POLICY
INFORMATION
PERSPECTIVE

# Early Literacy Assessment Systems:
## Essential Elements

by Jacqueline Jones


ETS
Educational Testing Service

Research & Development
Policy Information Center

Additional copies of this report can be ordered for $10.50 (prepaid) from:

Policy Information Center
Mail Stop 04-R
Educational Testing Service
Rosedale Road
Princeton, NJ 08541-0001
(609) 734-5694
e-mail: *pic@ets.org*
Copies can also be downloaded from:
*www.ets.org/research/pic*



June 2003

# Table of Contents

# Preface

As part of the No Child Left Behind Act, increasing attention is being paid to early literacy achievement. Even pre-kindergarten programs such as Head Start are being held accountable for the learning of very young children, most particularly with respect to early language and literacy acquisition. And certainly, we know that the achievement gap between those of different economic and ethnic/racial groups begins to evidence itself at the outset of formal education. My colleague Richard Coley presented a very fine analysis of the Early Childhood Longitudinal Survey last year in his Policy Information Center Report, *An Uneven Start: Indicators of Inequality in School Readiness*.

If we are to improve early literacy for all our young people, we are going to have to wisely make use of assessment. In *Early Literacy Assessment Systems: Essential Elements*, Jacqueline Jones carefully describes how assessment can support policy, teaching, and learning of those literacy skills that are the key determinants of individuals' future educational success.

Jones has worked with schools in South Brunswick, New Jersey, and New York City to develop systems that help teachers gain insight about their students' progress and make instructional decisions that improve student learning. For our youngest students, she convincingly urges us to avoid the seductive trap of relying on any single test to provide all the critical information needed to have an effective and accountable educational system. Rather, she helps us see how different information sources can be used together to fulfill different roles in providing critical information needed by different stakeholders in the system. But ultimately, she focuses most on how assessment can assist teachers in helping their students develop literacy skills.

Though this report focuses on assessments of early literacy, the lessons Jones provides are worth attending to for older students as well. We need fundamentally different forms of assessment to provide information appropriate to different needs. The granularity of assessment information needed by teachers is far different from that needed by policymakers. The qualities of educational leadership that are so compelling for schools inhabited by 5-year-olds are just as necessary for those attended by teenagers. If we can make the recommendations of Jacqueline Jones a reality in all our schools, we have the chance of achieving fully accountable, high-quality learning environments for all our students.

Drew Gitomer
Senior Vice President
Research and Development
Educational Testing Service

# Acknowledgements

This report was reviewed by Ted Chittenden, Richard Coley, Drew Gitomer, and Irv Sigel of Educational Testing Service; John Love of Mathematica Policy Research; Willa Spicer of the South Brunswick, New Jersey School Distict; and Ed Greene, Montclair State University. Lynn Jenkins was the editor; Carla Cooper provided desktop publishing; and Marita Gray designed the cover. Errors of fact or interpretation are the responsibility of the author.

# Introduction

*The ability to read and write is essential to successful participation in our society.*

Extraordinary attention is currently being given to early childhood education, with an emphasis on early literacy acquisition. Across the country, politicians, educators, and researchers are attempting to ensure that young children are provided with the most favorable opportunities to develop strong literacy skills. The desire to hold early childhood educators accountable for children's literacy acquisition is strong, and the accountability methods themselves have become a focus of discussion. This report will outline a system-wide framework for monitoring the literacy development of children in preschool through 2nd grade. Specific early literacy assessment instruments and instructional approaches will not be suggested. Rather, this report will focus on some of the essential elements of an assessment *system* intended to monitor the progress of young children's literacy development.

As a starting point, it is helpful to focus on two major factors that have led to the current emphasis on early literacy development and teacher accountability:

- Increased awareness of the importance of early development, and
- The achievement gap among kindergartners.

## Increased Awareness of the Importance of Early Development

Although U.S. public education has long been committed to K-12 education, there is a recent and growing emphasis on the importance of the first five years of life.[1] Merely providing a safe and nurturing environment for young children is no longer adequate. Greater attention is being paid to early cognitive development, with an emphasis on language and literacy. This shift has resulted from new insights into the extraordinary amount of learning that takes place during the first years of life. A subcommittee of the National Research Council and Institute of Medicine set out to update the state of knowledge on early development. Their final conclusions and recommendations were grounded on four broad themes:

- All children are born wired for feelings and ready to learn.
- Early environments matter, and nurturing relationships are essential.

[1] Committee for Economic Development, *Preschool for All: Investing in a Productive and Just Society*, New York, 2002; National Research Council, *Eager to Learn: Educating our Preschoolers*, Washington, DC: National Academy Press, 2001.

- Society is changing, and the needs of young children are also changing.

- Interactions among early childhood science, policy, and practice are problematic and demand dramatic rethinking.[2]

In sum, children enter the world trying to make sense of communication systems, rules of social interaction, and how things work. As a result, early childhood educators are now challenged to ensure that young children receive enriched cognitive, linguistic, and social-emotional stimulation even prior to the traditional age of compulsory education

## The Achievement Gap at Kindergarten

This new understanding of the importance of early learning opportunities has been accompanied by the realization that socioeconomic status can be an important factor in early language development. Economically advantaged children often demonstrate a significant lead in language development over their less economically privileged peers. Most disturbing has been the finding that these economically based discrepancies in language development can persist throughout the school years, resulting in overall poor literacy acquisition.[3]

Using data from the Early Childhood Longitudinal Survey (ECLS-K), a recent report from the ETS Policy Information Center examined differences in the reading readiness of kindergartners grouped by gender, race/ethnicity, socioeconomic status (SES), and age. The study revealed statistically significant differences in reading readiness among different subgroups of kindergartners. Socioeconomic status was strongly related to reading proficiency, as children in higher SES groups were more likely to be proficient than children in lower SES groups. The study also reported that Asian and White children were more likely than children in other racial/ethnic groups to be proficient across all reading tasks measured. However, nearly all racial/ethnic differences in reading disappeared when children were grouped into similar SES levels.[4]

## Concern About the Effectiveness of Public Education

Identifying such discrepancies in early reading has fueled concern that our public education system may not be effective in teaching all children to read and write, especially children from lower socioeconomic environments. Reactions to the achievement gap can be seen at both the state and federal

[2] National Research Council and Institute of Medicine, *From Neurons to Neighborhoods: The Science of Early Childhood Development*, Washington, DC: National Academy Press, 2000.

