

Assessing: → **Young Children**
→ **Investments in Early Care and Education**
→ **Early Care and Education Providers**

PAES Invest in Kids Meeting
April 25, 2007

Dr. Thomas Schultz, Early Childhood Accountability Project
The Pew Charitable Trusts – tschultz@pewtrusts.org

The Early Childhood Accountability Project is supported by the Foundation for Child Development, the Joyce Foundation and The Pew Charitable Trusts.

Overview/Road Map

- **Assumptions**
- **Current efforts**
- **Child assessment tools: challenges, approaches, strengths and weaknesses**
- **Points of view**
- **Recommended principles in the selection and use of child and program assessment tools and data**
- **Who is researching what/over what timeframe?**
- **3 state approaches in using child and program assessment data**

ASSUMPTIONS

- Investments in high quality preschool and other early care and education services will improve children’s readiness for school.

- “Readiness” will lead to early school success which will translate into better educational outcomes.

- Enhanced educational outcomes will yield more productive citizens and cost savings.

- Assessments of programs and children can:
 - Guide decisions and investments.

 - Inform the public.

 - Advance understanding of most effective approaches, practices, models of service.

 - Identify higher performing and lower performing local service providers.

 - Target and motivate program improvement efforts.

CURRENT EFFORTS: 3 “PURPOSE-DRIVEN” APPROACHES

- Assess on how ***all young children*** are learning and developing:
 - Setting priorities/monitoring progress for state early childhood systems efforts.
 - Baseline information for elementary schools.

- Assess effectiveness (quality, outcomes, impacts) of ***specific programs***/investments:
 - Evidence to support increased investment or reexamining program strategies.
 - Program-wide improvement efforts

- Assess performance of ***local service providers***:
 - Guide technical assistance to provider agencies.
 - Award incentives/public recognition for program improvement.
 - Target providers for defunding if serious deficiencies can't be remedied.

CURRENT EFFORTS AT-A-GLANCE

	State Pre-K	Head Start	Special Education	Child Care
Program Quality Standards	State Program Standards	Federal Program Performance Standards	IDEA regulations State program standards	State Licensing Standards State Quality Rating Systems (QRS) (13 states + 29 pilots)
Assessing Program Quality	State Program Monitoring (30 states)	Federal Program Monitoring	State Program Monitoring	State Licensing Visits QRS Assessments
Standards for Children's Learning	Early Learning Guidelines (49 states)	Head Start Child Outcomes Framework	3 functional goals	Early Learning Guidelines (49 states)
Child Assessments	Pre-K - 12 states Kg. - 16 states	National Reporting System	States report % of children in 5 categories on 3 goals	No current requirements
Program Evaluations	25+ State/Multi-State Pre-K Evaluations	Head Start Impact and FACES Studies	Federal Program Evaluation	NICHD Child Care Study Cost, Quality & Outcomes Study

CURRENT EFFORTS: WHAT'S NEW? (A Lot)

- New state/federal standards for children's learning
 - Head Start Child Outcomes Framework - 2000
 - OSEP - 2005
 - State Early Learning Guidelines – 10 states in '99, 49 states + DC in 2005
- New large-scale, ongoing child assessment systems
 - Head Start National Reporting System – 2003-4
 - OSEP – 2006-7 baseline assessments
 - Expansion of state assessments of kg. and preschool-aged children
- New/revised program quality criteria and assessment efforts
 - Quality Rating Systems - 5 states in 2001, 13 states + 29 in planning in 2006
 - Revised NAEYC Accreditation System – 2005
- Multiple forms/formats of new data on children and programs, including research and evaluation studies as well as ongoing reporting and program improvement initiatives.

CHILD ASSESSMENT TOOLS: KEY CHALLENGES

- Uneven pace and variable rates of development and learning.
- Broad range of areas of learning and development are of interest:
 - Cognitive/academic skills/knowledge
 - Social/emotional competence
 - Language development
 - Physical development
 - Approaches to learning
- Early childhood programs serve increasingly diverse populations, so cultural/language differences complicate assessments.
- Children's aren't capable of responding to inexpensive group-administered paper-and-pencil tests.
- Therefore, adults must assess each child individually, which means:
 - Time-consuming and expensive.
 - Large numbers of people must be trained and supervised.
 - Special risks to the consistency, integrity of data in large-scale efforts.

