Purposeful Early Childhood Assessment

Martha Zaslow & Tamara Halle
Child Trends
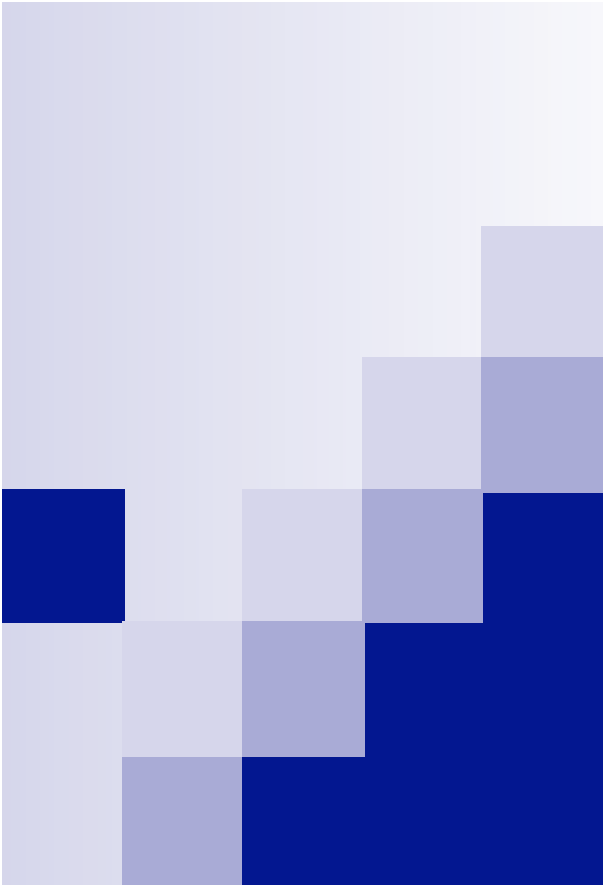
Presentation at the Child Care Policy Research Consortium
Meeting
Washington, DC
October 29, 2009

Goals of this Presentation

Two parts of this presentation:

☐ Part 1 will provide a brief overview of key themes in the recent report of the National Academies of Science *Committee on Developmental Outcomes and Assessments for Young Children*

☐ Part II will provide an update on a project sponsored by the Office of Planning Research and Evaluation, Administration for Children and Families, DHHS with input from the Office of Head Start, aimed at informing programs' selection of assessment instruments

Key Themes from *Early Childhood Assessment: Why, What and How*

Background on *Early Childhood Assessment: Why, What and How*

Policy Background

- Increasing emphasis on accountability, results, measurement in federal and state programs
- Efforts by Head Start and other early intervention programs to address these issues
- Multiple assessment and evaluation efforts, including many state pre-K programs

Background

Concerns motivating study

- Whether young children should be assessed at all
- How to assess them
- What domains to assess
- Appropriate implementation of assessments
- Appropriate uses of assessment results

Background

Congressional Mandate

- In FY 2007 HHS appropriation.
- Study intended to
  - guide future congressional actions regarding assessment of Head Start programs
  - inform users of early childhood assessments
- Sponsored by DHHS Office of Head Start, with OPRE involvement.

Background

Statement of Task

- Identify key outcomes associated with early stages of child development (ages 0-5).

- Review and evaluate quality and purpose of current techniques and instruments for developmental assessments of young children.

- Prepare report with recommendations for assessment policy and practice as well as future research priorities.

