DOCUMENT RESUME

ED 450 926 PS 029 268

AUTHOR Shepard, Lorrie A.; Taylor, Grace A.; Kagan, Sharon L.
TITLE Trends in Early Childhood Assessment Policies and Practices.
INSTITUTION Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 1996-10-00
NOTE 25p.
AVAILABLE FROM National Center for Research on Evaluation, CRESST/CSE, Graduate School of Education & Information Studies, University of California, Los Angeles, CA 90095-1522. Tel: 310-206-1532.
PUB TYPE Information Analyses (070)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Agency Cooperation; *Early Childhood Education; *Educational Change; *Educational Policy; *Educational Practices; Elementary School Students; Evaluation Methods; Grade Repetition; National Surveys; Preschool Children; School Readiness; Screening Tests; Standardized Tests; State Surveys; *Student Evaluation; Trend Analysis

ABSTRACT

This study documented current early childhood assessment policies and practices, focusing on changes since the 1988 nationwide survey and on changes occurring in response to Goal 1 of the National Education Goals, performance assessments, and integrated services for young children. Data were collected through telephone surveys and in-depth interviews with the state-level early childhood or elementary coordinator or testing director, as well as Part H and IDEA coordinators, Title I coordinators, Head Start directors, and the Goals 1 contact person. Documentary evidence of policies or legislative mandates was also obtained. The findings, in descending order of prevalence, follow: (1) most states have made an effort to move away from readiness testing and kindergarten retention, and there is a perceived reduction though not elimination of these practices; (2) almost all state-mandated standardized testing for school accountability has been eliminated for children below grade 3; (3) some states and local districts are moving to new assessment forms in the early grades that are more supportive of instruction; (4) misuse of screening instruments for instructional purposes has decreased since 1988; (5) professional training is needed to understand and use new assessment forms; (6) preschool testing is largely driven by mandates for categorical programs; (7) a few states are collecting data to report on progress toward Goal 1; (8) parent involvement was an issue both for identifying children with special needs and because of parent demands for standardized testing; and (9) lack of collaboration and coordination among agencies serving young children continues to be a problem. (Contains 16 references.) (KB)

Trends in Early Childhood Assessment Policies and Practices

Lorrie A. Shepard
Grace A. Taylor
University of Colorado, Boulder

Sharon L. Kagan
Yale Bush Center in Child Development and Social Policy

October, 1996

Running Head: Early Childhood Assessment Policies and Practices

This research was supported in part by the Office of Educational Research and Improvement (OERI), U. S. Department of Education, through the Center for Research on Evaluation, Standards, and Student Testing (CRESST). The findings and opinions expressed do not reflect the position or policies of OERI or the U. S. Department of Education.

Corresponding Author: Lorrie A. Shepard
Campus Box 249
School of Education
University of Colorado
Boulder, CO 80309-246
(303) 492-8108
Lorrie.Shepard@Colorado.edu

The present study was undertaken to document current early childhood assessment and testing practices in the states. For what purposes are young children assessed and how are they assessed? A decade ago, the most salient topics in early childhood education were inappropriate curriculum and its close corollary, inappropriate testing. Extensive efforts have been made by the National Association for the Education of Young Children (NAEYC), the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE), and other professional associations to offer guidance for more appropriate testing practices. How have assessment policies and practices changed in the ensuing years, if at all? How are assessment practices affected by other changes in the policy context such as efforts to measure progress toward Goal 1, standards-based reform, an emphasis on new forms of assessment, and the move toward integrated services for young children?

BACKGROUND

1988 Survey of States

In 1988, under the auspices of the National Academy of Sciences and the National Association of State Boards of Education, Gnezda and Bolig conducted a survey of prekindergarten and kindergarten testing in the 50 states. Their methodology was similar to that of the present study, relying on telephone interviews with early childhood specialists or testing and evaluation specialists in state departments of education. Because most testing practices were determined by local rather than state-level policies, their data necessarily reflected the perceptions of the people interviewed rather than a precise quantification of the prevalence of different practices. Nevertheless, findings from the Gnezda and Bolig survey provide a useful summary of testing practices in the late 1980s and a basis of comparison for interpretation of present-day reported practices.

Gnezda and Bolig encountered considerable confusion regarding testing purposes and terminology. Respondents were especially likely to confuse *screening* (a preliminary step in the identification of children as handicapped) and *readiness testing* (used both to plan instruction in a regular kindergarten or to make decisions to delay school entry and recommend retention in kindergarten). Screening upon entrance to kindergarten was mandated by state law in 19 states, with the most frequently used instruments being the Brigance, Battelle, Denver, DIAL-R, Early Prevention of School Failure, and the Gesell. Although screening requirements were originally intended for use in individualizing classroom instruction or for referral for additional assessment and special services, test results were reportedly also used to delay school entry or to place children in special programs such as developmental kindergarten. Seven states mandated readiness testing for kindergarten, with local use of readiness testing reported in 30 states, using such instruments as the DIAL-R, the Gesell, the Metropolitan, and the Peabody Picture Vocabulary Test. Six states mandated first-grade readiness testing with local use reported in an additional 37 states; measures included standardized achievement tests such as the CAT, CTBS, and ITBS. In 35 of these states, end of kindergarten test results were used to place children in transitional first grade or for kindergarten retention.

Most of the 22 state mandates for testing prior to kindergarten or first grade had been enacted in the mid-to-late 1980s and were attributed to "educational reform efforts, concern over the accountability of schools and of teachers, greater interest in identifying and serving children 'at risk' of school failure, and a downward extension of academics" (Gnezda & Bolig, 1988, p. 4). Local reasons for testing young children included a desire to "match children to the curriculum, increase the probability of success in kindergarten and first grade, demonstrate teacher success,

ensure placement of children into appropriate programs, identify learning problems, and individualize the curriculum" (p. 4). Some respondents also said that kindergarten teachers, "often encourage kindergarten readiness testing in order to 'protect' younger and immature children from skills-oriented kindergarten curricula" (p. 4).

Changing Policy Context

Many early childhood specialists were well aware that the testing practices captured by the 1988 survey did not meet professional standards for either sound measurement or educational decisions. In many cases IQ-like tests were being used inappropriately, and tests without adequate reliability and validity were being used to make major, career-altering placement decisions for children. Although these practices had arisen in response to a narrow skill-drill curriculum, holding out young and less ready children was exacerbating the problem of curriculum distortion rather than solving it. In the paragraphs that follow we summarize key policy efforts made by professional groups in the early childhood community in an effort to redress these negative practices. We also provide a brief overview of changes occurring in the larger context of educational reform.

Fostering developmentally appropriate practices in the early grades. In 1986, six professional organizations (Association for Childhood Education International, Association for Supervision and Curriculum Development, International Reading Association, National Association for the Education of Young Children, National Association of Elementary School Principals, National Council of Teachers of English) issued a joint statement expressing concern that pre-first grade children were being "subjected to rigid, formal pre-reading programs with inappropriate expectations and experiences for their levels of development" (International Reading Association, p. 110). Especially they noted that too much attention was being focused on isolated skill development and that "pressure to achieve high scores on standardized tests that frequently are not appropriate for the kindergarten child" (p. 111) was having a deleterious effect on the content of kindergarten programs. In 1987, responding further to these concerns, the NAEYC used the phrase "developmentally appropriate" (Bredekamp, 1987) -- now a commonplace slogan -- to represent the idea that curriculum and instructional practices should be consistent with the ways that children of a given age develop and learn and should be responsive to their individual and cultural needs and experiences.

