Appendix G

Best Practices in Early Childhood Assessment

The early childhood years are commonly held to include children ages birth through age eight. For the purposes of this document, the age range addressed is limited to children served in public schools, specifically children ages three through eight. While the purposes of assessment in early childhood classrooms and schools are the same as those for older children, the design of a comprehensive assessment system in early childhood is necessarily different because of the age of the child. Assessment of children in this age range is significantly impacted by the nature of the young child. Early childhood development and learning is rapid, episodic, and highly influenced by experience. In the preschool and early elementary years, rates of development in all areas outpace growth rates at any other time and because children develop and learn so rapidly, assessments given at one point in time might not give a complete picture of a child's abilities. Additionally, young children have uneven access to environmental supports prior to formal schooling. The young child's performance on assessment tasks is highly influenced by their emotional states and the conditions of the assessment. Young children are not consistent in demonstrating their abilities. Additionally, the younger the child, the less likely they are to be familiar with the goals of formal testing and the less likely they are to understand the need to perform well during assessments. It is more difficult to use assessment methods which require sustained, focused attention and cooperation with the examiner. Young children are better able to demonstrate their abilities, than to talk or write as a means of showing what they know. For these reasons, careful attention must be paid to the design of the early childhood assessment system and to the accuracy of the conclusions which can be drawn from the assessment information.

Guiding Principles

Assessment of young children should:

- Focus on goals which are developmentally and educationally important.
- Be aligned with learning standards.
- Include teacher observations, student work, checklists and rating scales completed by teachers and parents, criterion-referenced tests, curriculum-based measures, and norm-referenced tests.
- Rely on instruments selected by qualified professionals for reliability, validity, and appropriateness (e.g., include manipulatives vs. abstract pen/pencil tasks).
- Address all domains of learning, not just cognitive domains of literacy and mathematics.
- Be systematically obtained over time using repeated measures and using a variety of methods and sources in each domain.
- Rely on demonstrated performance during real, not contrived, activities.
- Not threaten children's psychological safety and self-esteem and be sensitive to children's motivation, interest, and attention span.
- Provide a clear benefit for children either in the services they receive or in the quality of their educational program.

Authentic Assessment

Authentic assessment generally results in the most valid information about what children know and are able to do. However, authentic assessment is often seen as time and cost intensive due to data collection, coding and entry, and data analysis requirements. Authentic assessment information needs to come from a variety of methods, including child observation, work samples, child interviews, and information gathered from a variety of sources, including parents and other relevant adults. Ongoing teacher observations of children have proven effective at shaping instruction to meet children's rapidly changing learning needs. However, these observations of children go beyond anecdotal notes and instead are used to complete developmental scales of proven reliability and validity. Examples of student work provide meaningful evidence of learning and development as long as the examples are aligned with learning goals and instruction.

Authentic assessment information is:

- Systematically obtained over time, across contexts, through multiple sources and methods.
- Generated using multiple methods for children to demonstrate what they know and can do – this is especially beneficial for children with disabilities.
- Collected in all domains of development and learning.
- Conducted in the natural environment as part of the child's daily experience – real knowledge measured in the context of real activities which are meaningful to children.
- Conducted in an ongoing manner, but should include more formal progress assessments at least twice a year.

Standardized Assessments

Standardized assessments, when administered, appropriately allow for fair comparisons among individual children and groups of children. They are considered objective, and both time and cost efficient. However, assessment experts advise that caution should be used when interpreting the standardized assessment results of young children's learning. There are a variety of issues which need to be taken into account when using standardized assessments as a part of a comprehensive early childhood assessment system. In general, obtaining valid scores on standardized assessments with children younger than age 8 is challenging because children may not understand the need to do well when tested, are inconsistent in their ability to demonstrate what they know and can do, and are easily influenced by their emotional states and testing conditions. In general, the long-term predicative validity of standardized assessments for children under the age of eight is not high. Additionally, adequate instruments do not exist to test in all domains or learning and development. Available tests primarily cover discrete components of language development, literacy, and mathematics. When standardized assessments are used, they should measure developmentally and educationally significant items and be aligned with early learning standards and program goals.