Background

Committee Areas of Expertise

- Infant and child development
- Early education, Head Start, and childcare programs
- Clinical psychology
- Pediatrics, infant and child health
- Children with disabilities and special needs
- English language learners and cultural/language minorities
- Assessment, psychometrics, statistics, research methodology
- Policy

Background

Committee on Developmental Outcomes and Assessments for Young Children

Catherine Snow (chair)	Harvard Graduate School of Education
Margaret Burchinal	University of California, Irvine
Harriet A. Egertson	Independent Consultant
Eugene K. Emory	Emory University
David J. Francis	University of Houston
Eugene E. García	Arizona State University
Kathleen Hebbeler	SRI International
Eboni Howard	Erikson Institute
Jacqueline Jones	New Jersey Department of Education
Luis M. Laosa	Educational Testing Service
Kathleen McCartney	Harvard Graduate School of Education
Marie C. McCormick	Harvard School of Public Health
Deborah Stipek	Stanford University
Mark R. Wilson	University of California, Berkeley
Martha Zaslow	Child Trends

BOCYF Liaison
Betsy Lozoff   University of Michigan

NRC Staff
Susan Van Hemel, Matt McDonough
**Rosemary Chalk, Stuart Elliott**

Background

Consensus of Committee

- Assessments of young children, well designed, implemented, and used, have important benefits for children and the programs that serve them.
- Flawed assessments and flawed use of assessment findings can result in harm to children and programs.
- Our task is an opportunity to maximize the benefits and minimize the harm of early childhood assessment.

Key Themes in *Early Childhood Assessment: Why, What and How*

Purposes: There are fundamentally different purposes for conducting early childhood assessments.

☐ The purpose of the assessment is critical to how the measure is designed, which measure is selected, how it is implemented and how results are reported and used
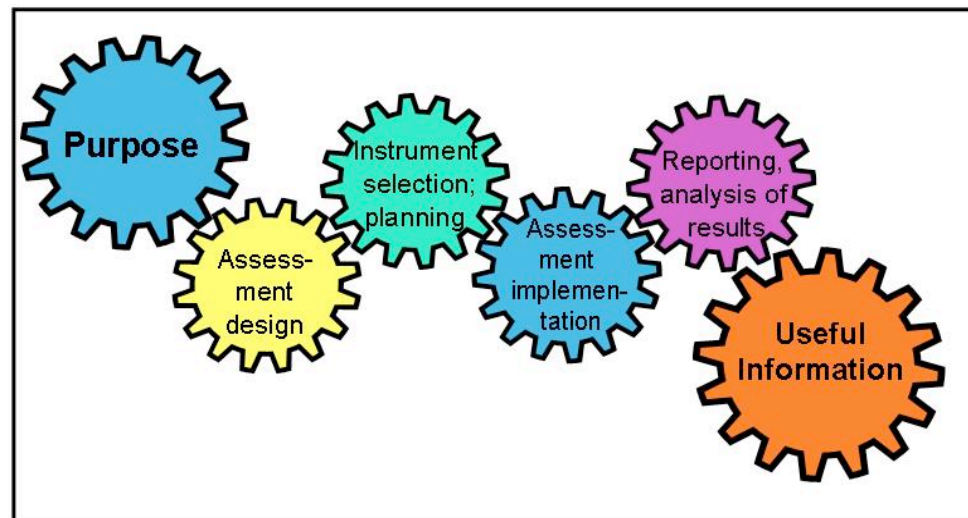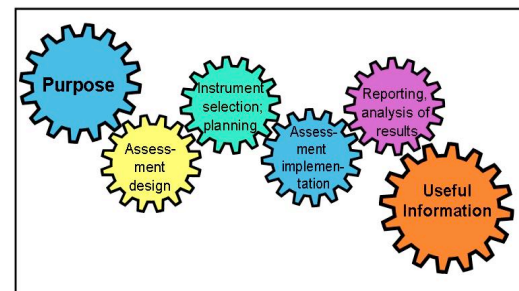
Key Themes

Systems: Conducting the assessments is only one part of system with multiple components that together can inform progress towards high quality ECE.

Key Themes

These are linked. The purpose will shape how the early childhood assessment is integrated with other components of the system.

