. In a recent study, researchers administered NLM:P daily to preschoolers receiving an oral language intervention and found the test to be sensitive to intervention effects (Spencer & Slocum, 2010).

To inform instructional decision making around early literacy, there is a need for instruments that measure preschoolers' performance on key early literacy skills, that have simple procedures and brief testing times, and that can be administered with a frequency sufficient to provide an index of progress. GGG and PALS-PreK meet several of these criteria. However, neither of these measures assesses all key domains (phonological awareness, alphabet knowledge, early writing, and oral language) nor were they designed for frequent administration. NLM:P has multiple equivalent forms for repeated administration, but it does not measure children's performance in areas of emergent literacy

besides language. Therefore, additional measures will need to be developed to inform instructional decision making.

In addition to measures that can be administered more frequently, there is a need for measures that assess multiple emergent literacy domains. Currently, researchers are developing measures that will be similar to the IGDIs tasks in GGG but that will assess additional domains, including alphabet knowledge and comprehension. Investigators in Minnesota are working on additional IGDIs that will be appropriate for assessing several emergent literacy skills (McConnell, Missall, Rodriguez, & Wackerle-Hollman, 2010). In contrast to the current GGG versions of IGDIs, the new items are scaled using an item response theory approach. This permits scaling of items for screening that covers a broad developmental age range. For the quarterly measures, the items selected can be scaled to take into account expected progress in critical language and literacy skills. Thus, the item pool would be different for different points in the school year or for different developmental levels.

There is a specific need for measures that can serve as progress-monitoring tools. Ideally, these measures could yield information to serve multiple purposes, assess skills in a number of early language and literacy domains, and align with early learning standards and other measurement tools. These measures should be able to be administered and scored quickly and reliably by practitioners (Deno, 2003). Additional research about the development of early literacy skills is necessary to guide the design of progress-monitoring measures that are appropriate for young children. For example, progress monitoring in the domain of early writing may not be appropriate for many preschool children if early writing skills are not expected to develop until late in the pre-kindergarten year.

Curriculum-based assessment (CBA) tools that include test stimuli drawn from the local classroom curriculum provide a viable alternative for monitoring progress to inform instruction. CBA is a general term encompassing methods to collect information about student performance in reference to the curriculum for the purpose of informing instruction (Tucker, 1985). Under this general umbrella of curriculum-relevant assessment, CBA can involve a variety of teacher-made tools such as observation recording forms, worksheets, and portfolios, as well as standardized, objective tests (McLoughlin & Lewis, 2008; Tucker, 1985). The direct correspondence between what is taught and what is assessed is an advantage of CBA. Teachers may use these tools to monitor mastery of the lessons taught each week or in each unit and use this information to differentiate instruction for children who may lack skills to progress to more advanced lessons. However, these mastery-monitoring CBAs may lack standardized administration and scoring procedures

and evaluations of their reliability and validity. Because this mastery-monitoring approach to assessment may provide limited information about the extent to which students have learned beyond the explicit context of the classroom curriculum (Fuchs & Deno, 1991), it should be supplemented with other standardized measures.

As is evident in Table 3.1, assessments for monitoring progress are limited in the phonological awareness and oral language areas and are absent in the alphabet knowledge and early writing areas. Consequently, early childhood education professionals may need to rely on general CRTs or CBAs that they develop themselves or that are recommended within existing curricula.

### **Identifying Children Who Require Intensified Intervention**

Identifying young children who need additional instructional support occurs in two ways. The traditional method involves screening and follow-up eligibility testing. Screening, typically a first step, involves a brief sampling of the young children's principal developmental skills for the purpose of detecting possible delays. If potential delays are detected, further in-depth eligibility assessment is conducted to determine the allocation of intensified intervention (Bagnato et al., 2010; National Research Council, 2008; Shepard et al., 1998).

Within the last decade, an alternative way to identify children who require intensified intervention has emerged in early childhood education based on response to intervention (RTI) conceptualizations. Although eligibility determinations are necessary before students receive special education services, early detection and prevention efforts such as RTI involve expanded options for children identified as needing intervention via screening measures. An at-risk identification at screening also could lead to increased monitoring or an immediate increase in instructional support without necessitating time- and resource-intensive eligibility assessments. As part of a RTI framework, universal screening occurs on a quarterly schedule (consistent with common early childhood assessment schedules). Assessment that is carried out in fall, winter, and spring is sometimes called benchmarking because students' development, skills, and achievements are compared with specific criteria or benchmarks for learning. In a RTI context, the extent to which student performance meets benchmarks and the extent to which students have progressed since the previous assessment point can be considered when determining students' needs for intensified literacy instruction.