[3] The Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), conducted by the National Center for Education Statistics, is following approximately 20,000 children in 1,000 public and private schools from kindergarten through fifth grade. For early findings and a description of the survey, see Jerry West, Kristin Denton, and Elvira Geronimo-Hausken, *America's Kindergartners*, National Center for Education Statistics, 2000.

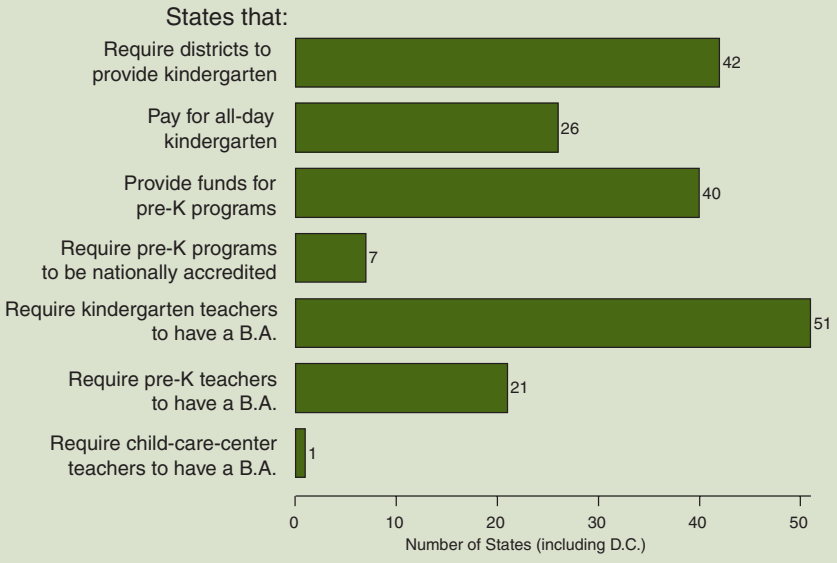
[4] Richard J. Coley, *An Uneven Start: Indicators of Inequality in School Readiness*, Policy Information Report, Policy Information Center, Educational Testing Service, March 2002.

levels. One example of the state response is the increasing number of states investing in non-compulsory education of 3- and 4-year-olds. As shown in Figure 1, according to *Education Week*, 39 states and the District of Columbia now provide state-financed pre-kindergarten for at least some of their 3- to 5-year-olds, up from 10 states in 1980. Annual state spending for such programs now exceeds $1.9 billion, and there are growing pressures to gather data to show that the programs are effective.

Despite the increase in state support for preschool education, it should be noted that teachers in these programs vary widely in training and experience, and inevitably the quality of instructional programs often varies. *Quality Counts* (2002) reported that only seven states require pre-K programs to be accredited, and while all states require their kindergarten teachers to have a B.A., only 21 require a B.A. for child care center teachers.[5]

One of the most ambitious state-level attempts to close the achievement gap between advantaged and disadvantaged young children is taking place in New Jersey. In 1988 the class action known as *Abbott v. Burke* resulted in the New Jersey Supreme Court ordering the State Department of Education to provide universal preschool to 3- and

FIGURE 1: STATE EARLY-CHILDHOOD POLICIES




Source: Education Week, *Quality Counts 2002.*

[5] Education Week, *Quality Counts 2002, Building Blocks for Success*, January 10, 2002.

4-year-olds in 30 of the state's most underachieving urban districts. The court ruled that, "Intensive preschool and full-day kindergarten enrichment programs are necessary to reverse the educational disadvantages these children start out with" (*Abbott v. Burke*, 1998). The New Jersey Supreme Court ordered the state's Department of Education to provide the following pre-K programming:

- ***Universal Eligibility*** - all 3- and 4-year-old children, with enrollment on demand;
- ***District-led Collaboration*** - preschool contracts with community and Head Start programs able and willing to meet the Abbott quality standards;
- ***Qualified Teachers and Small Classes*** - no more than 15 children per class, staffed by a state-certified (P-3) teacher and an assistant;
- ***Adequate Facilities and Funding*** - state-provided facilities and funding, adequate to meet district needs;
- ***Preschool Curriculum*** - developmentally appropriate curriculum, aligned with the New Jersey Core Curriculum Content Standards and elementary school reforms;
- ***Related Services*** - social and health services, transportation, and services for children with disabilities and with limited English proficiency, as needed; and
- ***District Support and Accountability*** - supervision, technical assistance, and professional development and evaluation to assure uniform high quality.[6]

Improvements to facilities, caps on class size, and new teacher certification requirements are part of the effort that is estimated to cost New Jersey $355 million in the 2002 fiscal year.

At the federal level, the No Child Left Behind Act of 2001 authorized two new reading programs: Reading First, funded at $900 million in the 2002 fiscal year, and Early Reading First, funded at $75 million. These programs are intended to enhance the language and literacy skills of all children and to eliminate the achievement gaps among racial/ethnic and socioeconomic groups. Funds may be used to select and administer screening, diagnostic, and classroom-based instructional reading assessments; purchase instructional materials; provide teacher professional development; and conduct program evaluation.

[6] Education Law Center, http://www.edlawcenter.org/ELCPublic/AbbottPreschool/AbbottPreschoolProgram.htm

# Issues in Literacy Instruction

*Educators have differed for some time on which instructional strategies are most effective in teaching children to read.*

Recently, groups of experts have come together to explore the issue of what constitutes appropriate and effective literacy pedagogy for young children. A 1998 report from the National Research Council sought to promote a balanced approach to reading instruction.[7] Its broad conclusions are included in Figure 2. Charged with "conducting a study of the effectiveness of interventions for young children who are at risk of having problems learning to read," the Council set three project goals:

(1) to comprehend a rich but diverse research base;
(2) to translate the research findings into advice and guidance for parents, educators, publishers, and others involved in the care and instruction of the young; and
(3) to convey this advice to the targeted audiences through a variety of publications, conferences, and other outreach activities.

In 2000, the National Reading Panel published a review of early literacy studies that they defined as "scientifically based" research.[8] That is, the group focused "exclusively on research that had been published or had been scheduled for publication in refereed (peer reviewed) journals... [N]on-peer-reviewed data were treated as preliminary/pilot data that might illuminate potential trends and areas for future research." The panel concluded that five factors constituted the most important components of reading instruction:

- ***Phonemic awareness*** - Phonemic awareness refers to the ability to focus on and manipulate phonemes in spoken words.