Because standardized instruments are so fallible, it is important that the measures selected meet rigorous standards of reliability and validity. Additionally, they must be administered and interpreted by trained professionals and scores should be interpreted within a broader assessment which includes information gathered from a variety of sources. Standardized assessments for young children must include enough items to

allow scores to represent a wide range of abilities and be sensitive enough to represent minor differences in skills. Assessments should be used for their intended purpose with their intended population and should be reliable, valid, and fair for that purpose; including culturally and linguistically appropriate. To some extent all standardized assessments are a measure of language, so it is critical that assessments be linguistically appropriate and that first and second language development are taken into account when selecting standardized assessment measures and interpreting the results. Lastly, standardized assessments must be administered in environments which correspond to the testing manual's specifications – usually controlled, relatively quiet areas with no distractions.

Conducting Early Childhood Assessment

Implementing comprehensive systems of early childhood assessment requires a substantial investment in training and professional development of teachers and assessors and ongoing quality checks. For effective child assessment, staff need to be educated about assessment principles and understand the limitations of standardized tests. Additionally, they need opportunities to practice classrooms assessment and interpret assessment information.

All assessors of young children should be knowledgeable about both early childhood development and learning and skilled in the use of early childhood assessment measures, whether they will be using authentic or standardized assessment measures. Assessors also must have knowledge about cultural differences and their impact on development and learning. When implementing systems of authentic assessment, care should be taken to ensure that both the selected tool and the use of that tool are both reliable and valid. Additional competencies related to objectively documenting observations and reliably interpreting those observations against recognized standards are also necessary.

Screening and Identification

Approximately 10% of all children born each year have developmental disabilities or live in environments that place them at risk for delays in learning and development. It has been clearly demonstrated that children with developmental delays who receive early identification and intervention services require less intensive services or no services at all when they are older. Early identification not only effectively promotes positive outcomes for young children and their families, but also has substantial cost benefits to our educational systems and to society. However, in special education, there is a tension between the need to identify children with disabilities early and to provide intervention and the possible harm of labeling children and subjecting them to ineffective treatments. This is complicated by the fallibility of standardized assessment instruments used to determine eligibility for special education. Screening serves as a first step in the process of identifying children who have special needs and ensuring that they receive appropriate services and interventions. Additionally, federal and state special education regulations require that LEAs have a process for identifying children with disabilities beginning at age three.

Developmental Screening

In Rhode Island, all LEAs have established developmental screening programs called Child Outreach and seek to annually screen all children, ages three through five, in the following areas: Vision, Hearing, Speech/Language Skills, Social/Emotional Development, and General Development (including, but not limited to gross and fine motor skills, language, and cognition). Developmental screening, as conducted by Child

Outreach, samples developmental tasks to determine whether a child may experience a challenge that will interfere with the acquisition of knowledge or skills. Developmental screening tests focus on a child's ability to acquire skills as opposed to other types of screening which seek to find out what skills the child has already acquired. Examples of the latter types of screening include literacy screenings and readiness testing. Screening measures should never be used as the sole measure to identify children for special services as they are limited assessments and often administered by staff who are not trained to make interpretations based on the results. Screening and diagnostic assessment measures used to determine whether a child has a disability are designed to assess a child's ability to learn and are traditionally designed to be "curriculum free" and therefore should not be used for instructional planning purposes.

Diagnostic Assessment

Although diagnostic assessment tied to eligibility determination for special education is the more common occurrence in early childhood education, increasingly, results from diagnostic assessment of early academic problems are being employed to guide instruction and intervention. The purpose of diagnostic assessment in early childhood is to identify and secure appropriate intervention services for children whose development and learning is delayed. Diagnostic assessment entails a comprehensive process that addresses specific questions about the development, knowledge and skills of young children. During diagnostic assessment, information is obtained to develop an in-depth analysis and description of a child's level of development in an area or areas of concern. This involves identifying the nature and the severity of the developmental or learning problems comprehensively and systematically. The diagnostic assessment of early academic problems typically considers criterion-referenced, grade-level academic expectations comparing the performance of the individual child to local norms and curriculum benchmarks. The use of norm-referenced diagnostic tests that are not directly connected with the curriculum should be limited in young children. A thorough diagnostic assessment in early childhood includes the following components:

- Use of a valid, reliable tool that is implemented with fidelity
- Developmentally appropriate evaluation tasks (e.g., manipulation of toys and materials for younger children versus pictures and pencil and paper tasks)
- Use of experienced diagnosticians well-versed in child development who have experience working with young children
- Collection of information from multiple sources, including families.