The allocation of supplemental instruction and intervention is contingent upon screening and eligibility assessment results. Because financial and personnel resources necessary to provide intensified intervention

are valuable and scarce, the consequence of an identification error can be costly. Therefore, screening and eligibility assessment instruments have stringent psychometric requirements. Educators will want to select screening tools (e.g., *Get Ready to Read! Revised [GRTR-R]*; National Center for Learning Disabilities, 2009) and eligibility instruments (e.g., *Test of Preschool Early Literacy [TOPEL]*; Lonigan, Wagner, Torgesen, & Rashotte, 2007) with sufficient evidence of reliability and validity. With respect to emergent literacy skills, screening instruments also should have evidence of predictive validity with conventional reading and writing. Screening tools should involve standardized administration and scoring procedures and yield either criterion-referenced or norm-referenced scores to help determine when potential delays exist. Nearly all tests used as the primary method for determining eligibility for special education are norm-referenced and standardized. To identify children who require intensive intervention, such as special education, educators will want to select measures that also have evidence of good sensitivity and specificity. Sensitivity and specificity are indicators of a test's accuracy in identifying a condition. If a test is sensitive, a person with a condition will test positive for the condition on the measure. If a test is specific, a person without a condition will test negative for the condition on the measure. Test development methods, such as receiver operating characteristic (ROC) analysis (Catts, Petscher, Schatschneider, Bridges, & Mendoza, 2009; Compton et al., 2010; Johnson, Jenkins, & Petscher, 2010), have improved practitioners' ability to design instruments that optimize sensitivity and specificity, as well as the accuracy of predicting developmental delays.

**Screening Instruments** Several assessments are available for screening purposes in early language and literacy. As can be seen in Table 3.1, the domain with the most screening measures is oral language; the majority of measures provide a score cutpoint to identify children who need further assessment. Other screening instruments sample skills across multiple domains. For example, GRTR-R is a composite instrument that measures alphabet knowledge, concepts about print, and phonological awareness and provides a score cutpoint.

Several of the measures discussed in the previous section on informing instruction have the potential to be useful as a first step in the identification of children who require intensive intervention. PALS-PreK and GGG, administered in the fall, might provide information to educators about children who should receive additional assessment. However, neither PALS-PreK nor GGG provide score cutpoints or benchmarks to identify such children. PALS-PreK provides a developmental range for spring of the prekindergarten year. GGG suggests using local normative information to create benchmarks. Educators will need to

make decisions about a child's performance on these measures to determine if additional testing is necessary. There is a need for the further development of instruments for early literacy screening purposes, especially instruments that provide norm-referenced or benchmark scores to indicate children who may be eligible for additional intervention.

**Eligibility Instruments** There are many norm-referenced and standardized measures available for the purpose of eligibility determination in the domains of early language and literacy. Educators also will need to make careful decisions about the domains of early language and literacy that are assessed to determine eligibility. NELP has identified key predictors for assessment; educators will need to determine which of these predictors will be assessed to identify children who require intensified intervention. It may be most appropriate to assess those domains that are potential intervention targets. However, educators may supplement eligibility evaluations with measures of rapid automatized naming, phonological memory, and visual processing, which NELP identified as moderate or strong predictors of later reading. Although they are less functional for emergent literacy instruction, rapid automatized naming, phonological memory, and visual processing are good indicators of risk and may be helpful in the identification of children who require intensive literacy intervention (Weismer et al., 2000).

Often, measures that assess a broad range of early language and literacy skills will be most appropriate. For example, some students may struggle to acquire many early literacy skills, including alphabet knowledge, phonological awareness, and vocabulary. Other students may have a weakness only in a particular domain, such as oral language. Educators will need to select assessments that determine not only eligibility for additional services but also the type of services (e.g., early literacy intervention, speech-language services). Tests that provide both a composite score and scores for subtests that relate to particular domains might serve this purpose. Table 3.1 provides examples of available measures that can assist with eligibility determination and that have adequate psychometric properties. More comprehensive information about assessment tools can be found in the report *Early Childhood Assessment: Why, What, and How* (National Research Council, 2008).

TOPEL is an example of a measure that can be useful for the purpose of eligibility determination. It assesses skills in several of the domains identified by NELP and includes three subtests: print knowledge, definitional vocabulary, and phonological awareness. Children receive standard scores for each subtest and a composite score. Although the test manual reports strong reliability and validity, it does not provide information about sensitivity and specificity in identification.

NELP reports that measures of complex oral language skills (i.e., grammar and listening comprehension) have been found to be stronger predictors of later decoding and reading comprehension than simple measures of vocabulary. Therefore, measures of oral language that assess a broad range of skills may be most appropriate for determining children's eligibility for intensified intervention. The Clinical Evaluation of Language Fundamental, Preschool–Second Edition (CELF-P2; Wiig, Secord, & Semel, 2004) is an example of a widely used measure for determining eligibility in the domain of oral language. CELF-P2's core language subtests assess the language skills of sentence comprehension, word structure, and expressive vocabulary. Children receive standard scores for each subtest and a composite standard score based on performance on the core language subtests. CELF-P2 also has subtests and supplemental measures that could be used to assess other early literacy skills, including phonological awareness and early writing. The test manual provides strong evidence of reliability and validity. Sensitivity and specificity are high (.85 and .82, respectively) in identification of children with language disorders when the criterion for a disorder was set at 1 standard deviation below the mean.