- ***Phonics*** - Systematic phonics instruction is a way of teaching reading that stresses the acquisition of letter-sound correspondences and their use to read and spell words.[9]

- ***Fluency*** - Fluent readers can read text with speed, accuracy, and proper expression.

[7] Catherine E. Snow, M. Susan Burns, and Peg Griffin (Eds.), *Preventing Reading Difficulties in Young Children*, National Research Council, Washington, DC: National Academy Press, 1998.

[8] National Reading Panel, *Report of the National Reading Panel: Teaching Children to Read, An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Impact for Reading Instruction*, Washington, DC: NIFL, NICHD, 2000.

[9] T.L. Harris and R.E. Hodges (Eds.), *The Literacy Dictionary: The Vocabulary of Reading and Writing*, Newark, DE: International Reading Association, 1995.

FIGURE 2: CONCLUSIONS FROM THE NATIONAL RESEARCH COUNCIL REGARDING APPROPRIATE READING INSTRUCTION

**Adequate initial reading instruction requires that children:**

- Use reading to obtain meaning from print,
- Have frequent and intensive opportunities to read,
- Be exposed to frequent, regular opportunities to read,
- Learn about the nature of the alphabetic writing system, and
- Understand the structure of spoken words.

**Adequate progress in learning to read English (or any alphabetic language) beyond the initial level depends on:**

- Having a working understanding of how sounds are represented alphabetically,
- Sufficient practice in reading to achieve fluency with different kinds of texts,
- Sufficient background knowledge and vocabulary to render written text meaningful and interesting,
- Control over procedures for monitoring comprehension and repairing misunderstandings, and
- Continued interest and motivation to read for a variety of purposes.

Source: National Research Council, *Eager to Learn: Educating Our Preschoolers*, Washington, DC: National Academy Press, 2001.

Fluency depends upon well-developed word recognition skills, but such skills do not inevitably lead to fluency.

- ***Vocabulary*** - The importance of vocabulary in reading comprehension has been recognized for more than half a century. In 1925, Whipple stated, "Growth in reading power means, therefore, continuous enriching and enlarging of the reading vocabulary and increasing clarity of discrimination in appreciation of word values."[10]

- ***Text comprehension*** - Comprehension has come to be viewed as "the essence of reading.... Comprehension strategies are specific procedures that guide students to become aware of how they are comprehending as they attempt to read and write."[11]

[10] G. Whipple (Ed.), *The Twenty-fourth Yearbook of the National Society for the Study of Education: Report of the National Committee on Reading*, Bloomington, IL: Public School Publishing Company, 1925.

[11] D. Durkin, *Teaching Them to Read* (6th ed.), Boston, MA: Allyn and Bacon, 1993.

Although these five components do not constitute an exhaustive list of reading skills, they do represent important elements of reading development. Sound reading programs will balance decoding, comprehension, and high-interest materials. While debate continues on how to implement early literacy instruction, the groups do not hold entirely antithetical notions of what represents the basics of becoming literate—cracking a code, understanding the text, and having frequent opportunities to read interesting material.

Regardless of the instructional approach to literacy development, educators need effective strategies to inform classroom practice and to monitor children's progress. Assessment needs and demands go beyond this, however. The increased focus on early literacy initiatives, the new thrust to provide enriched learning environments for young children, and the persistent achievement gaps among specific subgroups of students have raised the inevitable call for accountability. Meeting these diverse needs for assessment information is, of course, a complex undertaking. What is required is not merely a new set of standardized instruments, but a coordinated system for monitoring children's literacy development—one that incorporates appropriate learning goals, multiple measures of learning, administrative leadership, and ongoing professional development for teachers.

The following section outlines major issues in the design of an effective early literacy assessment system. The components described have emerged from the study of the design and implementation of early literacy assessment systems that inform instructional practice and meet accountability needs.[12]

[12] See Brent Bridgeman et al., *Characteristics of a Portfolio Scale for Rating Early Literacy*, Princeton, NJ: Educational Testing Service, 1995; Jacqueline Jones and Edward Chittenden, *Teachers' Perceptions of Rating an Early Literacy Portfolio*, Princeton, NJ: Educational Testing Service, 1995; Ruth Mitchell, *Testing for Learning: How New Approaches to Evaluation Can Improve American Schools*, New York: The Free Press, 1992; Joe Murphy, "Leadership for Literacy: Policy Leverage Points," paper presented at the Educational Testing Service/Education Commission of the States Conference on Leadership for Literacy, Washington, DC, October 9, 2001.

# Issues in the Assessment of Young Children

*The question of how to assess young children appropriately has challenged educators for some time.*[13]

In 1987 the nation's largest childhood organization, the National Association for the Education of Young Children (NAEYC), issued a position statement, shown in Figure 3, that calls for careful consideration of the purpose, selection, and use of standardized tests with young children. Early childhood experts in NAEYC and other professional groups, such as the National Association for Early Childhood Specialists in State Departments of Education (NAECS/SDE), have continued to consider what constitutes sound assessment of children below grade 3.[14]

The process of ongoing assessment should be distinguished from the administration of standardized norm-referenced tests. In their joint set of standards for educational and psychological testing, the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) have defined assessment as "Any systematic method of obtaining information from tests and other sources, used to draw inferences about characteristics of people, objects, or programs."[15]

Lorrie Shepard and her colleagues have proposed a set of principles and purposes to guide the overall assessment of young children (see Figure 4). These researchers argue that, particularly below 3rd grade, the primary purpose of assessment is to provide the teacher with information that can be used to guide and improve instruction. Therefore, assessment of young children below 3rd grade should be centered on classroom-based

[13] See Edward Chittenden and Rosalea Courtney, "Assessment of Young Children's Reading: Documentation as an Alternative to Testing," in *Emerging Literacy: Young Children Learn to Read and Write*, Dorothy S. Strickland and Leslie M. Morrow (Eds.), Newark, DE: International Reading Association, Inc, 1989; Henry S. Dyer, "Testing Little Children: Some Old Problems in New Settings," *Childhood Education*, 49 (7), 362-367, 1973; Celia Genishi (Ed.), *Ways of Assessing Children and Curriculum*, New York: Teachers College Press, 1992; Lorrie A. Shepard, "The Challenges of Assessing Young Children Appropriately," *Phi Delta Kappan, 76 (3),* 206-212, 1994.