Differing Purposes of Assessment

- Screening and diagnostic testing
- Guiding instruction and intervention
- Evaluating the performance of a program or society
- Advancing knowledge of child development

Differing Purposes of Assessment

Screening and diagnostic testing

- *What is the goal of assessment?* To provide an initial assessment of whether child appears to be progressing as expected or may need special supports; to follow-up by determining whether a particular concern is present.
- *Who is assessed*: Screening and diagnosis can be informative for individual children, children in an early childhood program, or at community level.
- *Who is the target of assessment information?* Family, medical professional, early childhood program staff, early childhood intervention staff.
- *How is the assessment information collected?* May be medical assessment, parent report or one-on-one assessment using instrument designed to provide screening or diagnostic information. On screeners, thresholds indicate whether there *may* be a physical or developmental issue that should be followed up with further in-depth diagnostic work

Differing Purposes of Assessment

Guiding instruction and intervention

- *What is the goal of assessment?* To get picture of what children know and can do, and to track their progress over time, to inform decisions about instruction at level of individual or group

- *Who is assessed? All* children in a class, center or program need to be assessed

- *Who is the target of assessment information:* Primarily *within* an early childhood setting: caregivers/educators and directors; family.

- *How is assessment information collected?:* Usually by observing children, collecting samples of their work, and talking to children. Progress is related to learning or behavioral criteria (criterion referenced) or to progress on curricular goals (curriculum referenced)

Differing Purposes of Assessment

Evaluating the performance of a program or society

- *What is the goal of the assessment:* To determine if a program is effective; to determine if a public investment has had desired results; to document and track well-being of children in a community, region or state to identify strengths/weaknesses to better target interventions and to monitor children's wellbeing

- *Who is assessed:* It is *not* necessary to assess every child. An appropriate sample is sufficient. Goal is not to provide input on individual children's instruction but to assess program as a whole.

- *Who is target of information:* Funders, policymakers

- *How is assessment information collected:* Often through standardized assessments to hold constant timing, content and procedures of data collection

Differing Purposes of Assessment

Advancing knowledge of child development

- *What is the goal of the assessment:* Deepen knowledge of particular domains of development with eventual aim of strengthening related aspects of professional development, curricula, and environmental quality

- *Who is assessed:* Sample appropriate for research question. Informed consent required because data are used for "generalizable knowledge"

- *Who is target of assessment information?* Researchers and those in practice community

- *How is assessment information collected?* Standardized assessments of development or new measures under development. Usually in coordination with measures of early childhood professional development and/or quality of the environment.

Differing Purposes of Assessment

- Purposes drive decisions about:
  - domains to be measured;
  - tools to be used;
  - implementation;
  - analysis, interpretation, reporting of results.

Issue: There is a tendency to focus on particular domains of development in conducting child assessments.

- Committee identified five important domains of development:
  - Physical health and wellbeing
  - Social-emotional development
  - Approaches to learning
  - Language and literacy
  - Cognitive skills (with math as an example)
- Purpose of assessment determines the domain or domains assessed.
- However Committee cautioned that where possible:

  "Domains included when assessing child outcomes and quality of educational programs should be expanded beyond those traditionally emphasized (language, literacy, and mathematics) to include others, such as affect, interpersonal interaction, and opportunities for self-expression" (p. 348)

- Noted need for measures development in these areas

Issue: There are advantages and disadvantages for each assessment approach.

Observation-based:
- Need for training to obtain reliable results
- Need for well-defined rubrics and scoring guides
- Need for ongoing monitoring
- Not all teachers/caregivers are good assessors
- Potential for bias if teachers/ caregivers perceive consequences

Examples are from Mathematica (2007)

Standardized assessments:
- Format may be unfamiliar to child; may not stay engaged
- May be contrary to cultural norms to respond to direct questioning
- Weighing alternatives may be unfamiliar
- Children may be more comfortable showing than telling what they know
- Approach may be more familiar to children in some programs than others

Issue: Caution is needed if one assessment approach is used for multiple purposes

- Burden to programs, costs of assessments, challenges of implementing standardized assessments with young children result in pressure/decisions to use assessments for multiple purposes
- We are seeing this especially with respect to use of data from ongoing observational assessments, originally developed to inform instruction, for accountability and social benchmarking
- Committee concluded that it is acceptable to use one measurement approach for more than one purpose.
- However caution is indicated. For example, when observation-based measures are used for accountability:
  - Need for standards with respect to initial and ongoing training (booster; follow-up as well initial training)
  - Need for checks on reliability