Educators selecting assessment tools for oral language also will need to consider dialectical variations. Assessments of oral language, both for screening and eligibility purposes, may overidentify children who speak a dialectical variation as needing intervention. Measures such as the Diagnostic Evaluation of Language Variation—Screening Test (DELV–Screening Test; Seymour, Roeper, & de Villiers, 2003) have been developed to distinguish between children who are speakers of a dialectical variation and children who have language impairments.

In summary, assessments developed to serve the function of identifying children's eligibility for additional services comprise the greatest concentration of assessments in Table 3.1. This is especially evident in the oral language area. As can be seen in the table, early writing assessment has been included only in one screening assessment (PALS-PreK). It also was sampled in a supplemental subtest of the CELF-P2. The findings of the NELP report have provided an impetus for further development in this area. Puranik and Lonigan (2011) are among the investigators who are working on the development of a test of early writing skills.

## Helping Programs Make Systematic Improvements

Improving the quality of early childhood educational programs is a third purpose of assessment. However, the use of assessment data for the formative evaluation of a program's early literacy curriculum and instruction is not commonplace. A major reason for this is that few assessment

instruments are designed with this specific purpose in mind. Nonetheless, making data-based systematic changes can utilize child-level data combined with program-level measures of teacher behavior, literacy environment, and curriculum content (Epstein et al., 2004).

Although child-level data serve as a reasonable basis for program improvement, there are a number of issues to be considered. First, using child-level data should not be an afterthought. Instead, administrators should plan assessment data collection using valid designs and procedures to properly answer questions about program effectiveness. Second, it is not necessary to test all children in the program, which can be costly for programs with limited resources. With large or more homogeneous programs, sampling procedures can be used strategically to assess enough children to represent the population in the program. Third, child-level data should be aggregated in meaningful ways that reflect the impact of teachers, classrooms, or curricula. Examining an individual child's assessment data reflects the child's ability to learn, but examining a class's annual progress compared with a different class's annual progress may reflect differences in instructional quality. Fourth, children's gains over time as opposed to a static performance assessment provide the best estimate of programmatic impact. Assessment of program effectiveness for the purpose of program improvement is similar to progress monitoring of student performance to determine the effectiveness of that student's instruction, but it occurs on a much larger scale (National Research Council, 2008). Because growth cannot be established using a single assessment score, it is necessary to design program improvement measurements with at least two (beginning and end of year) or three (quarterly evaluation) data collection times across a year. Fifth, to effectively address program improvement goals, results of child-level data should be used to identify professional development needs, because results can be analyzed to reveal strengths and weaknesses in curriculum and instruction (Epstein et al., 2004; NAEYC & NAECS/SDE, 2009).

Many of the assessment instruments discussed in the previous two sections, if executed properly, can be used in the collection of child-level program improvement data. However, the results of these assessments offer only one source of information for evaluation. Program-level information, such as the quality of the classroom literacy environment, the breadth and depth of emergent literacy coverage in the program's curriculum, and the quality of teacher–student literacy interactions, also should inform systemic improvement efforts. These types of data should inform systematic and focused professional development and the selection of evidence-based curricula and instructional approaches. For example, programmatic data indicating that a teacher provides limited opportunities for shared book-reading experiences and makes little effort to expand



children's spoken vocabulary should lead to customized training and coaching on how to encourage vocabulary development and incorporate shared book reading into classroom activities. Likewise, if an examination of the program's curriculum finds that it does not include instructional suggestions and objectives for teaching phonological awareness, then selection of a curriculum that does is warranted.

A number of environmental inventories require raters to observe classroom environments to characterize the availability of materials and organization conducive to learning (e.g., the Early Childhood Environment Rating Scale–Revised Edition [ECERS-R; Harms, Clifford, & Cryer, 2005], the Early Language and Literacy Classroom Observation Tool, Pre-K [ELLCO Pre-K]; Smith, Brady, & Anastasopoulos, 2008). ELLCO Pre-K was developed for the purpose of characterizing the classroom literacy environment in particular. It consists of an observational checklist and supplemental teacher interview with items that relate to classroom structure, curriculum, language environment, books and book reading, and print and early writing. The content of ELLCO Pre-K overlaps with the predictors identified by NELP. Although these instruments give general information about the classroom environment, they are limited in their ability to capture details of instructional and classroom quality. For example, observational indicators in the domain of print and early writing describe the availability of writing materials, the display of written material, and opportunities for children to practice early writing skills. These indicators do not directly measure instruction in early writing.