[14] National Association for the Education of Young Children (NAEYC) and National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE), *Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 Through 8*, 1992. www.naeyc.org. (Revised statement due out in Summer, 2003.)

[15] American Educational Research Association, American Psychological Association, and the National Council for Measurement in Education, *Standards for Educational and Psychological Testing*, Washington, DC: American Educational Research Association, 1999.

## Figure 3: NAEYC Position Statement on Standardized Testing of Young Children 3 Through 8 Years of Age

NAEYC believes that the most important consideration in evaluating and using standardized tests is the utility criterion: The purpose of testing must be to improve services for children and ensure that children benefit from their educational experiences. Decisions about testing and assessment instruments must be based on the usefulness of the assessment procedure for improving services to children and improving outcomes for children. The ritual use even of "good tests" (those that are judged to be valid and reliable measures) is to be discouraged in the absence of documented research showing that children benefit from their use.

The following guidelines are intended to enhance the utility of standardized tests and guide early childhood professionals in making decisions about the appropriate use of testing.

- All standardized tests used in early childhood programs must be reliable and valid according to the technical standards of test development.*
- Decisions that have a major impact on children, such as enrollment, retention, or assignment to remedial or special classes, should be based on multiple sources of information and should never be based on a single test score.
- It is the professional responsibility of administrators and teachers to critically evaluate, carefully select, and use standardized tests only for the purpose for which they are intended and for which data exists demonstrating the test's validity (the degree to which the test accurately measures what it purports to measure).
- It is the professional responsibility of administrators and teachers to be knowledgeable about testing and to interpret test results accurately and cautiously to parents, school personnel, and the media.
- Selection of standardized tests to assess achievement and/or evaluate how well a program is meeting its goals should be based on how well a given test matches the locally determined theory, philosophy, and objectives of the specific program.
- Testing of young children must be conducted by individuals who are knowledgeable about and sensitive to the developmental needs of young children and who are qualified to administer tests.
- Testing of young children must recognize and be sensitive to individual diversity.

*American Educational Research Association, American Psychological Association, National Council on Measurement in Education, *Standards for Educational and Psychological Testing*, Washington, DC: American Education Research Association, 1985.



Source: National Association for the Education of Young Children, *Standardized Testing of Young Children 3 Through 8 Years of Age*, Washington, DC, 1987.

## Figure 4: General Principles and Purposes of Early Childhood Assessment

Assessment should bring about benefits for children.

Assessments should be tailored to a specific purpose and should be reliable, valid, and fair for that purpose.

Assessment policies should be designed recognizing that reliability and validity of assessments increase with children's age.

Assessments should be age-appropriate in both content and the method of data collection.

Assessments should be linguistically appropriate, recognizing that to some extent all assessments are measures of language.

Parents should be a valued source of assessment information, as well as an audience for assessment results.

**Major Purposes of Early Childhood Assessment**

- Assessments to support learning,
- Assessments for identification of special needs,
- Assessments for program evaluation and monitoring trends, and
- Assessments for high-stakes accountability.

Source: Lorrie Shepard, Sharon L. Kagan, and Emily Wirtz, *Principles and Recommendations for Early Childhood Assessments*, Washington, DC: National Education Goals Panel, 1998.

evidence of learning rather than externally designed norm-referenced, standardized measures.

Finally, early literacy assessment was addressed directly in the Joint Position Statement of the International Reading Association (IRA) and NAEYC. The two organizations outlined a set of principles for appropriate literacy assessment of young children, and these are shown in Figure 5.

In general, these positions highlight appropriate early childhood assessment as a means of understanding how young children are developing competence and as a tool to inform instructional practice. The following sections will outline some of the elements of an effective early literacy assessment system.

To provide the appropriate educational experiences that can promote literacy development in all young children, educators will

## Figure 5: Joint NAEYC/IRA Position Statement

**Principles of Assessment in Reading and Writing**

Assessment should support children's development and literacy learning.
Assessment should take many different forms.
Assessment must avoid cultural bias.
Assessment should encourage children to observe and reflect on their own learning progress.
Assessment should shed light on what children are able to do as well as the areas where they need further work.

**Assessment Procedures**

Anecdotal notes
Narratives, story retelling
Writing folders
Instructional conversations
Emergent storybook readings
Informal reading inventories
Running records

Source: Susan B. Neuman, Carol Copple, and Susan Bredekamp, *Learning to Read and Write: Developmentally Appropriate Practices for Young Children*, Washington, DC: National Association for the Education of Young Children, 2000.

need to design appropriate learning goals, implement sound instructional strategies, institute comprehensive monitoring systems that reflect and inform classroom practice, and provide for a well-trained and adequately supported teaching force. To this end, those responsible for the education of young children must:

- Acquire the skills necessary to select, use, and interpret multiple measures of literacy development that are aligned to clear and appropriate standards;
- Recognize the critical role that sound leadership plays in a useful assessment system; and
- Provide the teaching staff with a coherent program of professional development to help them develop knowledge and skills that will enable them to successfully build children's literacy skills, monitor children's progress, and apply assessment information to their instructional practice.

# Elements of Effective Early Literacy Assessment Systems

*Perhaps the single most unifying concept in the arena of early literacy instruction and assessment is recognition of the need for multiple assessment measures.*

## Multiple Forms of Evidence of Early Literacy Development

The early childhood community, the national reading organizations, and the measurement community have issued policy statements opposing the use of a single assessment instrument to determine literacy progress for high-stakes purposes.[16] In testimony before the House Subcommittee on Education Reform, Reid Lyon, Chief of the Child Development and Behavior Branch, National Institutes of Health, spoke of the role of assessments in the No Child Left Behind legislation: "The President's proposed reading programs recognize both the importance of assessment and the fact that assessments have multiple purposes, including early identification, diagnosis, program evaluation, and accountability. A single test cannot address all these purposes."[17]

Further, as Catherine Snow and Jacqueline Jones have stated, "... tests, by themselves, cannot improve educational outcomes. They can lead to improvement only if they become a stimulus to change in the educational system—a basis for improved curricula, upgraded instruction, better professional development for teachers, and better distribution of resources."[18] No single test instrument will be sufficient to promote good teaching, track children's progress, signal when children are demonstrating potential problems, and provide useful feedback to teachers. As illustrated in Figure 6, some instruments are designed to provide more information about individual students than others.