Policy statements about testing of young children. Inappropriate curriculum and instructional practices were closely tied to inappropriate testing practices. Therefore, it was impossible to address one without addressing the other. In November of 1987, the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE) issued a statement against "Unacceptable Trends in Kindergarten Entry and Placement," and that same month, the NAEYC adopted its position statement on standardized testing. NAECS/SDE argued against denying school entrance to age-eligible children or segregating children into extra-year classes because such practices denied opportunities for cognitive growth through social interaction to children who most needed to be in school, labeled children as failures, and assigned the burden of responsibility (for readiness) to the child, rather than the school program. NAECS/SDE also addressed the problems of validly testing young children and recommended that tests be used for "initial program planning and information-sharing with parents" but not to "create barriers to school entry or to sort children" (p. 7). The NAEYC statement offered guidelines addressing the need for (1) reliable and valid measures, (2) use of multiple indicators when making important decisions, (3) the use of tests for their intended purpose, (4) professional knowledge and caution when interpreting test scores, (5) appropriate match of tests to local curricula, (6) knowledge of development needs, and (7) sensitivity to diversity. Given the

potential for measurement error and harm from testing, the stance of the NAEYC was that testing of young children should not occur unless the information gained was demonstrably of benefit.

Broader policy concerns about testing. Increased testing and misuse of testing were well documented phenomena in the 1980s and extended well beyond the confines of the early childhood years. A comprehensive report of these issues, Testing in American Schools(1992), was made to Congress by the Office of Technology Assessment in preparation for the reauthorization of the Elementary and Secondary Education Act (ESEA). In response to the minimum-competency movement and demands for accountability following A Nation at Risk in 1983, the amount of standardized testing increased greatly as did the pressure to improve test scores. Eventually evidence accumulated that these accountability pressures could have deleterious effects. For example, test scores could become inflated (i.e., student performance would drop dramatically when a new test was introduced) and curriculum could be distorted. Elementary teachers reported deemphasis of "non-basic" subjects such as social studies and science, and many teachers revised their instructional practices to closely resemble drill on standardized test formats. Studies also documented the relationship between accountability pressures in higher elementary grades and developmentally inappropriate practices in kindergarten and first grade (Hatch & Freeman, 1988; Smith & Shepard, 1988).

Standards-based reform and performance assessments. In contrast, "the second wave of reform" in the 1990s emphasizes problem solving and higher-order thinking skills and self-consciously rejects the focus on low-level skills of earlier reforms. A National Academy of Education report, Improving Education through Standards-Based Reform (McLaughlin & Shepard, 1995), summarizes both the vision of standards-based reform and criticisms of the standards movement. According to the vision, high expectations and challenging curricula will help all students reach high standards. Standards-based reforms "focus more on depth of understanding -- how well students can reason with and use what they have learned -- rather than on regurgitation of isolated facts" (McLaughlin & Shepard, p. 9) and call for teaching methods that engage students with meaningful content, involve hands-on experiments and projects, and require students to apply knowledge and skills to real-world problems. The development and use of more open-ended, authentic assessments that better represent important learning goals is an integral part of standards-based reform.

Changes in Title I accountability. Along with special education mandates, Chapter 1 evaluation requirements have been a major impetus for administering standardized tests to young children. Consistent with the trends already identified, a Commission on Chapter 1 convened to give advice about desired changes in the Elementary and Secondary Education Act, focused on higher expectations and reform of testing practices as the most pressing needs for improvement of Chapter 1. A critical component in the Commission's eight-part proposed framework was to remove the requirement for "low-level, norm-referenced, fill-in-the-bubble tests currently used to assess progress in Chapter 1. In their place, schools should develop ongoing means of evaluating the progress of individual students toward the standards, and states should administer new, richer, performance-based systems that measure school progress in enabling students to reach the state standards" (Commission on Chapter 1, 1992, p. 8). The spirit of these reforms is still very much evident in the draft document providing guidance to states for implementation of the new ESEA Title I (3/10/96). "The explicit purpose of the new ESEA is to support the broad-based reforms occurring in the States and localities" and to ensure that the same challenging standards set for other students will be held as expectations for Title I students. By the year 2000 states are expected to have developed or adopted high-quality assessments in at least reading and mathematics, which are aligned to the state content standards. These measures must assess

"complex thinking skills and understanding of challenging content" and be administered once in each of the grade intervals 3 through 5, 6 through 9, and 10 through 12. Note that under these new provisions, individual students participating in Title I need not be assessed on an annual basis for purposes of program accountability, and no accountability testing is required below grade 3.

The National Education Goals and Goal 1. The National Education Goals were first announced at the 1989 Charlottesville Summit by President Bush and Governor Clinton, speaking for the nation's governors. Bush and the nation's governors identified as their "first goal" that by the year 2000 all children in America will start school ready to learn. Goal 1 was supported by three enabling objectives which stated: that all disadvantaged and disabled children would have access to high quality preschool programs, that parents would have access to training to help their children learn, and that children and pregnant mothers would receive adequate nutrition and health care. The response from the early childhood community to the establishment of Goal 1 was mixed. Some believed that the readiness goal would keep the needs of young children at the forefront of a political agenda and help garner resources. Others feared that it would work to perpetuate exclusionary practices that kept low performing children out of school. The National Education Goals Panel reports annually on data relevant to infant health risks, immunizations of 2-year olds, and preschool participation, but there are no direct measures available on the level of children's physical, social, and cognitive development when they enter school.

At the same time, the two achievement goals -- that American students will demonstrate competency in challenging subject matter in English, mathematics, science, history, and geography and that by the year 2000 U.S. students will be first in the world in science and mathematics achievement -- drew public and political attention to the standards movement. In March 1994, Congress passed the Goals 2000: Educate America Act which set the goals into law. Although the education goals initially enjoyed bipartisan support, the Republican revolution in the fall of 1994 raised serious questions about the now politicized goals and standards. Although development of curriculum standards and performance assessments continue to be the focus of educational reform efforts in many states, the policy context is also colored by a political backlash against such standards.

Early childhood reforms outside of education. Numerous efforts are taking place to improve children's health status. Healthy Start programs exist in many states, as do parenting education and support programs that have their roots in health prevention and service. In addition, there are numerous efforts in the early care and education field designed to foster links between services for parents and families and services for children. Efforts to link these new programs to child care, Head Start, and pre-kindergarten programs are being made, although -- as will be seen -- these efforts may not be as well integrated as might be hoped.

METHOD

Study purpose and conceptual framework

The purpose of the study was to report on current early childhood assessment policies and practices. We were interested in what if any changes might have occurred since 1988 in response to direct policy efforts to curtail inappropriate practices; but we were also interested in other changes in testing and assessment practices that might be occurring in response to Goal 1, performance assessments, and integrated services for young children.

The approach taken to development of the interview protocol was guided by a conceptual framework based on categories of assessment purposes. In measurement theory, test validity and, indeed, all practical aspects of test content and format, depend on test use. For example, when tests are used for instructional purposes and for school accountability, test content should be

closely tied to curricular goals. In contrast, when evaluating learning "potential" to determine special needs, assessment tasks are intentionally designed to be as curriculum free as possible to avoid the confounding effects of opportunity to learn. Similarly, the stringency of technical requirements varies depending upon whether test results will be used to make important, potentially career-altering placement decisions about young children or are used on an ongoing basis in classrooms to plan instruction. Validity requirements and the effects on participating children are also different if assessments are used only to evaluate programs rather than to report reliably on the performance of individual children. To capture these distinctions, our questions were focused not just on how young children were being assessed but for what purpose? Three categories of assessment purpose were identified (Shepard, 1994): screening and identification of children with special needs, instructional improvement, and accountability. The first category included screening, developmental assessment, and other forms of testing for at-risk identification. The second category referred to both classroom and formal assessments used for instructional purposes. The third category included state or district level assessments and indicator systems used for school accountability, program evaluation, or for monitoring progress toward Goal 1.