The Classroom Assessment Scoring System, Pre-K (CLASS Pre-K; Pianta, La Paro, & Hamre, 2008) also is used to inform early childhood educational program improvement. This measure is an observational recording system designed to characterize teacher–student interactions in three domains: emotional support, classroom organization, and instructional support. CLASS Pre-K measures aspects of teacher–student interactions that relate to the development of early literacy skill. However, CLASS Pre-K is designed to measure characteristics of classroom interactions across all content domains.

To inform program improvement, there is a need for measures that can more accurately describe the instructional experiences of children in relation to the early language and literacy skills identified by NELP. Classroom CIRCLE: Code for Interactive Recording of Children's Learning Environments (Atwater, Lee, Montagna, Reynolds, & Tapia, 2009) is an example of an observational tool that can provide detailed information about instructional quality in the domain of early language and literacy. Using an event-recorder device (e.g., a personal digital assistant), the observer codes teacher and child behavior every 15 seconds for three



10-minute observations. The coding scheme allows observers to capture information on the focus of classroom instruction, the interactions of teachers and children, and the engagement, academic or otherwise, of children. Classroom CIRCLE yields a time-sampled record of teacher behavior and calculates the amount of time teachers spend teaching domains of early language and literacy instruction, including phonological awareness, alphabetic and print concepts, vocabulary, comprehension, or reading. This measure also provides estimates of the amount of time that children are engaged in early writing and early reading as well as other activities.

Several research groups have developed tools for the evaluation of preschool curricula. Some measures are designed to provide educators with a tool to examine curricular practices across content domains (e.g., socioemotional, mathematics, early literacy). For example, the Curriculum Rating Rubric (Pretti-Frontczak, Robbins, Jackson, Korey-Hirko, & Harjusola-Webb, 2008) allows educators to rate curricular practices as they relate to assessment, scope and sequence, activities and instruction, and progress monitoring. Other measures have focused on early language and literacy domains. Another curriculum evaluation tool, the Preschool Curriculum Review Rubric and Planning Tool (Virginia Department of Education, 2007) helps educators compare preschool curricula in the domains of oral language and vocabulary, phonological awareness, alphabet knowledge, print knowledge, and comprehension. Although the Preschool Curriculum Review Rubric and Planning Tool can provide information about early language and literacy instruction, the rubric's items are general and lack the level of specificity necessary to make decisions about program improvement. For example, in the domain of phonological awareness, the item related to rhyming is simply whether rhyming is taught in the curricula.

The extent to which curricula provide instructional support that teachers find useful (e.g., suggestions for explicit teaching strategies, recommendations for teacher- and child-led activities) is likely to guide programs in the selection of curricula. The Preschool Curricula Checklist (PCC; Kaminski & Carta, 2010) is a tool for examining the instructional design evident in preschool curricula in early language and literacy. Using the checklist, educators can evaluate the instructional support provided by a curriculum in the domains of phonological awareness, alphabet knowledge, vocabulary and oral language, and comprehension. The checklist can provide information about the scope of skills addressed, the sequence of lessons to address those skills, and the materials provided to teachers. PCC also gives an indication of the adaptations suggested for teachers to address the needs of children who struggle to acquire early literacy skills. Items on the checklist are

specific and can capture details that will assist educators in the selection of curricula. For example, in each domain educators rate the curriculum resources provided for implementing activities, including the materials required for the lesson or activity, a description of skills to be taught, suggested wording for how to teach the skills, and specific examples and content for teaching. Tools such as PCC can guide educators in the selection of curricula that provide strong support for instruction in early language and literacy.

In summary, evaluations of educational program quality can be informed by several sources of data: 1) test information about children's achievement and development, 2) instruments that rate the quality of literacy environments and teacher–student interactions, 3) instruments that summarize the extent to which teachers and students are engaged in literacy instruction, and 4) analyses of the adequacy of literacy instruction in classroom curricula. The results can be used to make decisions about targeted professional development and selection of new curricular programs.

## **EDUCATIONAL IMPLICATIONS**

This chapter is meant to help early education professionals construct a system of assessment that can address the multiple purposes of assessment: informing instructional decisions, identifying children who require intensive intervention, and helping educational programs make systematic improvements. The focus of the system of assessment should be on promoting successful outcomes for children and facilitating positive programmatic changes. Although the NELP findings help inform the selection of early language and literacy assessments that can guide decision making, they also highlight gaps in the availability of suitable assessments. Further research and development is necessary to provide educators with reliable and valid measures for assessment in early childhood. Optimally, teachers need measures that will allow them to easily and reliably assess how children are developing phonological awareness skills, alphabet knowledge (especially letter–sound correspondence), oral language, and early writing skills.

### **Recommendations for a System of Assessment**

In the next section, we make recommendations for a system of assessment that can address the multiple purposes of assessment. This system includes measures for universal screening, for instructional planning and progress monitoring, for determining eligibility, and for program improvement.

