

2009 CCPRC Annual Meeting
Plenary Session 2
Thursday, October 29, 2009, 12:00 p.m. – 2:00 p.m.

High-Stakes Testing and Assessment: Quality and Children's Outcomes

Description

The context for this plenary was as follows: There is a push from many fronts to link program data with child outcomes. For example, Head Start mandates that child assessment data be collected as one indicator of program performance. In addition, the Office of Special Education requires all States to report child outcome data for early intervention and early childhood special education. States are taking very different approaches in terms of responding to these mandates, and extending those mandates to other programs such as public pre-kindergarten.

This plenary session highlighted some innovative approaches taken by States to address the challenges of including child assessments in accountability efforts. Also discussed were the challenges and pitfalls of such approaches, especially in coordinating assessments across sectors and populations. Finally, the implications for children, programs, and providers were considered.

Moderator

Toni Porter, Bank Street College of Education

Presenters

Martha Zaslow, Child Trends
Tamara Halle, Child Trends
Beth Rous, University of Kentucky
Jennifer Park, Florida Agency for Workforce Innovation

Scribe

Meagan McSwiggan, Child Trends

1. Documents in Session Folder

- “Key Themes from Early Childhood Assessment: Why, What and How,” Martha Zaslow and Tamara Halle
- “High Stakes Testing and Assessment: Quality and Children's Outcomes: Kentucky's Approach to Measuring Child Progress,” Beth Rous
- “High Stakes Testing and Assessment: Quality and Children's Outcomes: Kentucky Early Childhood School Readiness System,” Beth Rous
- “CCDF Policies Database: Categories of Policy Information Included in Database,” Urban Institute
- “Purpose of KEDS”
- “Florida's Child Assessment Initiatives,” Jennifer Park, Ph.D.

2. Summary of Presentations

- **Introduction: Toni Porter**

- We have convinced policy makers that we have something to say. So, what are the policy makers doing? They're beginning to look at child outcomes.

- **Summary of Presentation #1: Martha Zaslow**

- *Key Themes from Early Childhood Assessment: Why, What, and How--The Policy Background:*

- How to assess young children, what domains to assess, and appropriate implementation of assessments.
- The 2007 Head Start (HS) appropriation guided future actions regarding assessment of HS programs with regard to assessment expertise, special needs children, children learning two languages, etc.
- The consensus was that well-designed and implemented assessments have important benefits for children.
- Differing purposes of assessment:
 - Screening and diagnostic testing.
 - Provide initial suggestion about whether a child needs additional help.
 - Guiding instruction and intervention.
 - Get an understanding of what the child knows and can do, and track his/her progress over time.
 - Information is usually collected by ongoing observations.
 - Important to observe a whole set of children.
 - Reliability is very important.
 - Evaluation performance.
 - To determine if an intervention is working.
 - Use a sampling of children.
 - Deepen our knowledge and put results into context with other developmental areas.

- *Emerging issues:*

- There is a tendency to focus on particular domains of development in conducting child assessments.
- There are advantages and disadvantages for every assessment approach (observation-based, standardized assessments, etc.).
 - "If teachers and caregivers are not well trained or do not maintain reliability, the results are useless" – National Academy of Sciences.
- Caution is needed if one assessment approach is used for multiple purposes- important to continue ongoing discussions about this issue.
- It is inappropriate to use assessment data in isolation to make decisions about programs. It's important to measure a child's progress rather than an end-of-year status.
- The purpose of assessment may be ambiguous or may not be communicated explicitly.
- Reliability and validity need to be established for specific purposes and populations.