Sample

Between May 1995 and August 1996 states were contacted by telephone to determine their early childhood assessment policies and practices. Initial contacts were made with early childhood specialists using a membership directory of the National Association of Early Childhood Specialists in State Departments of Education, which included 37 states, or by asking for the early childhood or elementary coordinator. If there was no early childhood coordinator or (in 3 cases) if that individual declined to participate, we asked to speak to the state testing director. Early childhood specialists or test directors, in turn, provided names and numbers of Part H and IDEA coordinators, Title I coordinators and Head Start directors. Using a web-page directory for Goals 2000 provided by the U.S. Department of Education, we also interviewed the Goals contact person in 47 states. In-depth interviews were conducted with at least one respondent from all 50 states and the District of Columbia. Whenever possible, we interviewed several officials from a state, but in three states we interviewed only the early childhood specialist or testing director.

Procedures

The formal interview protocol identified three different assessment purposes, screening and at-risk identification, instructional improvement, and accountability, and asked respondents to describe assessment policies and practices affecting 4- to 6-year olds in each of these categories. If one assessment was intended to serve more than one purpose, specifics of each use were noted. In addition to detailed questions about assessment requirements, e.g., the age and grade of students affected, the specific instruments used, respondents were also asked to reflect more broadly on any significant changes in early childhood assessment in their state in the last 5 years and to identify important issues still to be addressed. We also requested documents including specific policy statements or legislative mandates or to illustrate the types of reports produced or the support materials provided to local districts.

Data were recorded by note taking and by the use of audio recording to supplement note taking. Audio tapes were not exhaustively transcribed. In addition to simple tallies (for example, of states with and without mandatory screening programs), the open-ended narrative data were read and coded for emergent themes. After rereading, a final set of themes was identified that included all the issues discussed by multiple respondents. In the last stage of analysis, examples were culled from the data in support of each theme along with counter examples. As authors and interviewers, we also identified two themes by inference. These last two themes or findings were

not reported directly by respondents. Instead we noted that a) multiple phone calls were required because of lack of knowledge about and coordination of assessment requirements across educational programs (the assessment portion of finding 9) and b) instruments designed for one purpose were sometimes used inappropriately for other purposes (finding 4). Interview data were again examined systematically for evidence of these patterns.

FINDINGS

Important trends and issues are presented here using the organizing themes developed in data analysis. For each theme or finding, prototypical responses are summarized in support of each theme along with specific examples; we also cite counter examples where they occurred or use frequency information to give some indication of the generalizability of the stated theme. Findings are listed in roughly descending order of prevalence. Finding 1 was supported at least in part by nearly every state; in contrast, Findings 8 and 9 are based on comments from only a subsample of states.

Finding 1. Most states have made an effort to move away from readiness testing and kindergarten retention, and there is a perceived reduction though not elimination of these practices.

Given that some of the state early childhood coordinators who responded to this survey were among the members of the NAECS/SDE that issued a statement in 1987 condemning inappropriate uses of readiness testing, it is not surprising that they reported on efforts to discourage such practices. "We have attempted to get the word out," was the most frequent type of response. In some cases this involved workshops or other types of professional development activities; in other cases there were explicit policy directives or mandates forbidding certain uses of tests.

In California, for example, the California State Department of Education School Readiness Task Force issued a report entitled Here They Come: Ready or Not (1988) which discouraged the use of standardized tests for placement or exclusion of children. A 1989 advisory from the California State Department of Education urged "schools to examine their retention policies of young children" (Program Advisory, Child Development Division, June 12, 1989, p. 12). Noting that "standardized testing is particularly inappropriate for young children because each child comes from a unique set of family experiences," the advisory recommended that "standardized, norm-referenced tests not be used in kindergarten" (p. 13). California Education Code revised in 1992 (Section 8972) states that "standardized assessment tests may be used for diagnostic purposes only, provided those tests have been demonstrated to be nonbiased, valid, and reliable. These tests shall not be used to deny admission." In New Mexico, the state superintendent sent a letter to local school districts discouraging the use of junior first grades. In Massachusetts, the State Department of Education sent information to local school districts discouraging the use of the Gesell School Readiness Test and other instruments that may miscategorize children. They also discouraged the use of extra-year programs. Additional examples of state level efforts intended to discourage readiness testing and extra-year placements are shown in Table 1.

Many respondents described state-level policies against the use of tests to exclude or track students, but also stated that some local districts continued to use tests in ways that were considered to be inappropriate. In Indiana, the state has guidelines that say "kindergarten is an entitlement by law" and "all children should be welcomed into a one-year developmentally appropriate kindergarten experience." Although the state discourages retention at the kindergarten level and the number of transition classes has decreased over the last 5 years, some local districts still continue these practices. In Connecticut, A Guide to Program Development for

Kindergarten put out by the State Board of Education encourages developmentally appropriate measures and practices consistent with the NAEYC position statements on kindergarten placement, retention, screening, and assessment issues, but some local districts continue to use assessment for retention and transition placement.

As part of this trend, some states have eliminated or modified legislation that was the impetus for readiness testing and retention decisions. For example, Rhode Island had a law requiring that all children entering kindergarten be screened to determine "level of educational disadvantage." The law was not implemented beginning in 1991 and was eventually eliminated in 1994. In Oklahoma, the state requirement for screening of kindergarten students was eliminated in 1994 because some local school districts were using the screen to determine that some children were not ready for kindergarten; now screening, involving vision and hearing checks, a parent questionnaire, and developmental information, is required before entrance to first grade. Georgia's highly visible kindergarten retention test, mandated by the state legislature in 1987, was first modified by changing the content from the California Achievement Test to the more hands-on Georgia Kindergarten Assessment Program (GKAP). Then the law was changed so that retention in kindergarten was no longer automatically determined on the basis of a test score; instead retention decisions are to be based on multiple sources of information.

Despite this general trend away from formal testing in kindergarten, a number of states still mandate "screening" either as a first step in identification of children with special needs or to plan instruction. Examples of formal developmental screening requirements are shown in Table 2. In most cases, these state-level mandates are intended to satisfy the federal IDEA requirements that every state have a plan whereby children with disabilities can be identified, located, and evaluated. However, we should note that IDEA can be satisfied without every-pupil formal testing. In many states, in fact, only those children about whom there is concern are referred by teachers or parents for more in-depth developmental assessment. Mandated formal testing is problematic because it often leads to misuse of test results. For this reason, as noted in Table 1, some states such as Oklahoma and Rhode Island have eliminated their kindergarten screening requirements. Other states, as noted in Table 2, continue to require developmental screening but have attempted to clarify appropriate use of the results. For example, Louisiana state regulations require that "each local school district shall administer a nationally recognized screening instrument to every child entering kindergarten for the first time. The results of the screening shall be used in placement and planning instruction." The Louisiana State Board of Education clarified the regulation recently to note that the screening was not to be used to exclude children from programs. In South Carolina, all entering first graders are required to be tested within the first few weeks of school. The Cognitive Skills Assessment Battery has a cut off score that labels children either "ready" or "not ready," but the test does not determine eligibility. Every child that is 6 goes to first grade. The test is supposed to be used for curricular and instructional planning and as a starting point for further diagnosis. Children that score "not ready" may receive compensatory education. Although the CSAB is not to be used for program placement, in the past it has resulted in the use of transitional first grade classrooms. State level respondents expressed concern about this practice and are trying to prevent its occurrence.

Finding 2. Almost all state-mandated standardized testing for purposes of school accountability has been eliminated for children below grade 3; some local testing for accountability remains.

In asking respondents about assessments for the third purpose -- accountability -- a number of early childhood specialists commented that state-level testing had been eliminated for children below grade 3 in part because of the NAEYC policy statement on standardized testing

and the efforts of early childhood specialists and K-2 teachers in the state. Respondents described the grade levels and subject areas tested as part of their statewide assessment program. Grade levels tested are shown for the 50 states in Table 3. Interview data have been cross checked and updated based on the Profile of 1994-95 State Assessment Systems and Reported Results (NEGP, 1996).