**Universal Screening** All children entering prekindergarten should be screened using measures that can be administered reliably and yield a score that can serve as a first step in identifying children who may need additional intervention. Many children enter early childhood educational programs with limited emergent literacy skills. For many children who have limited language and literacy experiences, early childhood education could facilitate rapid development of emergent literacy skills. Other children may fail to make significant progress. Thus, screening should be readministered on the same assessment schedule that is already common in early childhood education (i.e., fall, winter, and spring) to make sure that children who are struggling are detected in a timely manner. Aggregated universal screening results at the local level could inform program improvement and can be used to establish benchmarks to help identify children in need of intervention.

**Instructional Planning and Progress Monitoring** Measures selected for this purpose should be brief assessments that align closely with the instruction provided in the classroom, which in turn should align with early learning standards. All children should participate in quarterly progress monitoring (fall, winter, and spring). Depending on the instruments employed, universal screening and progress monitoring can be accomplished using the same tests (e.g., GGG IGDIs). Quarterly progress monitoring can address general classroom needs as well as the needs of individual children. First, performance of the group of children in a class can inform educators about the effectiveness of classroom instruction. For example, if a classroom of children has made progress in the domain of alphabet knowledge but demonstrates limited improvement in phonological awareness, educators can modify instruction to emphasize phonological awareness. Thus, instructional planning can be reflected in revisions in the scope and sequence of emergent literacy skills targeted in the general classroom curriculum. Second, quarterly progress monitoring can identify children who fail to respond to instruction or who are falling behind in certain emergent literacy areas. This would allow teachers to target areas of need for children who require additional instruction.

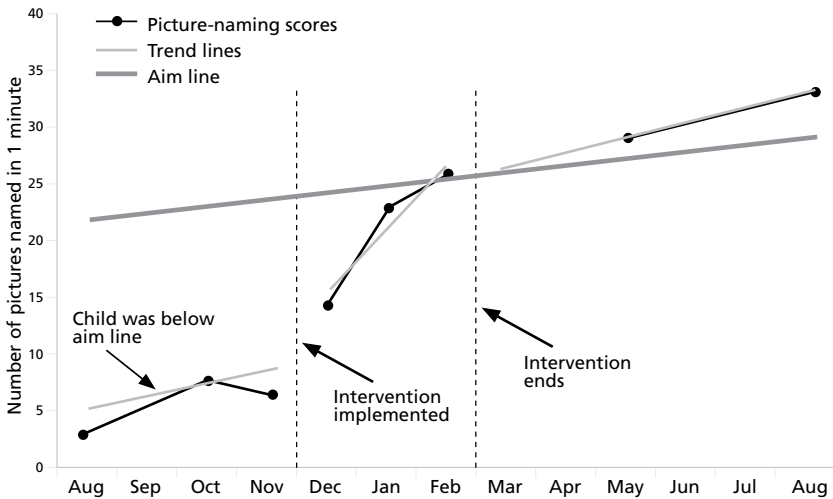
More frequent progress monitoring is necessary for children who have, or are at risk for, limited early language and/or early literacy skills, whether they receive intensified instruction or not. For example, a child who performed below benchmark on a screening measure at the beginning of preschool should be monitored more frequently to determine if he or she needs a more intensified level of instruction. Alternatively, educators can provide extra support immediately and monitor the child's progress.

In contrast to screening assessments that provide cutpoints, progress-monitoring assessments should provide a means of evaluating whether children are making progress consistent with their peers. For example, Figure 3.1 provides an example of vocabulary development based on the GGG picture-naming IGDI. The example shows a child who may have experienced limited home literacy opportunities and performed well below normative levels in August. After an opportunity to see the effects of the preschool's general curriculum, little progress was evident in October. With no progress evident a month later in November, the teacher decided that the child required supplemental language instruction. The effects of implementing this intervention were readily evident from IGDI measures over the next 3 months. Monthly IGDI assessments in December, January, and February reflected an upward trajectory as evidenced by the steep trend line. When the child reached the aim line for typical development in February, the supplemental language intervention ceased. IGDI measures were then administered quarterly, and the child showed development consistent with typical development at the end of the year and the beginning of the following school year. Figure 3.1 offers an illustration of how progress-monitoring measures could be examined to determine whether more intensive instruction is needed and whether it is successful in improving a child's developmental trajectory.

**Eligibility Assessment** Timely, comprehensive assessments are necessary to determine whether developmental delays are significant enough to warrant eligibility for intensive intervention services. Intensive intervention services typically include special education services (e.g., speech-language pathology services, reading specialists). Eligibility assessments work best in conjunction with screening and progress monitoring. If universal screening and progress-monitoring measures are in place, rather than a “wait-to-fail” model, educators should be able to efficiently identify the children for whom more information is needed. Consistent with RTI models, how well a child responds to small group or individual instruction may help determine the need for a comprehensive eligibility assessment. Assessments that are used to determine eligibility should yield information about the array of children's developmental needs. Therefore, these assessments need to be comprehensive, in that they assess across and within developmental domains. These assessments should serve to provide a profile with estimates of developmental status across domains that are indicative of areas of strength and the particular instructional needs of a child. Reading specialists, school psychologists, and speech-language pathologists are among the professionals who should have sufficient expertise in emergent literacy to contribute to this process.