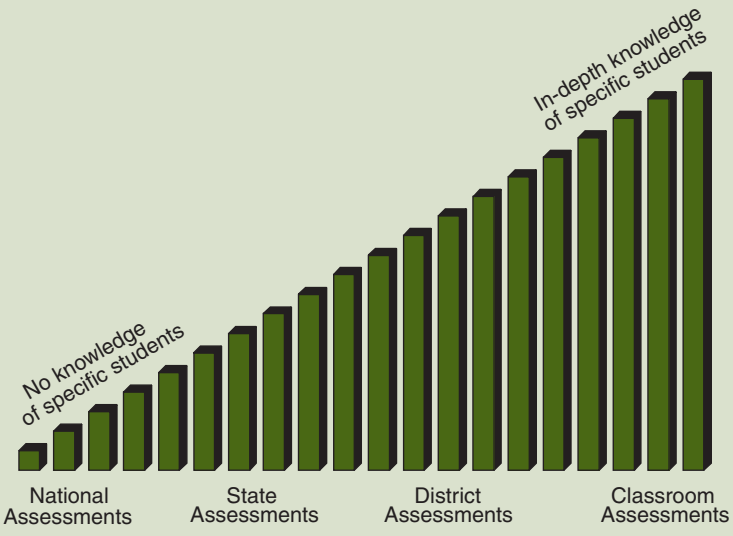
Classroom-based assessments that are directly linked to the instructional program can provide a rich source of information on a single child or a group of children. Conversely, large-scale national assessments, such as the National Assessment of Educational Progress (NAEP), are intended to provide a big picture of group progress. In the case of NAEP, individual children take blocks of the assessment; no child takes the entire test. Therefore, no generalizations can be made about a specific child. Sound early literacy assessment systems should include

[16] International Reading Association, *High-stakes Assessment in Reading*, Newark, DE, 1999; American Educational Research Association, 1999; and National Association for the Education of Young Children, *Standardized Testing of Young Children 3 Through 8 Years of Age*, Washington, DC, 1987.

[17] G. Reid Lyon, *Measuring Success: Using Assessments and Accountability to Raise Student Achievement*, Subcommittee on Education Reform, Committee on Education and the Workforce, Washington, DC, 2001.

[18] Catherine E. Snow and Jacqueline Jones, "Making a Silk Purse: How a National System of Annual Testing Might Work," *Education Week*, April 25, 2001.

FIGURE 6: RELATIONSHIP BETWEEN LEVEL OF ASSESSMENT AND KNOWLEDGE OF SPECIFIC STUDENTS




Source: Robert J. Marzano and John S. Kenall, *A Comprehensive Guide to Designing Standards-Based Districts, Schools, and Classrooms*, Copyright 1996, McREL. Reprinted with permission of McREL.

the range of assessments, from informing and reflecting classroom practice to serving school district and state program evaluation needs. Screening measures, classroom-based assessments, and program evaluation measures constitute the major types of assessment instruments that can yield important information in an early literacy assessment system. Each is considered below.

- Screening measures are gross indicators of whether children are generally progressing on course or whether there are major problems, such as visual or auditory obstacles to learning. They serve as a first step in identifying children who may be at high risk for delayed development or academic failure, or who are in need of further evaluation for special services. While screening instruments are not intended to guide instructional interventions, they can serve as red flags to alert educators of potential difficulties.[19]

- Classroom-based literacy assessments can provide powerful evidence of children's literacy acquisition. As Lorrie Shepard and her colleagues have argued, the major purpose

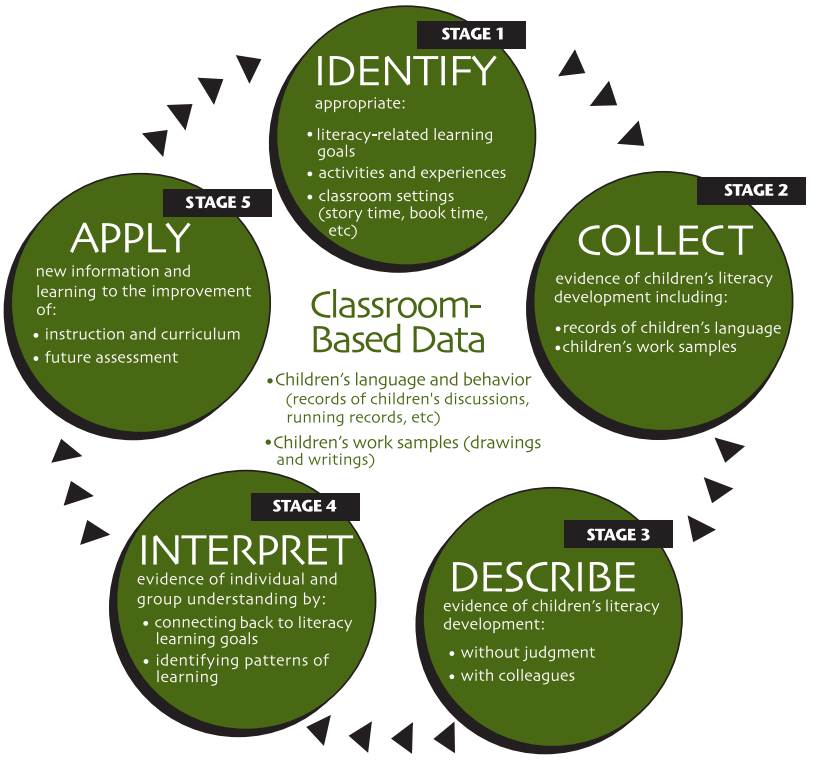
[19] John M. Love, *Instrumentation for State Readiness Assessment: Issues in Measuring Children's Early Development and Learning*, Princeton, NJ: Mathematica Policy Research, 2001.

of early childhood assessment is to inform instructional practice. Classroom-based assessments that are tied to classroom experiences in which children play, engage in conversations, and construct meaning can guide future instruction by providing valuable information about literacy development. This evidence should be continually measured against a set of clearly defined, challenging, and appropriate literacy goals. As illustrated in the 5-Stage Early Literacy Documentation-Assessment Cycle, developed by ETS Research and Development and shown in Figure 7, classroom-based literacy assessments are an integral part of the teaching/learning process.

In the documentation-assessment process, teachers are guided through a five-stage cycle.