- *Assessment systems:*
 - Systems of assessment are organized around specific goals. The components need to work together or the system will function poorly.
 - Components of assessment systems:
 - Standards for children’s learning.
 - Program quality standards.
 - Monitoring and evaluation.
 - Professional development.
 - Database.
 - Adequate resources.
- *What are the key pitfalls regarding assessment systems?*
 - Lack of alignment among system components.
 - Need for resources to assure that information is both useful and used. We need resources for ongoing training and reliability. Reports need to be prepared in a timely manner.
- **Summary of Presentation #2: Martha Zaslow, on behalf of Tamara Hall**
 - A document is being prepared that will summarize information to help programs pick appropriate screeners and assessments.
 - Funding is coming from OPRE and profiles of measures are being completed by a team at Child Trends.
 - Measures will be selected based on Head Start data.
 - Key issue is translating technical information. Each profile has background information, availability, cost, etc.
 - A definition section of psychometric terminology and crosscutting summary tables will be included.
 - Example of translating psychometric information:
 - Instead of asking, “Does the measure have high inter-rater reliability?” we ask, “Do different raters agree when they are assessing the same children?”
- **Summary of Presentation #3: Beth Rous**
 - *How is Kentucky (KY) looking at school readiness?*
 - KY has worked to align the professional development and child assessment system.
 - Work began in 1998 with guidelines on accountability systems in preschool settings through a U. S. Department of Education (USDOE)-funded study. Key findings included nomenclature differences between K-12 settings and preschool settings.
 - Why develop an assessment system?
 - There was an increase in interest in early childhood in Kentucky. The question became “Where are these millions of dollars going?”
 - *Issues:*
 - There was no common set of assessment tools that States and programs used. Different types of assessments were being implemented (naturalistic, standardized, etc.).

- No assessment is cheap but costs relate to the stakes involved (the lower the stakes, the lower the costs can be).
- Some disagreement in the field about assessments.
- Some were still saying that assessing children is bad.
- Is it better to do naturalistic or standardized assessments?
- Should we use teacher reported data?
- Can we use one assessment tool for multiple purposes?
- *Considerations about assessments:*
 - No assessment is perfect.
 - Some domains have more assessment options.
 - Still working to establish predictive validity and reliability with various populations.
- *Kentucky's principles:*
 - Assessments should bring about benefits, be tailored to purpose, be age appropriate in content and method, reliable and valid, and linguistically appropriate.
- *The Kentucky approach:*
 - We reviewed our child standards. We looked at assessments currently being used and linked our child standards to tools that we thought would be adequate.
 - We developed a continuous assessment guide. There were three distinctions: screening, diagnostic, and classroom/instructional assessment that were part of the continuous assessment process.
 - We then developed a data system with some funding support through the USDOE Office of Special Education Programs (OSEP).
 - A list of approved assessment tools is included in the PowerPoint presentation (linked to this summary).
 - Why the Kentucky Early Childhood Data System (KEDS)?
 - Support improved instruction.
 - Document child outcomes.
 - Align individual assessment items with State child level benchmarks and standards. Benchmarks are then aligned with OSEP child outcomes. Head Start child outcomes were also linked to State benchmarks and standards.
 - Underlying principle is that if outcomes change or are added for specific populations, they can be linked to the benchmarks so no change in assessment processes or reporting is needed at the local level.
 - Advantages: We have a rich dataset, inclusive standards, a normed sample specific to Kentucky, which eliminates work for teachers, allows choices, and reduces the burden on children
 - Challenges: Decision-making ability of tools, ensuring reliability and validity, assessing outliers, assessing children with severe disabilities, changing leadership across State programs, and requests to use outcome data for other purposes, which may not be appropriate.
 - Currently 17,000 children are in the system this year.
 - We report summaries by OSEP outcomes (currently) and KY Early Childhood Standards (in 2011).