**Program Improvement** Decisions about educational program improvement should be made using a combination of child-level assessments and program-level assessments of instructional quality and curricular support. Child-level data need to be reported in ways that can be meaningful for this purpose. For example, rather than only reporting the mean score of the classroom or center, distributions of child data should be provided. This information would inform program improvement that addresses the needs of all children in a classroom. Child-level data might identify classrooms in which teachers could benefit from professional development in a particular early language or literacy domain (e.g., phonological awareness) or in instructional strategies for a particular subgroup of children (e.g., English language learners). Data that are aggregated across the classrooms within a center might identify domains of early language and literacy that are not sufficiently addressed in the curriculum. As part of a coherent system of assessment, measures that inform program improvement should align with early learning standards.

In summary, an effective system of assessment in early language and literacy can only be put into place if appropriate measures exist to be a part of this system. As we have highlighted in this chapter, there is a need for further research to develop reliable, valid measures of early language and literacy.



**Figure 3.1.** An illustration of the use of progress-monitoring assessment to direct intervention efforts, based on the Get it, Got it, Go! picture-naming Individual Growth and Development Indicator (Early Childhood Research Institute on Measuring Growth and Development, 1998). The trend lines indicate the slope of growth before, during, and after intervention.

## CONCLUSION

An effective system of assessment in early language and literacy includes instruments that provide coherent information (National Research Council, 2008). Coherence should be demonstrated in multiple ways. First, educators should strive for consensus on goals for children's learning and the purposes of assessment. Second, educators should strive for assessments that align with early learning standards as well as curriculum and instruction. Third, educators should strive for assessments that provide in-depth information at the child's developmental level but have the ability to relate to a broader range of development. Fourth, educators should strive for alignment among measures. If educators are screening for a particular skill (e.g., phonological awareness), then they will want a measure that allows them to monitor progress of that skill.

One reason that educators may develop a system of assessment is to implement a RTI model. Although an in-depth discussion of RTI models is beyond the scope of this chapter, the recommendations we make for a system of assessment align closely with those that would be part of a RTI model. Additional recommendations for assessment in RTI for reading are available in the Institute of Education Sciences practice guide on assisting school-age children who are struggling with reading (Gersten et al., 2008). With an increasing acceptance of RTI in early childhood, it will be important to examine the extent to which assessment systems for school-age children apply to early childhood education (Greenwood et al., 2008). To a great extent, this chapter's recommendations for assessment of young children's early literacy skills are consistent with those applied in primary grades. Unfortunately, the availability of assessment tools in early childhood that can fulfill RTI assessment functions currently is limited. The importance of addressing literacy early, the extension of RTI to early childhood, and the need for increased early childhood assessment options make the development of preschool early literacy assessment instruments an urgent priority.

## REFERENCES

- Atwater, J., Lee, Y., Montagna, D., Reynolds, L.H., & Tapia, Y. (2009). *Classroom CIRCLE: Code for Interactive Recording of Children's Learning Environments*. Kansas City: University of Kansas, Juniper Gardens Children's Project.
- Bagnato, S.J., Neisworth, J.T., & Pretti-Frontczak, K. (2010). *LINKing authentic assessment and early childhood intervention: Best measures for best practices* (2nd ed.). Baltimore, MD: Paul H. Brookes Publishing Co.
- Barnett, W.S., Carolan, M.E., Fitzgerald, J., & Squires, J.H. (2011). *The State of Preschool 2011: State Preschool Yearbook*. National Institute for Early Education Research. Retrieved from <http://nieer.org/yearbook>