Issue: It is inappropriate to use assessment data in isolation to make decisions about programs

"It is not uncommon that inferences about program effectiveness are based on end-of-program performance of individual children. Such inferences are inappropriate without attention to the environment children experience both inside and outside the program, as well as the characteristics at entry of the children served by the program. " (p. 344)

Committee concluded that child assessment data should not be used in isolation. It is important to:

- Measure child progress rather than end of year status
- Collect direct indicators of program quality
- Collect information on risk status of families and children
- Collect information on program resources (e.g., funding, administrative support, professional development)
- Have a clear plan for program improvement

Issue: The purpose of assessment may be ambiguous or may not be communicated explicitly.

Committee concluded that it is an important responsibility to communicate the purpose of assessment clearly.

"There is a responsibility to articulate the purpose of any assessment…to those who participate and who might be influenced by outcomes. For example, if a program is being evaluated, program staff should understand whether there are plans to use the assessments to evaluate their performance on an individual level. They should also know whether the information will be made available to guide decisions about the program and individual children" (p. 345).

Assessors, teachers and administrators should be able to articulate the purpose of assessment to families (p. 353)

Issue: Reliability and validity need to be established for the particular purpose and population

- Selecting an assessment tool includes confirming that it has acceptable reliability and validity *for the specific purpose it will be used for.*
- Reliability and validity are also need to be acceptable *for the population to be assessed*
- "Program directors, policy makers, and others who select instruments for assessments should receive instruction in how to select and use assessment instruments"

Assessment Systems

Systems:

- Are organized around specific goals
- Have components each with own goal
- Components must work well tether for larger system to work
- Missing or poorly operating components can cause system to function poorly

Early childhood assessments should not be conducted in isolation, but should be part of an assessment system with the goal of providing information to guide progress towards high quality early care and education.

Assessment Systems

Components:

- **Standards for children's learning** that serve as the source of constructs focused on in assessments and performance levels considered acceptable
- **Program quality standards** that are aligned with standards for children's learning
- **Attention to inclusiveness**, respecting variation in children's culture and developmental challenges.
- **Assessments** *both* of program quality and children's development

Assessment as Part of a System

- **Professional development** on using and understanding the assessments and assessment reports
- **Data base** that provides for integrity of data and provides basis for reports to varied audiences according to purpose
- **Resources** to provide timely useful, high-quality assessment data
- **Monitoring and evaluation** of the system, including alignment of elements, monitoring of the effectiveness of the reporting system, monitoring effects of the system, and burden on teachers, children and administrators.

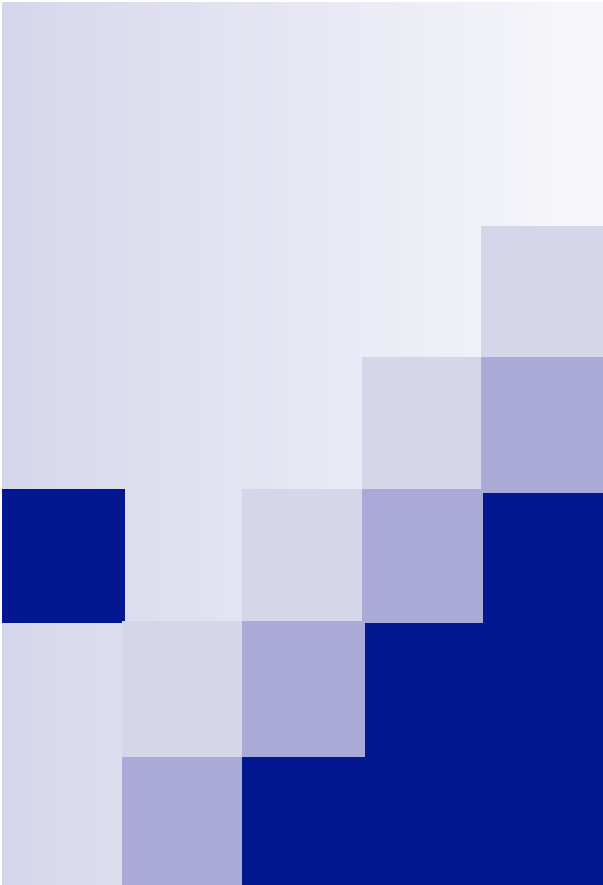# Issue: Lack of alignment of components of system