Respondents also noted two other trends. First, local testing for accountability, though not necessarily required by the state, persists. Such local testing might include readiness testing in kindergarten or standardized tests such as the Metropolitan and CTBS administered in grades 1 and 2 and typically might affect half the districts in a state. We note that district administration of standardized tests in grades 1 and 2 has very likely been influenced by previous Chapter 1 requirements for individual pupil pre- and posttest data. These practices could change in the future; but at present most early childhood specialists whom we interviewed seemed unaware of changes in Title 1 guidelines. Test directors were better informed about proposed new Title 1 guidelines but were more wary of moving away from standardized tests. A second trend reported by a few early childhood specialists and test directors was increasing pressure to reinstate standardized testing below grade 3 due to legislative or gubernatorial press for accountability. "The pendulum is swinging back," they said.

Finding 3. Some states and local districts are moving to new forms of assessment in the early grades that are more supportive of instruction.

When asked to describe changes in assessment practices in their state, many respondents said that there was less readiness testing and grade retention in local school districts and more interest in "developmentally appropriate assessment." This included an increased use of portfolios and teacher observation and what some respondents referred to as "curriculum embedded assessment." A number of respondents said that districts were implementing the Work Sampling System (Meisels, 1992) or the High/Scope Child Observation Record (COR) (1992). For example, in the District of Columbia, Work Sampling is being phased in and is currently used in 68 of 112 elementary schools.

Most of these efforts to implement assessment in support of instruction are occurring at the district or school level. Only in a few cases have state-level assessment programs been developed to support instruction for children before grade 3. The Missouri State Department of Education, for example, is encouraging the development of a constructivist approach to assessment through the Project Construct Assessment System. It is a process-oriented curriculum and assessment framework for children ages 3-7. Similarly, North Carolina developed an assessment system designed to be used in lieu of the CAT in grades 1 and 2. Rather than a one-time test, data for the North Carolina Assessment are based on teacher observation over the course of the school year. The content of the assessments reflect emergent literacy and language development goals and, in mathematics, the Standards of the National Council of Teachers of Mathematics. A sample checklist for first-grade mathematics is shown in Figure 1. In addition to planning instruction, a primary purpose for these assessments is reporting to parents. Because the performance indicators are detailed and concrete they help parents to understand what curricular expectations are being addressed in a given grade as well as to see their own child's progress.

A few other states have legislative mandates that may lead to programs like those in Missouri and North Carolina. For example, in Massachusetts, the 1993 Education Reform Act includes a provision for portfolio assessment in grades pre-k through 12. In South Carolina, Proviso 19.73 of the 1994 Appropriations Act calls for pilot testing of a continuous assessment system for Kindergarten through Grade 3. In Washington, a new statewide assessment system has been mandated to begin in the 1998-99 school year. While third grade will be the grade

where state-level testing begins, it is hoped that local districts will develop a portfolio system for grades prior to grade 3. We note that these shifts are consistent with proposed new Title I guidelines, which would not require school accountability before grade 3, but which do presume that individual-level assessments would be used to guide instruction in grades K-2. The California Learning Record, developed by Barr and Cheong (1991) under the auspices of the California Department of Education is an example of a classroom-based assessment system specifically designed as an alternative model for Chapter I (now Title I) assessment.

Finding 4. Misuse of screening instruments for instructional purposes has apparently decreased since 1988 but continues to some extent.

In the 1988 Gnezda and Bolig survey, respondents frequently cited examples of special education screening measures being misused to plan instruction or to make decisions to delay school entry or recommend retention in kindergarten. Problems associated with denial of school entry and kindergarten retention have already been discussed. Use of screening test results to plan instruction is also problematic for several reasons. First, results of screening measures may be unreliable if used alone without the follow-up of a more in-depth developmental assessment. Second, classroom teachers may lack sufficient training to take proper account of children's background and language experiences and may make invalid grouping decisions based on test results. Third, the IQ-like test content of screening measures is not closely aligned to curriculum and therefore is a poor guide for diagnosis and instructional planning.

In the present study, respondents reported on the decrease of readiness testing (Finding 1) but tended not to comment directly on the misuse of screening measures except in those states where screening was expressly eliminated to prevent misuse. When asked to report what instruments were used for each assessment purpose, state-level respondents only occasionally reported that districts were using the Dial-R or the Peabody or the Battelle for instructional decisions. Therefore, the frequency of such practices appears to be greatly reduced. Indeed, examples cited in support of Finding 3 indicate that instead districts are beginning to develop their own instructionally relevant assessments or are implementing the Working Sampling System or High/Scope COR.

Finding 5. There is a need for professional training to understand and be able to use new forms of assessment.

More than half of the respondents commented on the need for better training of professionals to understand what developmentally appropriate assessment means and how to use new forms of assessment. For example, "portfolio assessment is 'overwhelming' for some teachers. They collect 'all this stuff,' and they have checklists, but they are not sure what to do with it." As explained by one respondent:

> "Teachers 'do not use that loop' connecting instruction with assessment. Teacher training at universities is not equipping preservice teachers for developmentally appropriate assessment. Once teachers are in the schools we (state-level early childhood specialists) give them training but they are not 'sophisticated enough at this point' to make the connection between developmentally appropriate practice and developmentally appropriate assessment. It's not that they're resistant, it's just that they're still on the beginning part of the learning curve."

Another respondent echoed the complaint that universities are not preparing preservice teachers to do appropriate assessment. Yet another agreed that, "training will be required to help teachers implement a new assessment program, to become confident users of the test results, and to be able to translate results into improved instructional strategies."

A respondent from yet another state explained that more is needed than just training in assessment:

> "Teachers are having a hard time with appropriate assessment because you don't get good numbers from performance assessment. People don't know what they are seeing when they observe children because they don't know enough about child development. The usefulness and validity of assessment in early childhood is limited because teachers don't know what it means."

Good assessment can occur by observing children during instructional activities and by recognizing misconceptions or patterns of errors in their written work, but this means that teachers must have knowledge of underlying developmental sequences to know how to interpret student performance and recognize its implications for instruction.

Hawaii is one example of a state that has made a major investment in staff development. After eliminating statewide testing at the kindergarten level, the state stepped up training in developmentally appropriate assessment for teachers. Teachers have been encouraged to use observation and to develop their own instruments. A cooperative arrangement has been made between the State Office of Instructional Service and the University of Hawaii to offer graduate credit courses that include developmentally appropriate assessment components. Another example of in-depth staff support was the joint effort by the Iowa Department of Education and Nebraska Department of Education to develop a primary grades curriculum based on the British Columbia Primary Program. Along with guiding principles and instructional activities, the project materials provide detailed guidance for educators regarding the purposes of assessment, authentic methods for gathering evidence, methods of reporting to parents and the like. The Primary Program materials are also supplemented with Study Team resource materials to be used by teams of teachers, administrators, support staff, and parents working to implement the Primary Program in their schools. Vermont also provided a contrast to states complaining about lack of appropriate training of preservice teachers by universities. The University of Vermont has taken the lead in training early childhood professionals in play-based assessment.