CHILD ASSESSMENT TOOLS: APPROACHES, STRENGTHS, WEAKNESSES

Observational Tools	Direct/"On Demand"/ Adapted Direct Tools
<i>How They Work</i>	<i>How They Work</i>
Assessors use checklists, rating scales to document children's capabilities, based on observations over time of their behavior, language, interactions, performance in different contexts.	Assessors present children with standardized set of questions or tasks and record children's responses.
<i>Strengths</i>	<i>Strengths</i>
<ul style="list-style-type: none"> - Assessors have the opportunity to observe the children over time in a variety of contexts to confirm the accuracy of their ratings. - Available tools cover all domains of child development and learning. - Assessment process is inobtrusive and does not require removing children from their classrooms. - Risks of "teaching to the test" are minimized because these assessments are not composed of individual questions that teachers could use to "coach" children, and because tools are often aligned to the local program's curriculum. 	<ul style="list-style-type: none"> - Due to the structured nature of questions and method of eliciting direct responses from children, there are lower risks of errors based on the assessor's judgment. - Use of a common set of questions creates the perception that results are more objective than ratings generated by observers. -The scope, depth and costs of required training is typically lower than training for observational tools. - Allows programs to compare the performance of children to norms for nationally-representative samples of similarly-aged children.
<i>Limitations</i>	<i>Limitations</i>
<ul style="list-style-type: none"> - Assessors must be well-trained in order to carefully observe and analyze children's behavior, discourse, work samples and other evidence and generate appropriate ratings. - Bias in teacher ratings can occur when they don't share the same cultural and linguistic background as the children they are assessing. - Quality of ratings can decline over time or suffer from "drift". - If assessors are staff members in a local agency and their agency's reputation or funding will be influenced by test results, there are risks of bias or inflating of ratings to show more rapid progress or higher end-of-program outcomes . 	<ul style="list-style-type: none"> - Assessors must be trained in order to ensure consistent administration, to develop rapport with children and address behavioral challenges. - Children may be distracted or may not perform at their best if they aren't comfortable with the assessor. - Children must be able to process language well/ challenges in developing tools to assess English language learners. - Cultural differences and a program's curriculum may influence how children respond to questions or tasks. - Not appropriate for assessing some important goals, notably social and emotional development. - Use of a specific set of questions creates risks that teachers could coach children to inflate outcomes.

POINTS OF VIEW

- **General support** among assessment experts/professional leaders for the following uses of child assessments:
 - Teachers assessing individual children to plan instruction and communicate to parents.
 - Identifying children with disabilities and other barriers to learning.
 - Conducting research.
 - Evaluation of specific state or federal programs.
 - Reporting on the overall status of children in a state or local community.
- **General opposition** for the following uses of child assessments:
 - Use of a single test score to track, retain or deny access to services for an individual child or group of children.
 - Using assessments that are limited to a single curriculum area, e.g. literacy as the sole basis for judging programs or children.
 - Assigning rewards or sanctions for local provider agencies solely on the basis of children's end-of-program assessment results.
 - Using child assessment data to evaluate individual teachers.
- **Mixed opinions/lack of consensus** on:
 - Should child assessment data on children's progress over time be used, along with program quality measures, as one factor in judging the performance of local provider agencies?

RECOMMENDED PRINCIPLES

- Breadth in scope of assessments to capture full range of key elements in children's early learning and development and well-functioning programs.
- Assess samples of children/programs to limit costs and enhance quality of data.
- Design analysis and reporting to draw attention to context, connections, trends over time in child and program performance measures.
- Assessment tools should be valid, reliable, culturally and linguistically appropriate.
- Coherence, alignment, efficiencies of standards, assessments, professional development
 - Across funding streams and programs
 - Across early childhood-third grade continuum.
- Provide sufficient funding to enable programs to meet performance standards, support high quality assessments and program improvement efforts.
- Train teachers, managers and policymakers to understand and use child and program assessment data.

WHO IS RESEARCHING WHAT?

- **Reviews of child assessment tools:**

Early Childhood Measures Profiles – Berry, Bridges & Zaslow -
(http://www.childtrends.org/Files//Child_Trends-2004_09_01_FR_ECMeasures.pdf)

A Compendium of Measures for the Assessment of Young English Language Learners – Sandra Barrueco and Michael Lopez – Early Childhood Accountability Project – forthcoming, fall, 2007

- **The National Early Childhood Accountability Task Force**

Supported by The Pew Charitable Trusts, the Foundation for Child Development and the Joyce Foundation

Panel Chair: Sharon Lynn Kagan, Teachers College, Columbia University and Yale University

The Task Force will develop recommendations for a state accountability system for pre-kindergarten programs and for linking such efforts to standards-based assessments efforts in kindergarten and the primary grades.