The results of diagnostic assessment are used to guide targeted interventions, as well as to determine eligibility for special education services. Results of diagnostic assessments should be combined with information gathered using authentic assessment methods in a problem solving process to establish potential causality for the delay and develop intervention strategies. The individualized instructional plan stemming from this diagnostic assessment of early academic problems includes the clear articulation of goals and the monitoring plan for measuring progress. The diagnostic assessment of early academic problems should be a part of the responsive system of supports and interventions serving all students in elementary and secondary education.

Eligibility for Special Education

Eligibility determination by the evaluation team is perhaps the most common example of diagnostic

assessment in early childhood. Eligibility determination across all ages requires the development of a full and complete evaluation plan consistent with guidelines requiring assessment of the child in all areas related to the suspected disability. Procedures, methods, criteria, and timelines for determining eligibility for children ages 3 to 21 are regulated through the Rhode Island Special Education Regulations available at: http://www.ride.ri.gov/Special_Populations/State_federal_regulations/Default.aspx

The composition of the evaluation team and process for evaluation is individualized in response to the needs of the child and family. Diagnostic procedures include multiple sources of information collected over multiple points in time, with special attention to the family perspective in gathering information and interpreting results. Best practices include developmentally appropriate, evidenced based, comprehensive evaluation tools and practices administered by highly qualified professionals with expertise in early childhood development. Many young children have limited social exposure making the commitment to assessment and evaluation in a non-discriminatory, culturally and linguistically sensitive process the highest priority.

Given the challenges of standardized assessment inherent with young children, diagnostic assessment in young children relies on authentic assessment practices involving observation of children in their natural environments. While eligibility determination is a primary purpose of diagnostic assessment, the evaluation and assessment process must be embedded into a comprehensive system which guides instruction and intervention and informs the development of the Individual Education Plan.

Ascertaining Outcomes

Early childhood assessment information may be used to monitor trends in children's learning and development, inform program improvement and staff development needs, and to evaluate programs. When using assessment information for these purposes, the assessments must meet high standards of technical adequacy – observational assessments by teachers can only be used when there is sufficient information that the tool and the administration are valid and reliable. When evaluating programs, assessment data should be combined with program data that measure the overall classroom quality and teaching practices – it should not be used as the sole measure of program effectiveness. Additionally, there must be alignment between the assessment tools used for the purposes of classroom instruction and those used for program evaluation. Large scale assessments should use sampling so as not to over burden children and to protect against the potential misuse of assessment information at the individual child level. Authentic assessment measures do not meet the strict standards for technical accuracy required for high-stakes accountability purposes and therefore also should not be used as the only source of evidence when making high-stakes decisions.

Sources

[1] Trimble, S. (2003). NMSA research summary #20: What works to improve student achievement. Westerville, Ohio: National Middle School Association. Retrieved from: http://www.nmsa.org/portals/0/pdf/publications/On_Target/achievement/achievement_4.pdf

[2] Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research,* 77(1), 81-112. doi: 10.3102/003465430298487

[3] Popham, W.J. (2008) Transformative Assessment. Alexandria, VA: ASCD.

[4] Wiliam, D. (2009). *Assessment for Learning: Why, what, and how?* London: Institute of Education University of London.

[5] Rowe, M. B. (1974). Wait time and rewards as instructional variables, their influence on language, logic and fate control: Part one-wait-time. *Journal of Research in Science Teaching*, 11, 81-94. doi: 10.1002/tea.3660110202

[6] Shavelson, R. J. (2006). *On the integration of formative assessment in teaching and learning with implications for teacher education.* Paper prepared for the Stanford Education Assessment Laboratory and the University of Hawaii Curriculum Research and Development Group. Available at www.stanford.edu/dept/SUSE/SEAL

[7] Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). The role of interim assessments in a comprehensive assessment system: A policy brief. Retrieved from http://www.achieve.org/files/TheRoleofInterimAssessments.pdf

[8] Perie, M., Marion, S., & Gong, B. (2007). A framework for considering interim assessments. Washington, D.C.: National Center for the Improvement of Educational Assessment. Retrieved from http://www.nciea.org/publications/ConsideringInterimAssess_MAP07.pdf.