Some respondents also commented on the need for professional training of another type, that is to make teachers better informed about the appropriate uses of developmental screening information (which ties back to the misuse of screening measures noted as part of Findings 1 and 4). IQ-like tests are often used by teachers who are not trained in interpretation and who are not aware of how prior experience can affect performance. But equally serious problems arise if school psychologists, knowledgeable about tests, lack experience with pre-school age children. An example was given of a fidgety 4-year old being labeled as "hyperactive" by a psychologist used to working with older children. In addition, there seems to be some confusion in the early childhood community about the different purposes of assessment. As a result there is not a clear understanding of how the content and form of assessment should be adapted to the intended purpose. Although one Head Start director clearly distinguished between a portfolio, work-sampling, on-going approach to assessment as part of instruction and the one-day-per-year screening for possible disabilities, he estimated that most Head Start professionals were confused about the differences between screening and instructional assessment. The concerns mentioned about adequate training by different respondents included psychologists, Title 1 teachers, Head Start teachers, and classroom teachers without adequate knowledge of child development. Early childhood special education teachers "know kids" but often are misinformed about clinical assessment.

Finding 6. Testing of preschool-age children is largely driven by mandates for categorical programs.

Many respondents expressed concern about the pressure to test and categorize children in order to receive funds. In the midwest, one respondent said, "If we are only using assessment or screening to access funds, that is a dangerous precedent. Unfortunately that is the way the system is built right now. We need a national policy to address the issue. There is a need for flexibility and creativity, creating programs in the best interest of children rather than a program strictly based on separate labels." In the northeast, another respondent stated his concern about "labeling children to get funding." There is no funding for regular pre-kindergarten so children get categorized in order to get funding. He noted further that children who fall in the mild range and those who demonstrate behavior problems are more likely to be affected by the system's decision to require labels to get resources. "Once you get a label, it may become difficult 'to get out of the box.'" Washington state recently changed its funding formula for special education. In the past, the amount of money allotted varied with the severity of a child's problem. Under this system, people were assessing children to get more money rather than focusing on program development. Now that everyone will get the same amount of money per child, it is hoped that unnecessary testing will be reduced. The downside of this solution is that funding is the same for an autistic child and one who is mildly speech impaired despite the great differences in the cost of services that must be provided.

In contrast to the majority of comments, a respondent from Montana said that there was not a lot of assigning of labels to 3, 4, and 5-year olds because of non-categorical aspects of IDEA. Also, in North Dakota, they are making an effort to offer services to "borderline children." These are children who are not eligible under Part H or 619 programs.

Some respondents also stated concerns about the over identification of children in mild special needs categories. One respondent in the far west was concerned about labeling students as "delayed" when the perceived problem may be a language or cultural difference. Another respondent on the east coast also commented that "children who have special needs but do not need special education are being over identified, especially non-native English speakers." In South Carolina, labeling associated with the readiness score is acknowledged to be a major issue. Minority students are over represented, and once students get labeled "not ready" they tend to stay in compensatory programs indefinitely. Act 135 was adopted to help remedy this situation. The act allows for more flexibility in programming and the inclusion of more students in the program. There is less use of pull-out services and more effort to work with students in the regular classroom.

Finding 7. Only a few states are collecting data to be able to report on progress toward Goal 1. These data are primarily health and welfare indices.

To report on progress toward Goal 1, states might either gather data on the goal itself, by measuring children's health, social adjustment, and cognitive development, or states could monitor attainment of the enabling objectives by reporting on the quality and availability of preschool programs, parent training programs, and health care programs for pregnant mothers and infants.

Because so few early childhood specialists in state departments of education were aware of efforts to monitor progress toward Goal 1, the study sample was expanded to include designated Goals 2000 contacts. Although the interview questions focused on data being gathered to monitor progress toward Goal 1, goals contacts invariably began by talking about programs or the lack thereof. Even from the perspective of "goals" spokespersons, in more than half the states there is little Goal 1 activity to report. Note, however, that lack of "Goal 1" efforts in a state does not necessarily mean that there are no health or social services initiatives for young children that state; it means that these initiatives may not be identified with Goal 1, may have

predated the creation of the National Education Goals, or maybe primarily lodged outside the education system. When prompted, respondents returned to the question of data used to monitor progress. The summary that follows describes first the reported programmatic trends and then the responses regarding indicators of progress.

At the time of this survey, five states were not formally participating in Goals 2000. Additionally, respondents in more than 20 states answered that "not much is going on" regarding Goal 1. In several southern states, for example, Goal 1 was seen as too controversial "even to speak about" because of opposition from the religious right that sees government involvement in early childhood as an intrusion into the rights of the family. In Louisiana, for example, the state legislature voted to eliminate the early childhood initiative after protest from conservative groups. In a number of states, Goals 2000 efforts are not directed toward early childhood but toward technology, Goals 3 and 4, or staff development. In many states, local districts are free to make their own choices and have not opted for early childhood programs. In some states, state plans have just been completed or funds "have just been released from the Governor's office" so there was as yet no program implementation.

A few states have launched major early childhood initiatives but without creating a monitoring system of all children in the state. Washington state's Readiness to Learn initiative is an example. As part of its 1993 Education Reform Act, Washington funded 22 local projects aimed at improving services to children and families through interagency collaboration. Appropriately, evaluation of this effort has included information on program implementation and outcomes at the 22 local sites and a state-level evaluation of Readiness to Learn contracted with RMC Research. In another state, the respondent emphasized that the Goal 1 focus was intentionally on programs and not data collection to maximize the resources going directly to children.

When pressed about data gathering efforts, to report on progress on Goal 1 itself or on any of the objectives, respondents gave one (or more) of the following answers in roughly descending order of frequency: (1) There is nothing going on with Goal 1 and no data gathering effort; (2) Data on low birthweight babies and other "child well-being" statistics are available through Kids Count; (3) Local projects are responsible for evaluation; (4) Yes, we have a state progress report which includes Goal 1 indicators; (5) We have established a statewide data base for tracking students identified as at risk.

Health and social welfare data are available for all 50 states through the Kids Count project funded by the Casey Foundation that could be used to monitor progress on the third objective. Both national and state-by-state data are reported on indicators such as percent low birthweight babies, infant mortality rate, percent of children living in poverty, and percent of 2-year olds immunized; trend data and relative change since 1985 are also provided. An interesting finding from our survey was that many early childhood specialists and most Goals 2000 contacts were not aware of the availability of Kids Count data. Kids Count was acknowledged (by one or more respondents) in only 16 states. As shown in Table 4, some states incorporate Kids Count data in their reporting of progress on early childhood goals. In addition, respondents from eight other states mentioned the availability of the Kids Count document or said that Kids Count is "used a lot to inform policy." Of course, two or three phone calls per state are not sufficient to locate for certain the policy makers and program staff who are making use of Kids Count data; so lack of mention does not mean that the data are not being used. Nonetheless, infrequent reference to Kids Count when asked about available data suggests that attending to these health and poverty indices may not be salient for early childhood specialists in most states. In fact, early

childhood coordinators in only three states knew the name of the person or agency responsible for doing the Kids Count survey.

Respondents in nearly a dozen states interpreted our question about "monitoring progress toward Goal 1" to mean evaluation of programs funded through Goals 2000. They said that it was left to locals to evaluate local programs. Unlike the example from Washington state, cited above, however, there did not appear to be any systematic state-level collation or synthesis of local evaluations. Several other states said that their main data gathering efforts were to track children receiving special services. For example, in North Dakota the Early Childhood Tracking System identifies children at birth who may be at risk for health problems, poverty, or cognitive delays. The system includes an Infant Child Monitoring Questionnaire administered by parents to children 4 to 48 months and a professional screen administered by professionals at age 3 or 4. The purpose of the tracking system is to improve delivery of services to children who need them. Similarly, states such as Delaware and Maine have established tracking systems for children with disabilities. In theory, such data bases could also be used to determine whether progress is being made in providing appropriate preschool programming to disadvantaged and disabled children (objective 1); but that is not the intention at present.