Stage 1. ***Identifying*** learning goals and the evidence of learning
Stage 2. ***Collecting*** the evidence over time

FIGURE 7: 5-STAGE EARLY LITERACY DOCUMENTATION-ASSESSMENT CYCLE



Stage 3. ***Describing*** evidence in a non-evaluative manner
Stage 4. ***Interpreting*** evidence in order to identify mastery of established goals
Stage 5. ***Applying*** this new information to inform instruction to support learning.

In Stage 1 of the cycle, appropriate literacy goals and the classroom activities and experiences that will help children master the goals are identified. Stage 1 also includes identification of what will constitute evidence that children have achieved the goals.

The teacher collects the classroom-based evidence of early literacy development in Stage 2. This evidence may consist of records of children's conversations and samples of children's work, such as drawings, writings, and constructions.

In Stage 3 of the Documentation-Assessment Cycle, teachers work collaboratively to describe what they see in the records of children's language and in work samples. This stage is modeled on Patricia Carini's descriptive review process in which teachers engage in a close analysis of children's work.[20]

In Stage 4, the classroom-based evidence of early literacy development is weighed against the goals identified in Stage 1 of the cycle. It is in Stage 5 that the information gained in Stages 1-4 is used to plan future instruction and to plan additional assessment opportunities.

Teachers benefit most from the documentation-assessment process when they work collaboratively, sharing strategies for collecting evidence of children's literacy development and engaging in descriptions and interpretations of that evidence. The process can enhance teachers' recognition, observation, and understanding of the evidence of young children's emerging literacy learning. At the heart of identifying and collecting classroom-based data is the teachers' ability to become a keen observer and listener of children.

- Program evaluation and accountability are essential aspects of any system-wide work. Assessments such as NAEP, currently administered at grades 4, 8, and 12, can provide an overall look at how specific groups of children are progressing along toward a defined set of outcomes. However, these results are not intended to give specific information on the progress of individual children or to provide instructional strategies for teachers.

As educational systems move toward monitoring children's literacy development, the types of assessments described above must meet technical specifications if the results they yield are to be of value to teachers and

[20] Patricia Carini, *Observation and Description: An Alternative Methodology for the Investigation of Human Phenomena*, Grand Forks, ND: University of North Dakota Press, 1975; Patricia Carini, *The Descriptive Review of the Child: A Revision*, North Bennington, Vermont: Prospect Center, 1993.

children. Figure 8 shows a set of standards for accountability systems outlined by staff from the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

In general, effective early literacy assessment systems should include multiple types of assessments for multiple purposes, assure that all the measures are as valid and reliable as possible, and provide for clear and accurate reporting. Systems need to develop strategies to ensure that assessment instruments are administered and interpreted in comparable and consistent ways.[21] No system can accommodate all of these factors without a committed and informed administrative and teaching staff. The critical role of leadership in the success of a system-wide assessment system follows.

## The Role of Leadership

If the achievement gap is to be closed, system-wide solutions are needed. Factors beyond the assessment instruments can be critical determinants in the design and implementation of an effective assessment system. Research suggests that school districts that have been most successful in educating all children share some common attributes, and paramount among those commonalities is leadership that creates a shared vision of success and provides the staff with the tools to meet that vision. No matter how dedicated or well-trained, teachers cannot affect lasting, systemic change by themselves. Good teaching and good teachers need to be valued and supported by a rational system.

Leadership is not an individual. No single administrator can design, implement, and nurture an early literacy assessment system. Rather, leadership consists of the coordinated efforts among an administrative staff that shares a common vision. Joe Murphy has outlined the leadership components of successful literacy programs and these are provided below (in italics).[22]

- ***Establishing clearly defined, challenging, and public standards that are the focus of policymaking, institutional structures, and activities. Everything is centered on accomplishing the learning goals that have been defined.*** No assessment system can be effective and useful to teachers unless it is grounded in a clearly defined and public set of standards. So it is that early literacy assessment depends on a common definition of the literacy goals for young children. Whatever the emphasis on phonemic awareness, phonics, fluency, vocabulary, comprehension, or appreciation of text, the literacy goals must be clear to all educators and aligned with each

[21] Eva L. Baker, H.F. O'Neil, and Robert L. Linn, "Policy and Validity Prospects for Performance-Based Assessment," *American Psychologist*, 48 (12), 1993. Also see *Standards for Educational and Psychological Testing* (cited earlier).

[22] Murphy, 2001.

FIGURE 8: STANDARDS FOR SYSTEM COMPONENTS

**Accountability Systems**

- Accountability expectations should be made public and understandable for all participants in the system.
- Accountability systems should employ different types of data from multiple sources.
- Accountability systems should include data elements that allow for interpretations of student, institution, and administrative performance.
- Accountability systems should include the performance of all students, including subgroups that historically have been difficult to assess.
- The weighting of elements in the system, including different types of test content and different information sources, should be made explicit.
- Rules for determining adequate progress of schools and individuals should be developed to avoid erroneous judgments attributable to fluctuations of the student population or errors in measurement.
- Longitudinal studies should be planned, implemented, and reported evaluating effects of the accountability program. Minimally, questions should determine the degree to which the system:

  a. builds capacity of the staff;
  b. affects resource allocation;
  c. supports high-quality instruction;
  d. promotes student equity access to education;
  e. minimizes corruption;
  f. affects teacher quality, recruitment, and retention; and
  g. produces unanticipated outcomes.

- The validity of test-based inferences should be subject to ongoing evaluation. In particular, evaluation should address aggregate gains in performance over time and impact on identifiable student and personnel groups.

Source: Eva L. Baker, Robert L. Linn, Joan E. Herman, and Daniel Koretz, *Standards for Educational Accountability Systems*, Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), 2002.

assessment. A major responsibility of early childhood leaders is to guide their educators in defining a set of challenging and appropriate literacy goals for young children.

- ***Creating a shared belief that all children can learn and that the educators have the knowledge, skills, and resources to engage in effective teaching.*** No program can be successful without the deeply held belief that all children can and will become literate. While individual differences will certainly be demonstrated, and some children will be better readers or writers than others, teachers must feel they have the necessary skills and resources (human as well as material) to guide each child to successful literacy development. Accordingly, teachers need a wide range of literacy materials, as well as support from other staff within the school—including special education staff—in order to provide fully differentiated instruction that addresses students' diverse needs. As noted in a later section, professional development is also vital.