- Bodrova, E., Leong, D., & Shore, R. (2004, March). Child outcome standards in pre-k programs: What are standards; what is needed to make them work? *Preschool Policy Matters*. Retrieved from <http://nieer.org/resources/policybriefs/5.pdf>
- Brigance, A.H., & Glascoe, F.P. (2010). *BRIGANCE Inventory of Early Development II*. North Billerica, MA: Curriculum Associates.
- Burns, M.S., Midgette, K., Leong, D., & Bodrova, E. (2003). *Prekindergarten benchmarks for language and literacy: Progress made and challenges to be met*. Washington, DC: National Institute for Early Education Research.
- Carrow-Woolfolk, E. (1995). *Oral and Written Language Scales*. Circle Pines, MN: American Guidance Service.
- Catts, H., Petscher, Y., Schatschneider, C., Bridges, M., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities, 42*(2), 163–176.
- Compton, D., Fuchs, D., Fuchs, L., Bouton, B., Gilbert, J., Barquero, L., ... Crouch, R. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology, 102*(2), 327–340.
- Connor, C., Piasta, S., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., ... Morrison, F. (2009). Individualizing student instruction precisely: Effects of Child x Instruction interactions on first graders' literacy development. *Child Development, 80*(1), 77–100.
- Dawson, J., Stout, C., & Eyer, J. (2003). *Structured Photographic Expressive Language Test 3*. DeKalb, IL: Janelle.
- Deno, S.L. (2003). Developments in curriculum-based measurement. *Journal of Special Education, 37*, 184–192.
- Dodge, D.T., Colker, L.J., & Heroman, C. (2006). *The Creative Curriculum Developmental Continuum*. Washington, DC: Teaching Strategies.
- Dunn, L.M., & Dunn, D.M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition (PPVT-4)*. Minneapolis, MN: NCS Pearson Assessments.
- Early Childhood Research Institute on Measuring Growth and Development. (1998). *Research and development of individual growth and development indicators for children between birth and age eight* (Technical Report No. 4). Minneapolis: University of Minnesota, Center for Early Education and Development.
- Education Commission of the States. (2000). Setting the standard: Will higher expectations improve student achievement? *The Progress of Education Reform, 1*(5), 1–6. Denver, CO: Education Commission of the States.
- Ehri, L., & Wilce, L. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology, 71*(1), 26–40.
- Epstein, A.S., Schweinhart, L.J., DeBruin-Parecki, A., & Robin, K.B. (2004, July). Preschool assessment: A guide to developing a balanced approach. *Preschool Policy Matters*. Retrieved from <http://nieer.org/resources/policybriefs/7.pdf>
- Fluharty, N.B. (2000). *Fluharty Preschool Speech and Language Screening Test* (2nd ed.). Austin, TX: PRO-ED.
- Fuchs, L.S., & Deno, S.L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*(6), 488–500.
- Fuchs, L.S., Deno, S.L., & Mirkin, P. (1984). Effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and



- student awareness of learning. *American Educational Research Journal*, 21(2), 449–460.
- Fuchs, L.S., Fuchs, D., & Hamlett, C.L. (1993). Technological advances linking the assessment of students' academic proficiency to instructional planning. *Journal of Special Education Technology*, 12(1), 49–62.
- Gersten, R., Compton, D., Connor, C.M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W.D. (2008). *Assisting students struggling with reading: Response to intervention and multi-tier intervention for reading in the primary grades. A practice guide* (NCEE 2009-4045). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Greenwood, C.R., Carta, J.J., Baggett, K., Buzhardt, J., Walker, D., & Terry, B. (2008). Best practices integrating progress monitoring and response-to-intervention concepts into early childhood. In A. Thomas, J. Grimes, & J. Gruba (Eds.), *Best practices in school psychology V* (pp. 535–548). Washington, DC: National Association of School Psychology.
- Hardin, B.J., Peisner-Feinberg, E.S., & Weeks, S.W. (2005). *Learning Accomplishment Profile–Diagnostic (LAP-D)* (3rd ed.). Lewisville, NC: Kaplan Early Learning.
- Harms, T., Clifford, R.M., & Cryer, D. (2005). *Early Childhood Environment Rating Scale* (Rev. ed.). New York, NY: Teachers College Press.
- Invernizzi, M., Sullivan, A., Meier, J., & Swank, L. (2004). *Phonological Awareness Literacy Screening for Preschool*. Charlottesville: University of Virginia.
- Johnson, E., Jenkins, J., & Petscher, Y. (2010). Improving the accuracy of a direct route screening process. *Assessment for Effective Intervention*, 35(3), 131–140.
- Johnson-Martin, N.M., Attermeier, S.M., & Hacker, B.J. (2004). *The Carolina Curriculum for Preschoolers with Special Needs* (2nd ed.). Baltimore, MD: Paul H. Brookes Publishing Co.
- Kaminski, R., & Carta, J. (2010). *The Preschool Curricula Checklist*. Unpublished manuscript. Dynamic Measurement Group, Eugene, Oregon.
- Lonigan, C.L., Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (2007). *Test of Preschool Early Literacy (TOPEL)*. Austin, TX: PRO-ED.
- Macmillan, B. (2002). Rhyme and reading: A critical review of the research methodology. *Journal of Research in Reading*, 25, 4–42.
- McConnell, S., Missall, K., Rodriguez, M., & Wackerle-Hollman, A. (2010, February). Monitoring progress on progress monitoring: Recent innovations in the design and evaluation of individual growth and development indicators. Panel presentation at the 7th Biennial Conference on Research Innovations in Early Intervention (CRIEI), San Diego, CA.
- McLoughlin, J.A., & Lewis, R.B. (2008). *Assessing students with special needs* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education. (2009). *Where we stand on curriculum, assessment, and program evaluation*. Retrieved from <http://www.naeyc.org/files/naeyc/file/positions/StandCurrAss.pdf>
- National Center for Learning Disabilities. (2009). *Get Ready to Read! Revised*. Minneapolis, MN: NCS Pearson Assessments.



- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
- National Research Council. (2008). *Early childhood assessment: Why, what, and how*. Washington, DC: National Academies Press.
- Neuman, S.B., & Roskos, K. (2005). The state of state pre-kindergarten standards. *Early Childhood Research Quarterly*, 20, 125–145.
- Newborg, J. (2005). *Battelle Developmental Inventory—Second Edition (BDI-2)*. Itasca, IL: Riverside.
- Newcomer, P.L., & Hammill, D.D. (2008). *Test of Language Development—Primary: Fourth Edition (TOLD-P:4)*. Austin, TX: PRO-ED.
- No Child Left Behind Act of 2001, 20 U.S.C. §§ 6301 *et seq.*
- Paris, S., & Luo, S. (2010). Confounded statistical analyses hinder interpretation of the NELP report. *Educational Researcher*, 39, 316–322.
- Pianta, R.C., La Paro, K.M., & Hamre, B.K. (2008). *Classroom Assessment Scoring System (CLASS), Pre-K*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Pretti-Frontczak, K., Robbins, S.H., Jackson, S., Korey-Hirko, S., & Harjusola-Webb, S. (2008). *Curriculum Rating Rubric*. Kent, OH: Kent State University, Center for Excellence in Early Childhood Research and Training. Retrieved from [http://www.ehhs.kent.edu/odec/documents/CurriculumFrameworkRatingRubric\\_000.pdf](http://www.ehhs.kent.edu/odec/documents/CurriculumFrameworkRatingRubric_000.pdf)
- Puranik, C., & Lonigan, C.J. (2011). From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and Writing: An Interdisciplinary Journal*, 24(5), 567–589. doi:10.1007/s11145-009-9220-8
- Reid, D.K., Hresko, W.P., & Hammill, D.D. (2001). *Test of Early Reading Ability—Third Edition (TERA-3)*. Austin, TX: PRO-ED.
- Sattler, J.M. (2000). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M. Sattler.
- Seymour, H.N., Roeper, T.W., & de Villiers, J. (with de Villiers, P.). (2003). *Diagnostic Evaluation of Language Variation—Screening Test (DELV—Screening Test)*. San Antonio, TX: Harcourt Assessment.
- Shepard, L., Kagan, S.L., & Wurtz, E. (Eds.). (1998). *Principles and recommendations for early childhood assessments*. Washington, DC: National Education Goals Panel.
- Smith, M.W., Brady, J.P., & Anastasopoulos, L. (2008). *Early Language and Literacy Classroom Observation Tool, Pre-K (ELLCO Pre-K)*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Spencer, T.D., & Petersen, D.B. (2010). *Narrative Language Measures: Preschool*. Retrieved from <http://www.languagedynamicsgroup.com>
- Spencer, T.D., & Slocum, T.A. (2010). The effect of a narrative intervention on story retelling and personal story generation skills of preschoolers with risk factors and language delays. *Journal of Early Intervention*, 32, 178–199.
- Tucker, J.A. (1985). Curriculum-based assessment: An introduction. *Exceptional Children*, 52, 199–204.
- U.S. Department of Education. (2001). *The Longitudinal Evaluation of School Change and Performance (LESCP) in Title I schools*. Washington, DC: Author.
- VanDerHeyden, A., Snyder, P., Broussard, C., & Ramsdell, K. (2008). Measuring response to early literacy intervention with preschoolers at risk. *Topics in Early Childhood Special Education*, 27, 232–249.

- Virginia Department of Education. (2007). *Preschool Curriculum Review Rubric and Planning Tool*. Richmond: Department of Education, Commonwealth of Virginia.
- Weismer, S.E., Tomblin, J., Zhang, X., Buckwalter, P., Chynoweth, J., & Jones, M. (2000). Nonword repetition performance in school-aged children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 865–878.
- White House. (2002). *Good Start, Grow Smart: The Bush administration's early childhood initiative*. Washington, DC: Government Printing Office.
- Wiig, E.H., Secord, W.A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals—Preschool—Second Edition (CELF-Preschool-2)*. San Antonio, TX: Harcourt Assessment.
- Williams, K.T. (2007). *Expressive Vocabulary Test (EVT)* (2nd ed.). Minneapolis, MN: NCS Pearson Assessments.
- Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (2002). *Preschool Language Language Scale—Fourth Edition (PLS-4)*. San Antonio, TX: Harcourt Assessment.
- Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (2005). *Preschool Language Scale-4 Screening Test (PLS-4 Screening Test)*. San Antonio, TX: Harcourt Assessment.