- Poor coordination in focus of
  - Early learning standards and program standards
  - Early learning standards and child assessments
  - Program standards and quality measures
- Lack of joint consideration of program quality and child assessments in providing input into program improvement

Issue: Need for resources to assure useful information and use of information

Need resources to assure:

- Ongoing training to assure that child assessment and program quality data are collected with reliability
- Timely preparation of reports
- Training so that those receiving assessment reports can interpret and make use of them
- *That those selecting assessments make informed decisions*

Child Outcome Screening and Assessment Instruments Used by

Head Start Programs

Project Rationale

- Head Start legislation requires programs to use assessment measures that are:
  - ☐ Developmentally, linguistically, and culturally appropriate
  - ☐ Reliable and valid in the language in which they are used
- This document will summarize information to help programs pick appropriate assessments and screeners
- While helping to address a requirement in Head Start programs, the information provided will be useful to all types of early childhood programs

Project Team

- Funding is from Office of Planning Research and Evaluation (OPRE), Administration for Children and Families, US Department of Health and Human Services
- Input into the project is being provided by OPRE and the Office of Head Start
- Project does not necessarily reflect the views or opinions of the Administration for Children and Families of the US Department of Health and Human Services
- Profiles of measures are being completed by a team at Child Trends.

Objectives of Project

- Highlight evidence on reliability and validity of measures for key populations

- Present technical information about measures in a non-technical manner

- Identify areas where more work is needed to develop assessments and screeners for key populations

Selection of Measures

- Include profiles of instruments that are most widely used by local programs for:
  - Individualization, classroom planning, and program improvement (including professional development) efforts. These are primarily observation-based measures that may or may not be curriculum-based
  - Screening children to identify those for whom further follow-up diagnostic evaluation for developmental problems is warranted.
- Measures were chosen based on data from:
  - Head Start Program Information Reports (PIR) data
  - Head Start Family and Child Experiences Survey (FACES) data

Selection of Measures

Assessments

- Creative Curriculum Assessment Toolkit
- Desired Results Developmental Profile
- Learning Accomplishment Profile-Diagnostic (English and Spanish versions)
- Learning Accomplishment Profile-3rd Edition
- Preschool Galileo Scales
- High Scope Child Observation Record
- Work Sampling System for Head Start
- Mullen Scales of Early Learning

Screeners

- Child Development Inventory
- Denver II
- Ages and Stages Questionnaire/ Ages and Stages Social-Emotional
- Battelle Developmental Inventory (English and Spanish versions)
- Brigance Preschool Screen
- Developmental Indicators for the Assessment of Learning- 3rd Edition (DIAL-3) (English and Spanish versions)
- Early Screening Inventory
- Learning Accomplishment Profile-Diagnostic Screens
- Parents Evaluation of Developmental Status

Translating Technical Information

- Create a profile for each measure
- The profile is a comprehensive summary of the measure including:
    - Background information
    - Availability and cost
    - Training requirements
    - Information reporting systems
    - Approaches to family/parent input
    - Appropriateness for children of different backgrounds
    - Reliability, validity, sensitivity and specificity information for key populations

Translating Technical Information

- Document will also contain background information including
  - A definitions section with accessible definitions of psychometric terminology
  - Cross-cutting summary tables providing a brief overview of measures on content, reliability and validity, and populations the measure has been used with

Translating Technical Information

- Eliminated technical psychometric terms from profile, for example:
  - Instead of asking, "Does the measure have high inter-rater reliability?"
  - We ask, "Do different raters agree when they are assessing the same children?"