Only a few states have initiated early childhood initiatives in the name of Goal 1 and at the same time established indicator systems to monitor improvement as a result of these initiatives. Examples of states reporting year-to-year progress on various indicators are shown in Table 4. Additional states are in the process of developing indicators. South Carolina uses Kids Count and has recently established benchmarks for reaching the national goals but has not yet begun to report annual data in relation to the benchmarks. The Ohio annual reports are particularly noteworthy because they tie statistical indicators, such as the percent of children in poverty participating in Head Start, to programmatic efforts aimed at expanding services. Some of the "indicators" selected by states serve the political purpose of focusing attention on early childhood, but do not meet strict criteria as statistical indicators. For example, simply reporting the number of children participating in preschool programs could be misleading because increasing numbers could reflect either a growing population or an improved percentage of children in need being served. To ensure accurate comparisons across years, it would be preferable to report the percent of 4-year olds being served or the percent of children meeting a poverty criterion being served. Two states use teachers' reported opinions about children's readiness as an indicator, but over time this could be problematic because change could be the result of changes in teachers expectations or changes in the characteristics of children.

Finding 8. Parent involvement was an issue both for the identification of children with special needs and because of parent demands for standardized testing.

The interview protocol did not include specific questions about parental involvement, yet the need to involve parents or educate parents was mentioned in one-third of the initial interviews as an important part of the assessment picture. States such as Florida and West Virginia are trying to provide "one stop shopping for assessment of young children and family services" through community service centers. The North Dakota Tracking System for at-risk children involves parents in the collection of data on their children but also helps parents understand assessment results and appropriate follow-up activities. Several respondents stressed the importance of involving parents and "demystifying" the special needs identification process.

Equally important was the need to educate parents about the dangers of overinterpreting test scores and to increase receptivity to new forms of assessment. Several respondents said there was pressure from parents to test and to test early. One respondent commented on the need to educate the public, politicians, and teachers about the limitations of testing. Test scores are

emphasized as if they were "magic." In another state, the respondent went further: "Parents' perceptions of what assessment is and isn't needs to be addressed. Parents are used to tests and grades. It takes a while for them to get used to looking at it another way. There is an incubation period for parents (to become familiar with) parent/teacher conferences and portfolios." Similarly, in another state, "In the past, report cards were sent home and there was not much of an effort to bring parents into the school for conferences to explain the assessment. Now the schools are making an effort to show parents the connection between assessment, curriculum, and instruction." Comments about educating and involving parents were closely tied to state efforts to eliminate inappropriate uses of readiness testing and to support development of new forms of assessment.

Finding 9. Lack of collaboration and coordination among agencies serving young children continues to be a problem. In particular, specialists lack knowledge about assessment requirements across programs.

Although the issue of coordination of services among agencies was raised by only a half-dozen respondents, the similarity of concerns when mentioned warrants documentation as a theme. Respondents expressed concern that the early childhood system is "still very fractured." One said regrettably that the state had been making progress in working across state departments but the breaks had been put on as a result of the most recent election. Another said, "there was an effort to put all the agencies dealing with young children in one department, but now they are split up and are more fragmented than before." This last comment came from a state where the Goal 1 report announces a plan for collaboration among state and local government and private agencies. National reports document these problems on broader scale. According to a 1994 report by the Government Accounting Office (GAO), there are 90 different early childhood programs administered by 11 different federal agencies. One disadvantaged child may be eligible for 13 different programs but the majority of disadvantaged preschool-aged children are in no program at all (GAO, 1994). Lack of sufficient funding, narrow mission statements, and inability of programs to work together prevent early childhood centers from providing children with a full range of services (GAO, 1995). "Nevertheless, some state and local initiatives, sometimes with additional private funding, have demonstrated that a full range of services can be provided by funding the programs that offer more services; investing state money in Head Start; and locating centers, services, and children together even when funds come from different programs" (GAO, 1995, p. 5).

Respondents from two states described efforts to offer coordinated services. In Massachusetts, the report of the Special Commission on Early Childhood (1995) identified several steps to help develop a coordinated system across public and private programs including: creation of Board of Early Care and Education overlapping with the Board of Education, unified state standards, coordination of professional development projects across agencies, and an integrated data management system for early childhood programs. The Washington Readiness to Learn Initiative has as its main strategy "to enable schools and social service organizations to work together to provide comprehensive services for young children and their families."

In parallel to the finding that programs and services lack coordination, we also found lack of coordination and knowledge about assessment requirements affecting young children. However, data supporting this finding are not in the form of quotations from respondents because it was not something that respondents stated directly as a concern. Rather, we noted that early childhood specialists and test directors lacked detailed knowledge about Head Start, Part H and IDEA, and Title 1 assessment requirements because of the number of times we were referred to other departments to obtain answers to our questions. For example, in one state the elementary

coordinator referred us to IDEA for answers to questions about screening questions and to Head Start and Title 1 directors for information on other populations; the Head Start director said we should try the child development divisions for information about authentic assessment; and the special education director said try Migrant Education for evaluation of LEP students. In most states, our survey sample included both a Goal 1 contact and an early childhood coordinator, but in several instances only one of the two respondents knew of Goal 1 activities going on in the state. As noted under Finding 7, the majority of early childhood specialists and testing directors were not aware of Kids Count data available for their state. Early childhood specialists tended not to know how Chapter 1 requirements had shaped local district testing policies in the past and were not aware of how things might change under the new Title 1 guidelines.

An important exception to this trend of compartmentalized knowledge was that early childhood specialists tended to be informed about the state assessment program at least as it affected the early grades and sometimes even had been instrumental in affecting state level assessment policy. Also, as noted under Finding 1, most (but not all) early childhood specialists were able to report on state level mandates that required formal screening of children upon school entry. However, some of these respondents might not have had detailed enough knowledge of the federal requirements to be aware that there were other means available to satisfy IDEA than by every-pupil testing.

SUMMARY AND RECOMMENDATIONS

Based on telephone interview data gathered from early childhood specialists, test directors, Goals 2000 contacts, and other program specialists representing the 50 states, there appear to be significant changes since the late 1980s in assessment policies and practices affecting young children. Most states have made efforts to move away from readiness testing and associated practices such as denial of school entry, kindergarten retention, and extra-year transition programs. Although these practices have reportedly decreased in response to state department advisories, in-service training, and the like, they nonetheless continue in some school districts. Some states have also eliminated formal cognitive screening for potential handicaps because such tests were being misused to determine readiness for school; other states have attempted to clarify how screening results should be used. Misuse of screening measures appears to be less widespread than was reported in 1988 by the Gnezda and Bolig survey but continues to some extent and may be of particular concern in those states that still mandate every-pupil testing. Concomitant with these trends, and influenced by policy statements of the NAEYC and NAECS/SDE and efforts of early childhood specialists within particular states, there has been a trend to eliminate state-mandated standardized testing for children below grade three. However, many local districts still administered standardized tests in grades 1 and 2.

Elimination of inappropriate testing has been replaced in some states and districts by attempts to develop new forms of assessment that are more supportive of instruction. Examples include curriculum-embedded assessments, portfolios, the Work Sampling System (Meisels, 1992) and the High/Scope Child Observation Record (COR) (1992). However, survey respondents identified a serious need for better training of professionals to understand what developmentally appropriate assessment means and how to use new forms of assessment. Although many states may have invested in early childhood programs in recent years, the majority of states do not have initiatives specifically identified with Goal 1. Fewer than a dozen states have reporting mechanisms in place to report on progress toward Goal 1. The Kids Count project, funded by the Casey Foundation, reports both national and state-by-state data each year on indices such as percent low birthweight babies, percent of children living in poverty, and percent of 2-year olds

immunized. Kids Count data could be used to report on progress toward Goal 1, but most respondents did not know of its existence. Similarly, we documented lack of knowledge among specialists about assessment requirements across programs. For example, early childhood specialists and test directors could not give accurate information about testing done for IDEA. Other issues identified by multiple respondents were: the need for parental involvement in assessment, the undesirability of having to test and label pre-school-age children in order to obtain services, and lack of collaboration and coordination among agencies serving young children.

