Toward a Unified Vision of Assessment for Young Children

Kathleen Hebbeler
SRI International

Listening and Learning about Early Learning Tour
May 11, 2010
What is assessment?

“Assessment is a generic term that refers to the process of gathering information for decision-making.”
(McLean, Wolery, and Bailey, 2004)
Starting point

- Good early childhood assessment is good early childhood assessment
- Principles that apply to typically developing children apply to children with disabilities
  - See NAEYC position statement on assessment
Uses of assessment with young children with disabilities

- Screening
- Eligibility determination
- Intervention/Instructional planning and monitoring progress
- Measurement of outcomes (Accountability, program improvement, program evaluation)
Screening

- **Purpose:** Identify children who may have developmental problems and who need additional assessment
- **Tools** have been developed for this purpose
- **Possible uses**
  - General screening (e.g. pediatricians)
  - As part of program entry process to limit inappropriate referrals
Eligibility Determination

• Most of the children being served in early intervention (0-3) and early childhood special education (3-5) do not have a medical diagnosis.

• Assessment is key to program entry because assessment results determine the eligible population.
Importance of Assessment to Identification for Special Services

• Early Intervention
  – Serves 2.5% of population of 0 - 3 year olds
  – 64% eligible because of developmental delay

• Early Childhood Special Education
  – Serves 5.7% of 3 - 5 year olds
  – 84% have primary disability of speech and language (46%) or developmental delay (38%)
Eligibility Determination

- Eligibility is determined by
  - percent delay (e.g., 50% delay in one area)
  - standard deviation (e.g., 1.5 SDs below the mean in one area)
- Programs tend to use standardized, norm-referenced tools
  - Need to answer the question: “Does this child have a delay?”
Intervention/Instructional Planning

• Tools are curriculum or criterion referenced
• Assessment content provides a framework for working with children
• Provide a way to assess growth
• Many tools used in programs with typically developing children work well for many children with disabilities and delays
  • E.g., Creative Curriculum, Work Sampling, High Scope
Wasted Resources in Current Practice

• One or more assessments administered to establish eligibility
  – Provide limited to no information that is useful for working with the child.
  – Get the child into services

• Different assessment(s) administered to plan and monitor change over time
  – Useful information for providers
  – Are not age anchored
...but what we want to know is not all that different.

**Norm-referenced or criterion-referenced?**

1. The child identifies the sources of common action.
2. The child notices when one object in a series of objects is out of place.
3. The child answers simple logic questions.
4. The child understands and uses positional words correctly.
Assessment and Tiered Program Models

- Response to intervention (RTI), Recognition and Response, Pyramid Models
- Use a different type of assessment to screen for children requiring more intensive intervention (different tier) and monitor progress
- More research needed on applying RTI to young children
  - especially the blending of assessment for tiers and curriculum based assessment
Measurement of Outcomes (Aggregated)

- All States are reporting on outcomes for children receiving EI and ECSE services to U.S. Department of Education.
  - All States are using tools designed for eligibility or intervention/instructional planning.
  - Most using information collected for other purposes to report on outcomes.
- Children with disabilities also attend inclusive programs and participating in other accountability systems and program evaluations.
Interesting dilemma

- Observation-based assessment widely regarded as the better way to assess young children.
- Many large scale assessment efforts (especially program evaluations) use standardized measures.
  - Few notable exceptions: statewide efforts in KY, CO, NE, PA.

**Standardized assessments pose far more problems for assessing children with disabilities**
Large scale assessment

- Exclusion of children with disabilities is not an acceptable option.
- All children and their families are entitled to know if the program works for them, if they are making progress, etc.
Construct Irrelevant Variance

- Child has the concept but does not get credit for the item because
  - Can’t point
  - Can’t speak
  - Can’t attend for even short periods of time
  - Can’t understand the instructions, etc.

- Major problem with standardized assessments for children with disabilities because of the standard administration.