- ***Guiding the assessment system.*** An assessment system cannot survive without administrative guidance and support. In addition, a person or group within the school must have a deep knowledge of (a) how young children learn to read and write and (b) what constitutes appropriate literacy goals and instructional programs. A variety of assessments can then be aligned to the learning goals.

- ***Providing the early literacy expertise to ensure that teachers are implementing a high-quality instructional program that can help children meet the district's learning goals.*** Early childhood leaders must look carefully at the plethora of reading programs and decide which instructional program is best aligned with their literacy goals and has a proven record of success. Someone with expertise in early literacy development must be either on staff or available to the staff as these instructional decisions are made.

- ***Providing, honoring, and protecting the time that teachers and other staff need to carry out their work.*** There must be sufficient time to do the work. The most precious commodity in most schools is not money; it's time. Supportive administrators recognize that teachers need time to plan, teach, reflect, and collaborate with their colleagues. That time must be honored, provided, and protected.

- ***Creating a system of monitoring progress toward the specified goals and ensuring that this information is used and understood.*** Early childhood administrators must establish and support the means through which assessment information is communicated to educators and parents to show how children are progressing toward the learning goals. This assessment information may be conveyed through report cards, parent conferences, or other means.

- ***Providing for a coherent system of professional development activities*** that enable teachers to plan appropriate instructional programs, teach, gather the evidence of children's learning, reflect on that information with colleagues and individually, and apply what they have learned to their ongoing classroom practice.

- ***Maintaining a stable district administration*** that keeps the vision of the work, ensures consistent implementation, and offers a coherent system of policies and procedures. Frequent changes in administrators can create changing visions and shifting definitions of success. The focus on literacy may be lost, instructional programs can cease to be aligned to a set of standards, or the literacy goals may shift and become unclear. Long-term stability in administration helps to sustain the focus of early literacy programs over time.

- ***Creating an educational partnership between the school and the home.*** Parents of young children play a critical role, and the work of improving early childhood education cannot be accomplished without a strong educational partnership between parents and educators.

## Teacher Professional Development

Research tells us that teachers are at the heart of any successful reform initiative.[23] The preparation and ongoing development of well-trained teachers and administrators is therefore key to closing the achievement gap. If improving instruction is the primary purpose of early childhood assessment, teachers need to be able to design classroom assessments and use the resulting information to guide their practice.

In addition, administrators must be able to select screening and accountability instruments that are aligned with their instructional program and to accurately interpret and report the results. These are daunting tasks for most educators, however. Richard Stiggins has argued that educators need to develop a much greater degree of "assessment literacy" if they are to monitor student progress in a meaningful manner.[24] The lack of attention to the design and appropriate use of classroom-based assessments in early childhood teacher preparation programs has played a major role in creating a teaching force that holds negative views toward testing and lacks the expertise to develop or use the classroom-based strategies that are appropriate for assessing young children's progress.

[23] National Commission on Teaching and America's Future, *What Matters Most: Teaching for America's Future*, September 1996.

[24] Richard J. Stiggins, "Assessment Literacy," *Phi Delta Kappan*, 72 (7), 1991; Richard J. Stiggins, "Evaluating Classroom Assessment Training in Teacher Education Programs," *Educational Measurement: Issues and Practices*, 18 (1), 1999; and Richard J. Stiggins, "The Unfulfilled Promise of Classroom Assessment," *Educational Measurement: Issues and Practices*, 20 (3), 2001.

An essential element of an effective early literacy assessment system is a coherent and comprehensive set of professional development activities that are targeted to enhancing classroom teachers' and administrators' assessment-related knowledge and skills. Educators must be equipped to select the most suitable set of instruments, administer them appropriately, and interpret the results to other teachers, parents, school boards, and state departments of education. In addition, it is critical that teachers play a key role in the design and ongoing revision of their assessment system. As Scott Paris and colleagues suggest, "It may take several years for all teachers to understand assessment practices and use them in similar ways. Consensus is built upon regular reflection and discussion among teachers about what assessments are working well, how the assessments support parent conferences and report cards, and how assessments help individual children."[25]

[25] Scott G. Paris, Alison H. Paris, and Robert D. Carpenter, *Effective Practices for Assessing Young Readers* (CIERA Report #3-013), Ann Arbor, Michigan: Center for Improvement of Early Reading Achievement, 2001.

# An Effective Early Literacy Assessment System

*In considering these many prerequisites for an effective early literacy assessment system, it is helpful to examine an actual system—a place where the necessary components have been brought together successfully.*

The South Brunswick, New Jersey, school district designed and implemented an early literacy portfolio and scale that has been operational for over 10 years. The portfolio and scale were designed to monitor the literacy development of all kindergarten through 2nd grade children in the district. The assessment system was intended both to inform teachers' instruction and to serve as an accountability measure.

Components of the South Brunswick Early Literacy Profile are listed in Figure 9. As the figure shows, a variety of forms of information or "literacy evidence" is collected for each student. This information is then used by teachers, at least twice during every school year, to assess and document each child's literacy development using the district's Early Literacy Scale. This scale is organized around six major phases in the normal development of children's emerging abilities to make sense of print (see Figure 10).

The committee of teachers who created the initial instrument drew upon the existing body of research on early literacy development in defining each stage of emergent literacy. The constellation of behaviors at any given stage grows out of preceding stages. The scale is not a checklist of specific behaviors. Instead, it attempts to embody the assumptions about children's language, reading, and writing that underlie the district's notions of appropriate curriculum. The scale represents the district's public set of literacy goals for kindergarten through 2nd grade children, with each stage serving as a benchmark as children become "Advanced Independent" readers.

Each point on the early literacy scale is described in terms of the strategies and understandings that children typically bring to text at a given stage. The descriptors are explicitly referenced to types of behaviors that teachers might observe in the portfolio documents. Furthermore, there is a consistent attempt at each point to characterize children's literacy learning in terms of what they *can* do rather than what they *can't* do. For example, the child is described as "indicating a beginning sense of one-to-one correspondence" in attempts to read text, word for word.