Screening and identification of children with special needs. It is imperative that children at risk for developmental delays be identified so as to receive services. Although identification in the early years may be more problematic -- because measures are less reliable for very young children and because it is more difficult to distinguish between temporary differences in rate of development and more persistent delays, early identification is important because early intervention may be more effective than attempts to intervene later. However, the importance of identifying children with special needs must be weighed against the misuse of cognitive screening measures. Tests such as the DIAL-R, the Peabody Picture Vocabulary Test, and the Battelle continue to be misused in some local school districts to deny school entry, to group students by ability, or to plan instruction. Screening is only the initial step in identification of children with special needs. When formal screening of every child is mandated by a state, tests are most often administered by classroom teachers who are not trained in psychometrics or clinical assessment. Lack of training contributes to the likelihood of misinterpretation. In all cases, whether referrals are based on teacher and parent concerns or formal test scores, initial referrals must be followed up by an in-depth developmental assessment administered by trained professionals. Because we know of no evidence to suggest that identification of children with special needs is either more accurate in those states with mandated testing or reaches more children earlier, we recommend that states eliminate formal, every-pupil cognitive screening requirements and instead establish informal referral mechanisms. Except for vision and hearing checks, the majority of states no longer has mandated screening programs; this trend should continue.

Instructional improvement. Many districts and schools are attempting to implement new forms of assessment that are more supportive of instruction. Instead of IQ-like tests, which are intentionally not tied to school curricula, assessments designed to support learning should explicitly model the learning progression intended by the curricula, whether in language development, emergent literacy, early numeracy and learning about patterns, awareness of self, family, and community, and so forth. And, of course, the form of such assessments must be appropriate to the age and experiences of individual children. This represents both a huge instrument development task, i.e., to develop appropriate content and observational methods closely tied to the curriculum, and a huge professional development task to ensure that classroom teachers are able to use new assessments as an integral part of instruction. As one respondent noted, teachers are collecting "all this stuff" for portfolios, but it is overwhelming, they don't know what to do with it. States could support this trend by developing model assessments, or by serving as a clearinghouse to identify effective practices, and by providing support for staff development.

School accountability and indicator systems. We support the general approach taken in the proposed new Title 1 guidelines which distinguishes between assessments which must be done to monitor and support the learning of individual children and assessment for accountability which must be done once in each key grade interval. Consistent with the trend to eliminate standardized testing below grade 3, the new Title 1 guidelines permit waiting until grade 3 to test for program evaluation purposes.

When academic achievement is assessed for purposes of accountability it is essential that measures have the same close correspondence to the desired curriculum as is needed for instructional measures. However, because scores are publicly reported and may be used to make important decisions, accountability measures must meet stricter standards for reliability and comparability across schools than classroom assessments. If cognitive measures are administered to children below grade 3 for purposes of monitoring readiness or achievement trends, we recommend that matrix sampling be used to reduce the testing burden on individual children and to avoid the misuse of individually reported scores.

For purposes of monitoring progress toward Goal 1 we recommend first that policy makers seek out and review the data already available in each state from Kids Count and various state agencies. Then it is desirable that policy makers set programmatic goals -- what health and educational conditions are most important to improve? Lastly, policy makers should seek indices that can be reliably reported across years and that are closely tied to programmatic efforts. For example, if participation in high-quality preschool programs by disadvantaged children is a targeted objective, then criteria should be set to define high-quality programs, to specify the age range of eligible children, and to standardized the definition of disadvantaged; then data could be collected annually to see if an increased proportion of such children are being reached.

Education of teachers, administrators, parents, and policy makers regarding assessment. As indicated throughout this article, there is a serious need for more adequate pre-service and in-service preparation of teachers and administrators in the area of assessment. Those working with young children must understand not only the different purposes of assessment, but they must learn how to match assessment purpose and strategy. Parents are legitimately concerned about how their child is doing, just as legislators demand information about the performance of schools and school districts. But parents and policy makers may not be aware of the difficulties in measuring the achievement of 5-year olds or of the problems arising when IQ-like screening instruments are misused -- to determine readiness or plan instruction. Good assessment models being developed to support learning in the classroom and those being developed for accountability purposes are important because they will provide better measurement and data. They are also important because they will help to educate parents and policy makers about the kinds of information that are the most useful and meaningful.

REFERENCES

Bredekamp, S. (Ed.) (1987). Developmentally Appropriate Practice in Early Childhood Programs Serving Children from Birth through Age 8. Washington, DC: National Association for the Education of Young Children.

Commission on Chapter 1. (1992). Making Schools Work for Children in Poverty. Washington, DC: Council of Chief State School Officers.

Office of Technology Assessment. (1992). Testing in American Schools: Asking the Right Questions. Washington, DC: Office of Technology Assessment, Congress of the United States.

Gnezda, M. T., & Bolig, R. (1988). A national survey of public school testing of prekindergarten and kindergarten children. Paper prepared for the National Forum on the Future of Children and Families and the National Association of State Boards of Education.

General Accounting Office. (1994, October). Early Childhood Programs: Multiple Programs and Overlapping Target Groups. Washington, DC: Author.

General Accounting Office. (1995, March). Early Childhood Centers: Services to Prepare Children Often Limited. Washington, DC: Author.

Hatch, J. A., & Freeman, E. B. (1988). Who's pushing whom? Phi Delta Kappan, 70, 145-147.

High/Scope. (1992). Child Observation Record. Ypsilanti, MI: High/Scope Press.

International Reading Association. (1986). Literacy development and pre-first grade: A joint statement of concerns about present practices in pre-first grade reading instruction and recommendations for improvement, Childhood Education, 63 (2), 110-111.

McLaughlin, M. W., & Shepard, L. A. (1995). Improving Education through Standards-Based Reform. Stanford, CA: National Academy of Education.

Meisels, S. J. (1992). The Work Sampling System. Ann Arbor, MI: Center for Human Growth and Development, University of Michigan.

National Association for the Education of Young Children. (1988). NAEYC position statement on standardized testing of young children 3 through 8 years of age, Young Children, 43, 42-47.

Shepard, L. A. (1994). The challenges of assessing young children appropriately. Phi Delta Kappan, 76(3), 206-212.

Special Commission on Early Childhood. (1995). Children First: A Plan for an Early Care and Education System for Massachusetts. Boston, MA: Massachusetts Department of Education.

Smith, M. L., & Shepard, L. A. (1988). Kindergarten readiness and retention: A qualitative study of teachers' beliefs and practices. American Educational Research Journal, 25 (3), 307-333.

Special Commission on Early Childhood. (1995). Children First: A Plan for an Early Care and Education System for Massachusetts. Boston:

Figure I
Sample goals and performance indicators from the North Carolina on-going assessment of Grade 1 mathematics.

Goal 1: The Learner will identify and use numbers, 0 to 100.

1. 1 Count using 1 to 1 correspondence		
1.2 Make sets; Match numerals		
1.3 Compare and order sets		
1.4 Identify ordinal position		
1.5 Conserve numbers		
1.6 Read and write numerals; Read number words 0 to 10		
1.7 Count on		
1.8 Recognize one more/ less/ before/ after/ between		
1.9 Compare/ sequence numerals		
1. 10 Rote count by 1's, 10's, 5's, and 2's; Group objects to count by 10's, 5's and 2's		
1. 11 Count by using tallying		
1. 12 Make reasonable estimates of "how many"		
1. 13 Group objects into tens and ones; Record		
1. 14 Recognize models; Build 2 digit numbers; Write numerals		
1. 15 Represent numbers in a variety of ways		

Goal 2: The learner will identify and use geometric ideas.