- **Summary of Presentation #4: Jennifer Park**
 - *Florida's early learning structure:*
 - Department of Health is in charge of health and safety.
 - Department of Children and Families is in charge of licensing and provider qualifications verification.
 - Department of Education—Voluntary Prekindergarten (VPK) standards, curriculum, and assessment.
 - Agency for Workforce Innovation—School readiness, VPK operations and funding.
 - *Complicated web of relationships:*
 - We have multiple partnerships that start at our agency, and we have community partners, local child care licensing agencies, school boards, etc.
 - We were given stimulus dollars and have been investing in statewide supports. The coalitions at the local level have asked if we would step in and fulfill this role.
 - *American Recovery and Reinvestment Act of 2009 (ARRA) early learning projects:*
 - We have achieved consensus with all 31 early learning coalitions to retain a portion of the ARRA quality dollars at the State level for statewide investment in standardization in three areas.
 - How do we determine program success? We measure child gains.
 - How do we develop quality-learning environments? We measure environmental quality.
 - How do we support successful program delivery? We have competency based expectations.
 - *Creating a strong framework grounded in evidence-based research:*
 - The first thing our providers need to do is get to know the children. This is the foundation.
 - This foundation provides a basis for understanding children's progress as demonstrated through data.
 - This data can be used for program planning.
 - *Program environment:*
 - Ultimate goal is to inform program planning.
 - We've made the decision to invest in a professional development registry so we can track where a teacher is on a professional development trajectory.
 - We then put these pieces together to create a strong framework. Each leg of the stool is critical. The stool is not sustainable, if you take one leg away. Also, it's impossible to talk about one without referencing the others.
 - *Developmental screening:*
 - Determining the need for further evaluation at the earliest identification.
 - Using a single instrument for statewide developmental screening.
 - Defining consistent procedures for screening, referral, and follow-up.
 - *Child assessment:*
 - Statewide single measure? Yes and no. The coalitions will continue to assess as they choose, but we will support them as it relates to ongoing child assessment. We are moving toward procurement of a single tool to be administered by a third

party assessor. This assessment approach will involve only a sample of children and data used in the aggregate only for state-level use only.

- *Early learning coalitions:*
 - The early learning coalitions are implementing everything. We need to be thoughtful about how we release all training opportunities. They are the ones that partner with local partners to provide trainings. This is another challenge.
- *Alignment:*
 - Alignment is a big issue. We have early learning and developmental standards, but authority is lacking to align them with screening and assessment data.
 - In the context of child assessment, we want to look and see what teacher practices exist. We'll be able to take child assessment data, teacher-child interaction data, etc., and guide programs based on these results.
- *Populations:*
 - Almost 400,000 children were served last year. We have special populations that we want to make sure are appropriately assessed, such as English Language Learners, children from migrant families, and children with disabilities. We don't want to exclude any groups from testing but rather assess appropriately.
- *Summary:*
 - The point of this work is to be able to provide a referral for a child. To take a given child and ask, "How can we track his or her development based on teacher competency, classroom environment, etc.?"

3. Summary of Discussion with Presenters and Participants

- How these systems relate to what other partners are doing, and what the expectations are for those other systems was discussed. There was agreement that it's a complicated process. In Florida, for example, work is underway with the Department of Education to develop one unified guidance document for assessing all children birth to 5, as well as developing assessment plans, and collecting data that will inform the process of determining an appropriate way to assess how successful a program has been.
- A question was asked about whether there are particular challenges in accessing children for program and assessment purposes. Over five years, Kentucky intends to phase-in public pre-k programs. It also received a grant to start a pilot child care program. Kentucky doesn't have pure programs anymore; programs are mixed (pre-k, Head Start, etc). Kentucky is working to incorporate children in early intervention (they're currently in different settings). It also has a home visiting program called Hands and hopes to bring in this group of children.
- Kentucky has a student identifier system; recently, it has been agreed to assign those student identifiers to each child in the early intervention group.
- Discussion of the need for training for recipients of information.
 - Kentucky has been struggling with this. It does very little to prepare State administrators. It was mentioned that one weak link could really cause great harm in the system. It's hard and it's an important issue for the field to address.
 - In Florida, it's easy to point out this challenge to the Department of Education. One lesson learned is that decision-making needs to center on supporting early learning programs as opposed to getting hung up about how a program is funded, where it's housed, etc.

- Discussion of how the system of child assessment fits into the bigger picture of Quality Rating and Improvement Systems (QRIS). The question was asked if this is the prelude to developing a QRIS.
 - Jennifer Park said they hadn't planned for this yet, but they want to develop strong components to put them in a good position to go in that direction.
 - Beth Rous said they are collecting data in terms of different programs, hoping to build the system so they have a repository where they can connect data.

4. Key Themes and Issues

- Understand the purpose of using assessments—and understand that they are part of a larger system.
- Figure out how you'll use the assessments and train those involved so that they know what they're doing, what the purpose is, etc.
- Be sure to clarify how you'll use the information you gather from assessments. How will you provide information to policy makers, parents, and practitioners so that each can understand what the results mean for them?