  **Standardizing the conditions does not standardize the experience for the child.**
How to Gather Information on Young Children

• Recommended practice is through observation, child’s products

• Allow multiple ways for a child to demonstrate competencies
  – Very important for children with disabilities

• Requires well defined criteria for the scoring the items

• Requires training of teachers and other providers in the application of the criteria.
Universal Design and Accommodations

- Develop assessments to allow the widest range of participation; minimize the need for accommodations
  - E.g., refer to “communication,” not “spoken language”
- Accommodations – modifications in presentation, response format, timing, setting (Some of which assessors do in EC anyway)
  - Validity of accommodations
- Example:
  Desired Results Developmental Profile – Access
Floor effects and sensitivity

- Floor effects – not enough (or even any) items for children with slower skill development
  - E.g., assessment is for 3 - 5s – developmentally the child is 2
- Lack of sensitivity – increments between items are too large to capture growth of children who progress slowly
The Importance of Developmental Trajectories
Vocabulary Growth (Hart & Risley, 1995)
The Importance of Developmental Trajectories
The Importance of Developmental Trajectories
What we have learned recently through large scale assessments

- There are children in programs for children from low income families whose skills are lower than children receiving early intervention and early childhood special education.
  - The difference between these two groups of children is not so clear cut.
  - Early Head Start, FACES, State data
  - Many entering kindergartens in CA’s low performing schools had skills in the 3 - 4 year old range.
Why do we intervene?
Age Anchored Vs. Norm Referenced

• Norm-referenced
  – Associated with standardized assessment

• Age Anchoring
  – Based on research on learning and development.
  – Essential knowledge for all who work with young children
    • Is this child delayed?
    • Is this child making sufficient progress?
  – The (potential) link between early learning guidelines and assessment.
  – Accounts for individual variation AND identification of children whose skill acquisition is problematic.
Current assessments

Norm referenced
- Designed to provide information on where a child is developmentally compared to age peers
- Don’t provide useful information for planning
- Structured administration not consistent with recommended practices, problematic for children with disabilities

Criterion referenced
- Consistent with best practice
- Provide meaningful information for teachers, providers
- Criteria for scoring are not well defined (i.e., potential problems for reliability).
- Don’t provide information related to age anchoring or benchmarking
The need for assessments that can address multiple purposes

• If we can agree on …
  – A set of behaviors that are worth measuring
  – What constitutes evidence that the child does or does not display that behavior

• And conduct research on the ages at which children display these behaviors, then...

• We can develop assessments that serve multiple purposes.
If we don’t conduct meaningful age anchored assessment…

- We don’t know the child is behind.
- We don’t know how far behind.
- We don’t know if what we are doing with the child is making a difference in moving the child to a higher trajectory.
Recommendations

The federal government needs to support research and development on assessment tools for young children that can be used for multiple purposes such as:

– identifying children who are acquiring skills too slowly,
– planning for intervention and instruction,
– monitoring progress in meaningful outcomes,
– guiding ongoing program improvement, and
– documenting program effectiveness.
Content Specifications (Wish list)

• Measure multiple areas of development and learning that are important for children in the present and provide the foundation for important future outcomes.
  – Not just predictive validity or downward mapping from K-12
  – Starting points: Early learning guidelines, OSEP outcomes.

• Based on well researched developmental sequences and learning progressions.
  – May need research to establish some of these
  – Provide the basis for good instruction/intervention
Design Specifications (Wish list)

- Assess development and learning continuously between birth and age 8/Grade 3.
  - needed for trajectory tracking (longitudinal growth)
  - needed for children with slower skill acquisition
- Use naturally occurring behaviors (including products) from the child’s daily activities as evidence
  - multiple ways to demonstrate behavior or skill
- Provide detailed behaviorally-based scoring criteria for each item
  - Promote good reliability
Design Specifications (Wish list)

• Includes multiple methods for documentation
  – Observation notes, work samples, photos, video clips, etc.
• Include age-anchored descriptions or benchmarks across the full age range
  – promote knowledge of typical development and early learning guidelines
  – allow for the identification of children who are progressing so slowly that more intensive intervention is required.
• Can be administered by teachers and providers
  – Reasonable number of items
  – Not difficult or complicated
Design Specifications (Wish list)

• Allow for input from multiple team members working with child
• Provide ways to engage families in the assessment process
  – source of information on child behavior
  – informer user of the results
• Include information from multiple settings
• Incorporate principles of universal design; minimize the need for accommodations.
• Provide validated accommodations when necessary.
Design Specifications (Wish list)

• Have sufficient sensitivity at each age level to document status and progress of all children
  – Adaptive assessment? not all items (or levels) for all children
• Appropriate for children from all cultures and/or who are learning English.
• Have well developed training materials.
• Reflect the current science of assessment development.
• Provide for web-based data entry and immediate access to results and multiple reports for different users.
Conclusion

• Good assessment is an essential component of high quality programs for all young children, including children with disabilities.

• Investing in better tools is an investment in improving program quality and an important step to producing better outcomes for young children.