- **The National Academies Panel on Developmental Outcomes and Assessments for Young Children**

Supported by U.S. Department of Health and Human Services

Panel Chair: Catherine Snow, Harvard University.

The committee will focus on: (1) the identification of key outcomes associated with early stages of child development for children ages 0-5, and (2) the quality and purpose of different state-of-the art techniques and instruments for developmental assessments. In the first area, the committee will review the research base associated with developmental outcomes for children ages 0-5 in different domains, including physical, cognitive, social, psychobiological, and emotional. In the second area, the committee will examine the available range of techniques and instruments for assessing these outcomes, paying particular attention to the empirical evidence available about the reliability, validity, fairness and other considerations related to the quality and use of the developmental assessments. The committee will also examine the criteria that should guide the selection of assessment techniques for different purposes, such as guiding curriculum and instructional decisions for individual children, or program evaluation and program accountability, and the ability to link early childhood interventions such as Head Start with wider community goals for young children.

3 State Approaches: Maryland

- All entering kindergarten children assessed using shortened version of Work Sampling System assessment.
- Results published annually for state, county/school districts, subgroups of children, and types of early childhood programs.
- No reporting for individual local early childhood provider agencies, centers or classrooms.
- Strengths
 - Data on all children useful to early childhood system planning/advocacy and K-12 educators.
 - Assessment at kg. entry minimizes pressures for teacher assessors to inflate results.
 - Broad-based observational tool highlights significance of multiple domains.
 - Byproduct is sharpening kindergarten teachers' early awareness of children's abilities and differences.
- Limitations
 - One-time assessments don't provide feedback on children's progress in preschool or kg.
 - Reporting children's readiness without context of program quality/learning opportunities.
 - Risk of misusing data to retrospectively evaluate state or federal programs or local provider units without accounting for differences in populations served or program resources.
 - Need to ensure reliability/validity of shortened form of an assessment tool and method of calculating overall readiness scores.

3 State Approaches: New Jersey

- Ongoing external evaluation of representative samples of preschool classrooms (using ECERS-R, Support for Early Literacy Assessment and Preschool Classroom Mathematics Inventory) and children at kindergarten entry (assessing oral language, early literacy, mathematics); reporting for state prekindergarten program as a whole.
- All teachers use state-developed Early Learning Assessment System three times/year to assess oral language and literacy skills of all preschool children for use in instructional planning. To be expanded to include mathematics; results are not reported to the state.
- Local programs administer ECERS-R assessments annually to all preschool classrooms and use results for professional development. Programs also administer program self-assessment tool guided by state program quality standards; state validation visits are conducted in 1/3 of all programs to confirm ratings; results are reported to the state and used for program improvement.
- Strengths
 - Provides multiple forms of feedback on program quality and children's learning, including ongoing, credible child and program assessment data on state pre-k program as a whole.
 - Builds local capacity to understand and use combination of classroom and child assessment data.
- Limitations
 - Lacks capacity to answer questions on children's progress/readiness levels for local school districts or smaller community-based provider agencies.
 - Requires considerable state investment to support ongoing program improvement and external assessment efforts.

3 State Approaches: Florida

- All school districts will administer subsets of the Early Childhood Observation System (ECHOS) and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) to each entering kindergarten student.
- Reporting of “kindergarten readiness rate” for each public and private prekindergarten provider.
- Legislation mandates progressive sequence of consequences for lowest performing 15% of provider agencies (based on average readiness rates of children): technical assistance, new curricula and eventual defunding if outcomes do not improve.
- Strengths
 - Data on all children useful to early childhood system planning/advocacy and K-12 educators.
 - Assessment at kg. entry minimizes pressures for teacher assessors to inflate results.
 - Combination of assessment tools highlights significance of multiple domains.
- Limitations
 - Readiness index metric doesn’t reflect differences in children’s skills at entry to pre-k or amount of progress made during pre-k or kg.
 - Reporting children’s readiness without context of program quality/learning opportunities.
 - Concern that outcomes-based consequences could lead programs to selectively recruit children with higher levels of skills.