2.1 Identify open and closed figures		
2.2 Identify, describe, and model plane figures; recognize in environment		
2.3 Use directional/ positional words		
2.4 Describe likenesses and differences		
2.5 Identify and describe solid figures; recognize in the environment		

Goal 3: The learner will demonstrate an understanding of classification, patterns, and seriation.

3.1 Describe objects by their attributes; Compare and order		
3.2 Sort by given attribute; by more than one attribute; explain sorting rules		
3.3 Sort objects by own rule; explain sorting rule		
3.4 Copy/ continue patterns; translate into different forms		
3.5 Create patterns with actions/ words/objects		
3.6 Find and correct errors in patterns		
3.7 Identify patterns in the environment		

Code: M = Most of the Time
S = Sometimes N = Not yet

Table 1
Examples of State-Level Efforts to Eliminate Readiness Testing, Denial of School Entry, and Extra-Year Placements

Arizona	Arizona Dept. of Ed. has made an effort to get the word out to parents and local districts that districts do not have the right to deny a child entrance to kindergarten.
Delaware	State law mandates that all 5-year olds must be in school.
Illinois	Early Childhood Office discourages use of readiness tests which are (nonetheless) used by some districts.
Iowa	State guidelines recommend "assessment, screening, and evaluation of children should only be used to improve instructional practices, not label, track or retain children in the earliest years of their development."
Kentucky	State regulations prohibit the use of screening for "determining placement" or to "restrict entry to exit from preschool programs."
Louisiana	State Board of Education has clarified requirement to screen entering kindergartens by stating that screening should be used for planning and instruction and not to exclude children from programs.
Nebraska	State school accreditation regulations prohibit the use of screening for kindergarten entrance.
New Hampshire	The Dept. of Ed. discourages the use of extra-year classes.
Oklahoma	Prior to 1994, screening was used by local districts to label children "not ready" so the requirement was eliminated at the kindergarten level.
Pennsylvania	In 1990, in a Basic Education Curricular, the Commissioner strongly urged "all districts to curtail readiness and screening tests which are not used solely for diagnostic purposes. Readiness and screening tests are to be used to assist the instructional process, not to deny service and programs to children."

Table 2
Examples of States that Require Formal Developmental Screening of 5- and 6-Year Olds

Arkansas	State requires that all kindergarten students must undergo a comprehensive health and developmental screening prior to school entry.
Florida	A 1996 state regulation requires children's readiness for kindergarten be measured by an instrument selected by local districts. Results are to be used for program evaluation.
Kentucky	State requires local districts to screen all entering kindergartners within 30 days of enrollment. Screening is not to be used for program placement purposes but for planning activities and evaluating progress.
Louisiana	1987 state regulation states that "each district shall administer a nationally recognized screening instrument to every child entering kindergarten for the first time," to be used for planning and instruction.
Maine	State requires all children to be screened prior to school entrance to identify possible eligibility for exceptional child services.
Maryland	State requires administration of the Maryland Observation Screening Checklist for kindergartners to determine need for further assessment and possible eligibility for special services.
Minnesota	State requires that all children must be screened with a health and developmental screen within 30 days of entry to kindergarten to determine need for further assessment and possible eligibility for special services.
Mississippi	No state mandate but state recommends that local school districts screen all children who are age 5.
New Mexico	State requires screening of all 5-year olds and all new entering students to determine need for further assessment and possible eligibility for special services.
New York	State requires screening for all new entrants to school in 6 areas: physical, cognitive, expressive language, receptive language, vision and hearing.
Ohio	State requires that local districts screen children either when entering kindergarten or first grade. Screen includes: vision, communication/ language , physical/health, and a section to detect developmental delays.
Oklahoma	State requires readiness screen prior to first grade. Screen is used to determine need for further assessment and possible eligibility for special services and for instructional improvement.
Tennessee	State requires all kindergartners be screened upon entry into kindergarten. A readiness component is included in the screening process.

Table 3
Grade Levels Tested in State Assessment Programs

	K	1	2	3	4	5	6	7	8	9	10	11	12
Alabama				3	4	5	6	7	8	9	10	11	
Alaska					4		6		8				
Arizona					4			7			10		
Arkansas						5		7			10		
California										9	10	11	12
Colorado													
Connecticut					4		6		8		10		
Delaware				3		5			8		10		
Florida					4				8		10	11	
Georgia	K			3					8			11	
Hawaii				3			6		8		10	11	12
Idaho					4				8			11	
Illinois				3	4		6	7	8		10	11	
Indiana			2	3			6		8	9			
Iowa													
Kansas				3	4	5		7	8		10	11	
Kentucky					4	5			8			11	12
Louisiana	K			3	4	5	6	7			10	11	
Maine					4				8			11	
Maryland			2	3	4	5	6	7	8	9	10	11	12
Massachusetts													
Michigan					4	5		7	8		10	11	
Minnesota					4	5	6		8	9		11	
Mississippi					4	5	6	7	8	9		11	
Missouri			2	3	4	5	6	7	8	9	10	11	
Montana					4				8			11	
Nebraska													
Nevada					4				8			11	12
New Hampshire				3									
New Jersey									8			11	12
New Mexico		1	2	3	4	5	6		8		10		
New York				3	4	5	6		8	9	10	11	12
North Carolina				3	4	5	6	7	8	9	10	11	
North Dakota				3			6		8			11	
Ohio					4		6		8	9			12
Oklahoma				3		5		7	8			11	
Oregon				3		5			8		10		
Pennsylvania						5	6		8	9		11	
Rhode Island					4				8		10		
South Carolina				3	4	5	6	7	8	9	10	11	
South Dakota					4				8	9		11	
Tennessee			2	3	4	5	6	7	8	9	10	11	12
Texas				3	4	5	6	7	8	9	10	11	12
Utah		1	2	3	4	5	6	7	8	9	10	11	12
Vermont					4	5			8		10		
Virginia					4		6	7	8	9	10	11	12
Washington					4				8			11	
West Virginia		1	2	3	4	5	6	7	8	9	10	11	
Wisconsin				3	4				8		10		
Wyoming													

Table 4
Examples of States Using Data to Monitor Progress on Goal 1

Colorado	Colorado Kids Count is used to monitor progress on the Colorado Kids Campaign for the Decade of the Child Goals. Colorado Department of Education tracked Goals 2000 progress with 20 indicators for 3 years but has discontinued this effort.
Florida	State and county trends are monitored using indicators such as births, incidence of low birth weight, infant/child death rates, and school success indicators such as grade retention and performance on standardized tests.
Hawaii	Kids Count document is done through University of Hawaii Family Center. The Governor's Office on Children and Youth has collected data and surveyed teachers to determine the number of children served in pre-K, but the effort was suspended in 1996.
Kansas	State goals documents lists two early childhood performance indicators: participation rates in Parents as Teachers and in Even Start.
Maryland	Report on Maryland's Progress Toward National Education Goals includes trend data on percent of births to single teens, pre-K enrollment, and immunization by age two.
New Jersey	New Jersey's Progress Toward the National Education Goals Report lists 3 indicators of success on Goal 1: in 1995 9 cities offered Good Starts programs, 2,603 children enrolled in Good Starts programs in 38,352 children received social and health transitional services.
Ohio	Ohio's Annual Progress Report on Education reports data such as percent of 3- and 4-year olds in poverty who participate in Head Start or preschool, percent of school districts with family support programs, and percent of 2-year olds immunized.
Oregon	Oregon Benchmarks documents reports health and social indicators and percent of kindergartners meeting developmental standards in language and literacy as reported by teachers.
Vermont	State Dept. of Ed. is working with Dept. of Human Services to collect data. Developed indicators including percent low birth weight, immunization rate, percent of children perceived by kindergarten teachers to be ready for school. Also conduct annual survey of parents and teachers to find out what young children need to be successful in school and to what extent these conditions exist in their communities.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☒ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (3/2000)