

Running head: AUTHENTIC ASSESSMENT FOR EARLY INTERVENTION ELIGIBILITY

Research Foundations of Authentic Assessments Ensure Accurate and Representative
Early Intervention Eligibility

Marisa Macy
Penn State University

Stephen J. Bagnato
Jennifer Salaway
Cathryn Lehman
Children's Hospital of Pittsburgh/
University of Pittsburgh

Bagnato, S.J., Macey, M., Salaway, J., & Lehman, C. (2007). Research foundations of authentic assessments ensure accurate and representative early intervention eligibility. Pittsburgh, PA: TRACE Center for Excellence in Early Childhood Assessment, Early Childhood Partnerships, Children's Hospital/University of Pittsburgh; US Department of Education, Office of Special Education Programs, and Orelena Hawks Puckett Institute.

Abstract

In this research synthesis, we identified six characteristics of authentic tests and testing practices that are necessary for use with young children to determine their eligibility for early intervention services; we, then, applied these test characteristics to ten commonly used authentic assessment measures. Next, we reviewed the available research on the most frequently used authentic tests and found 26 studies. Findings of this synthesis will help professionals to critically identify characteristics of authentic tests and testing practices that influence the accurate and representative documentation of a young child's degree and pattern of delay or disability to determine early intervention eligibility.

Research Foundations of Authentic Assessments to Ensure Accurate and Representative Early Intervention Eligibility

A child who has patterns of development and learning that are different from peers will be referred for testing to determine if early intervention (EI; birth to three) or early childhood special education (ECSE; three to six) services are warranted. The Individuals with Disabilities Education Improvement Act (IDEIA, 2004; P.L. 108-446) requires that children be evaluated to determine if they are eligible to receive EI/ECSE, and the process often involves the use of standardized, norm-referenced tests on young children under the age of six. While conventional tests are widely used for eligibility determination, their shortcomings require the consideration of alternative formats (the reader is referred to the Bagnato et al. *Cornerstones* article for more information on conventional tests used to determine EI/ECSE eligibility). An approach that is gaining popularity is the use of authentic assessment when determining eligibility for EI/ECSE (Bagnato, 2005; Downs & Strand, 2006; Macy, Bricker, & Squires, 2005; McLean, 2005).

“Authentic assessment refers to the systematic recording of developmental observations overtime about the naturally occurring behaviors of young children in daily routines by familiar and knowledgeable caregivers in the child’s life.” (Bagnato & Yeh-Ho, 2006, p. 16). An authentic assessment involves observations conducted in the child’s natural developmental ecology of the child’s naturally-occurring behaviors or the use of analogue arrangements within daily routines under conditions in which the functional behaviors/skills are elicited (Bagnato, 2007). When functional skills are measured, the data gathered can be linked to service delivery components (i.e., goals, intervention, and evaluation) (Grisham-Brown, Hallam, & Brookshire, 2006; Gulikers, Bastiaens, & Kirschner, 2004). “Authentic” refers to natural observations of typical child functioning in typical daily routines. An authentic assessment process involves

practices where children participate in activities in typical environments (e.g., at home with a parent), are sensitive to children's interests/motivations, and includes the input of family members. Authentic assessment practices have been used to monitor children's progress once they are already in their specially designed programs (Layton & Lock, 2007), however a growing body of research is being conducted to explore the use of authentic assessment for eligibility determination.

Purpose

The aim of this research synthesis is to examine the technical adequacy of authentic measures used for determining a young child's eligibility for EI/ECSE. Two objectives of this research are: (1) to identify *characteristics of authentic measures and associated assessment practices* that are necessary to document a child's degree of delay and disability in an accurate and representative manner in order to qualify children for EI and ECSE services; and (2) to synthesize the body of *research on specific authentic measures* of early development and to appraise their capacity to determine IDEIA eligibility for EI/ECSE services in an accurate and representative manner. This practice-based research synthesis examines available evidence regarding the following ten authentic measures:

- Adaptive Behavior Assessment System (ABAS)
- Assessment Evaluation and Programming System (AEPS)
- Carolina Curriculum for Preschoolers with Special Needs (Carolina)
- Creative Curriculum for Preschool (CCP)
- Child Observation Record (COR)
- Developmental Observation Checklist System (DOCS)
- Hawaii Early Learning Profile (HELP)
- Pediatric Evaluation of Disability Inventory (PEDI)
- Transdisciplinary Play-Based Assessment (TPBA)
- Work Sampling System (WSS)/Ounce

These authentic measures (i.e., ABAS, AEPS, Carolina, CCP, COR, DOCS, HELP, PEDI, TPBA, and WSS/Ounce) were chosen through a review of literature and surveys of practitioners

investigating the most commonly used measures in preschool and early intervention settings (Bagnato, Neisworth, & Munson, 1997; Pretti-Frontczak, Kowalski, & Brown, 2002). There were multiple editions for following measures: AEPS, Carolina, ABAS, COR, CCP and TPBA. Table 1 offers information about these measures.

<Insert Table 1 here>

Characteristics of Assessment Measures

Assessments and assessment practices used for the purposes of determining a child's eligibility for early intervention should meet specific standards. After an initial review of the measurement literature in psychology, special education, early intervention, and early childhood special education, we identified six characteristics that represent quality indicators for authentic tests and testing practices. Procedures used to determine eligibility for EI/ECSE should contain the following characteristics: (1) disability samples, (2) procedural flexibility, (3) comprehensive coverage, (4) functional content, (5) item density, and (6) graduated scoring.