[9] The Council of Chief State School Officers (2008). Interim assessment practices and avenues for state involvement TILSA SCASS interim assessment subcommittee. Retrieved from http://www.ccsso.org/Documents/2008/Interim_Assessment_Practices_and_2008.pdf

[10] American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing.* Washington, D.C.: Author.

[11] Trochim, W. (2006). Research methods knowledge base: Reliability & validity. Retrieved from http://www.socialresearchmethods.net/kb/relandval.php

[12]National Research Council. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, D.C.: National Academy Press.

[13] Edwards, P.A., Turner, J.D., & Mokhtari, K. (2008). Balancing the assessment *of* learning and *for* learning in support of student literacy achievement. *The Reading Teacher*, 61(8), 682-684. doi: 10.1598/RT.61.8.12

[14] Rhodes , L. K., & Shankin, N. L. (1993). *Windows into Literacy: Assessing learners K-8*. Heinemann: Portsmouth, NH.

[15] Torgesen, J.K. (2006). A comprehensive K-3 reading assessment plan: Guidance for school leaders. Portsmouth, NH. RMC Research Corporation, Center on Instruction.

[16] Webb, N.K. (2002) Depth-of-Knowledge Levels for Four Content Areas. Wisconsin Center for Educational Research.

[17] Olinghouse, N. G. (2009). Writing assessment for struggling learners. *Perspectives in Language and Literacy*, Summer, 15-18.

[18] NCTE Writing Assessment: A Position Statement (2009).

[19] NCTE Writing Assessment: A Position Statement (2009).

[20] Routman, R. (2005). *Writing Essentials: Raising Expectations and Results While Simplifying Teaching.* Portsmouth, NH: Heinemann.

[21] Strickland, K. & Strickland, R. (2000). Making Assessment Elementary. Portsmouth, NH: Heinemann.

[22] Newkirk, T. & Kent, R. (2007). *Teaching the Neglected "R": Rethinking Writing Instruction in the Secondary Classrooms.* Portsmouth, NH: Heinemann.

[23] National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics.* Reston, VA: Author.

[24] American Federation of Teachers. (2008). Making classroom assessments work for you: Training Materials. Washington, DC: Author.

[25] Webb, N.& Collins, M. (2008). *Depth of knowledge: Descriptors, examples and question stems for increasing depth of knowledge in the classroom.* Kirksville, MO: Truman State University.

[26]Kentucky Department of Education (2007). Support materials for core content for assessment version 4.1: Mathematics. Retrieved from http://www.education.ky.gov/kde/instructional+resources/curriculum+documents+and+resources/core+content+for+assessment/core+content+for+assessment+4.1/content+specific+core+content+for+assessment+dok+support+materials.htm

[27] Bush, W.S. & Leinwand, S.(Eds.). (2000). *Mathematics assessment: A practical handbook for grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.

[28] Sutton, J. & Krueger, A. (Eds.). (2002). *ED thoughts: What we know about mathematics teaching and learning.* Aurora, CO: Mid-continent Research for Education and Learning.

[29] Ibid.

[30] Bush, W.S. & Leinwand, S.(Eds.). (2000). *Mathematics assessment: A practical handbook for grades 6-8*. Reston, VA: National Council of Teachers of *Mathematics*.

[31] Lesh, R. & Lamon, S.J. (1992). *Assessment of authentic performance in school mathematics.* Washington, DC: American Association for the Advancement of Science.

[32] National Science Teachers Association. (2010). Science education assessments - NSTA's official position on assessment. Arlington, VA: Author. Retrieved from http://www.nsta.org/about/positions/assessment.aspx

[33] Keeley, P. (2008). *Science formative assessment: 75 practical strategies for linking assessment, instruction, and learning.* Thousand Oaks, CA: Corwin Press.

[34] Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Education: Principles, Policy and Practice*, 5, (1), 7-71.

[35] Porter, A. (1993). *Brief to policymakers: Opportunity to learn*. Center on Organization and Restructuring of Schools. Retrieved from http://www.wcer.wisc.edu/archive/cors/brief_to_principals/BRIEF_NO_7_FALL_1993.pdf

[36] National Council for the Social Studies. (2008). A vision of powerful teaching and learning in the social studies: Building effective citizens. *Social Education*, 72 (5), 277-280.

[37] Ibid.

[38] Alleman, J. & Brophy, J. (1999). The changing nature and purpose of assessment in the social studies classroom. *Social Education,* 65 (6), 334-337.

[39] Ibid.