At least twice a year, each kindergarten through 2nd grade child in the school district is rated using the early literacy scale. As stated in the district's guidelines, "The ratings are based solely on information in the student's portfolio. Placing students on the scale is intended to serve two purposes. First, the process requires that teachers review and summarize each student's progress at a designated point in time. Second, the ratings are used by the

Figure 9: Components of the South Brunswick Portfolio

- Self-portrait
- Student Interviews
- Parent Questionnaire
- Concept About Print Test
- Phonemic Awareness
- WAWA II
- Alphabet Test
- Word Analysis
- Word Lists
- Running Record Assessment
- Writing Sample
- Story Retelling
- Nonfiction Retelling
- Higher Order Comprehension

district, in place of first- and second-grade reading tests, to monitor and evaluate the success of our students and the early childhood program."[26]

It is important to emphasize that establishing a common set of procedural guidelines for collecting work samples would not be sufficient by itself to sustain the assessment system; mechanisms must be in place to foster teacher professional development and collaboration.

Indeed, a critical factor in the success of the South Brunswick, New Jersey, school district's early literacy portfolio and scale is the district's emphasis on intense teacher professional development. Leaders in the district recognized that successful systemic implementation of the Early Literacy Portfolio and its developmental scale required several layers of ongoing district support. Beyond the definition of a common set of assumptions or guidelines, the following institutional mechanisms were needed to foster and sustain professional conversation among teachers.

***Development Committee*** - The Early Literacy Portfolio Development Committee, consisting of teachers, was responsible for designing the original portfolio and the developmental scale and for reviewing and revising the assessment.

***Training Workshops*** - Throughout the school year, the district sponsors training workshops and "same-grade" meetings for all primary-level teachers, as well as workshops for all new staff.

***Portfolio Scale/Calibration Meetings*** - These meetings provide one of the most powerful professional development opportunities in the district. They are viewed by teachers as useful in building a consensus and a common language, refining their rating skills, helping to standardize the portfolio procedures, and broadening their understanding of the entire rating process.

[26] *South Brunswick, New Jersey, Early Literacy Portfolio Guidelines*, 1999 (Portfolio Manual Directions, p. 6).

## Figure 10: South Brunswick Early Literacy Scale: Scoring Form

**Child's Name** _____ **Grade K 1 2**

**Primary Reading/Writing Scale**
**Development of children's strategies for making sense of print**

Directions: To record score, insert date and circle appropriate level
Record child's name and grade
Place form in the student's portfolio

Date _____

**Level 1: EARLY EMERGENT**
Displays an awareness of some conventions of reading, such as front/back of books, distinctions between print and pictures. Sees the construction of meaning from text as "magical" or exterior to the print. While the child may be interested in the contents of books, there is as yet little apparent attention to turning written marks into language. Is beginning to notice environmental print.

**Level 2: ADVANCED EMERGENT**
Engages in pretend reading and writing. Uses reading-like ways that clearly approximate book language. Demonstrates a sense of the story being "read," using picture clues and recall of story line. May draw upon predictable language patterns in anticipating (and recalling) the story. Attempts to use letters in writing, sometimes in random or scribble fashion.

**Level 3: EARLY BEGINNING READER**
Attempts to "really read." Indicates beginning sense of one-to-one correspondence and concept of words. Predicts actively in new material, using syntax and story line. Small stable sight vocabulary is becoming established. Evidence of initial awareness of beginning and ending sounds, especially in invented spelling.

**Level 4: ADVANCED BEGINNING READER**
Starts to draw on major cue systems; self-corrects or identifies words through use of letter-sound patterns, sense of story, or syntax. Reading may be laborious especially with new material, requiring considerable effort and some support. Writing and spelling reveal awareness of letter patterns. Conventions of writing such as capitalization and full stops are beginning, but are not consistent.

**Level 5: EARLY INDEPENDENT READER**
Handles familiar material on own, but still needs some support with unfamiliar material. Figures out words and self-corrects by drawing on a combination of letter-sound relationships, word structure, story line and syntax. Strategies of re-reading or of guessing from larger chunks of texts are becoming well established. Has a large stable sight vocabulary. Conventions of writing are understood.

**Level 6: ADVANCED INDEPENDENT READER**
Reads independently, using multiple strategies flexibly. Monitors and self-corrects for meaning. Can read and understand material when the content is appropriate. Writes with some fluency. Ideas remain on the topic and are conveyed in a logical sequence. Conventions of writing and spelling are —for the most part—under control.

**Rating scale developed by South Brunswick teachers and ETS staff - January 1991**

During the scale calibration meetings, teachers across the district read, discuss, and rate their colleagues' Early Literacy Portfolios. Working in pairs, teachers are presented with the portfolio of a child they do not know, other than by the evidence in the portfolio. The teachers use these portfolio documents to assign a rating on the portfolio's six-point scale. This rating is then compared with the rating assigned by the child's own classroom teacher. Agreement coefficients have consistently ranged from the mid .80's to the low .90's.[27] In the unlikely case of a disagreement of more than one scale point, a third party mediates the discussion until consensus is achieved. As a result of these meetings, the teacher whose portfolio has been rated might revise the score and review the documents in a specific portfolio. The double scoring procedures provide an opportunity to discuss documents, review scoring standards, and assure that consistent scoring standards are maintained across the district.[28]

[27] Bridgeman et al., 1995.

[28] Jones and Chittenden, 1995.

# Summary and Conclusions

*Monitoring the literacy development of young children and evaluating the effectiveness of programs cannot be accomplished by administering a single test during the academic year.*

Improving literacy development for all children and closing the achievement gap require a system-wide commitment to ongoing professional development in early literacy acquisition, instruction, and assessment. A useful assessment system requires:

- ***Multiple literacy assessments.*** The coherent use of an array of instruments consisting of screening, diagnostic, and classroom-based instruments to inform instructional practice and to evaluate program effectiveness is essential.

- ***Strong and stable leadership.*** While classroom teachers play a critical role in children's learning, the success of any sustained system-wide change requires solid administrative support. Someone must guide the process by providing sufficient time and resources to teachers, and someone must hold fast to the vision for the work.

- ***Coherent professional development programs.*** Good intentions alone are not sufficient. Effective assessment systems rely on well-trained and reflective teachers who understand cognitive and literacy development as well as the basic concepts of appropriate assessment.

These are difficult systems to design and implement, as evidenced by the paucity of model programs. The financial resources being allocated for literacy testing will only be useful to children and teachers if we recognize the critical role played by multiple forms of assessment in the teaching and learning process and begin to examine how to make system-wide changes that incorporate professional development for teachers and administrators.

88502-023990 • U63E8 • Printed in U.S.A.
I.N. 997906