Disability sample refers to the importance of ensuring that the standardization included young children with delays/disabilities in the normative group and any field-validation samples. Research indicates that it is critical that children with functional characteristics similar to the child being tested to be included in the standardization/field-validation sample. The sample should have at least 100 people per each age interval.

The second assessment-related practice characteristic is *procedural flexibility*. Procedural flexibility refers to the extent to which the administration procedures allow professionals to modify the method of assessment (i.e., table-top vs. play), the stimulus attributes of items, and the response modes of the young child to accommodate their functional impairments leading to a more realistic and representative estimate of capabilities.

Authentic assessment measures and processes must ensure *comprehensive coverage* of multiple domains of developmental functioning. Such broad coverage generates results that profile the young child's capabilities across multiple and interrelated functional competencies (i.e., cognitive, motor, adaptive, communication, self-regulatory). Simply, the results of eligibility assessment must reflect the "whole child."

Functional content indicates that assessment items should be comprised of content that requires functional and meaningful skills for everyday life rather than discrete and isolated tasks.

Authentic scales have sufficient *item density* when enough tasks or content are included in the survey of each developmental domain so that even low functional levels can be profiled--the lowest range of standard scores can be obtained when a child does not pass assessment items, or only a few items were scored correctly on a measure or sub-domains. The assessment items should be low enough to discriminate age (i.e., very young children) and level of functioning. There should also be a comprehensive set of items in each age interval to describe a child's performance.

Authentic measures should encompass a *graduated scoring* system to reflect varied skill levels on individual items. For example, a three-point scoring rubric (e.g., yes, emerging, and not yet) provides more information about a child's development rather than a simple two-point scale (e.g., yes/no), and may also indicate the conditions under which a young child can or cannot perform (i.e., with physical prompts, with verbal prompts, with general assistance, independently). The practice characteristics were evaluated by gathering information from test manuals and Buross reviews.

Research-based Use Characteristics

Research studies on the identified authentic assessments were reviewed in this synthesis. More specifically, we examined *research characteristics* associated with the **effective use** or application of each measure. We examined the authentic measures regarding how accurate, representative, and reliable they were in generating results applicable to eligibility determination. The studies were reviewed to provide support for specific “use” or applied effectiveness characteristics determined to be significant when using conventional tests to establish early intervention eligibility.

Research characteristics, determined by data synthesis and professional judgment, include *accuracy* (reliability) and *effectiveness* (validity, utility). Accuracy refers to the extent to which a tool identifies young children with disabilities, specifically. This includes the reliability of the tool, which includes consistency across test items and the use of cut-off scores in order for the tool to precisely or accurately measure a skill or behavior. Examples of accuracy include: test retest reliability, inter-rater reliability, intra-rater reliability, and inter-item consistency.

Effectiveness refers to the extent to which a tool successfully identifies young children with disabilities. This includes the validity of the measure (i.e., including to what extent does the tool measure what it was designed to measure) and how it relates significantly to similar measures. Examples of effectiveness include: predictive validity, concurrent validity, construct validity, test floors, and item gradients.

Search Strategy

Search Terms

Relevant published literature and unpublished position papers, literature reviews, and research studies were identified using the following search terms: assessment (authentic), testing,

early intervention, preschool, early childhood, eligibility, pediatrics, disabilities, handicap identification, referral, and specific assessment tools. More general terms of special schools, state programs resource, centers and evaluations were also used.

The ten authentic assessments were also included within the search. The search was done broadly in the fields of psychology, developmental disabilities, special education, allied health fields (speech and language therapy, physical therapy, occupational therapy), as well as early intervention.

Sources

The primary databases included the following sources: CINAHL, Cochrane Library, Digital Dissertations, Ebsco Host, Education Resource Information Center (ERIC), Google Scholar, Health Source, Illumina, Medline, Ovid/Mental Measurements Yearbook Buros, Psychological Abstracts (PsycINFO), and Social Sciences Citation Index. Additionally, we conducted selective searches of unpublished master's theses and doctoral dissertations. Hand searches of select journals and ancestral searches were also conducted.

Selection Criteria

The study had to meet the following criteria for inclusion: (a) researched one or more of the selected authentic assessment, (b) involved the evaluation of young children with disabilities or at-risk for developing a disability due to environmental or biological risk conditions, (c) examined the accuracy of the measure at testing infants, toddlers, and preschool children with disabilities, and (d) disseminated in a scientific and scholarly publication, which included dissertation and thesis studies. This overall synthesis was conducted as part of literature reviews and syntheses conducted at the Tracking, Referral and Assessment Center for Excellence (Dunst, Trivette, & Cutspe, 2002).

Results

Practice Characteristics

The purpose of Table 2 is to show test characteristics that are needed to determine young children eligible for early intervention. Characteristics were: (1) disability sample, (2) procedural flexibility, (3) comprehensive coverage, (4) functional content, (5) item density, and (6) graduated scoring. Of the ten measures, there were six tests that included children with disabilities in their validation sample. Seven included a graduated scoring system. All but the CCP were comprised of procedural flexibility, comprehensive coverage, and functional content practices characteristics. There were six measures with item density.

<insert Table 2 here>

Research Characteristics

A total of 26 studies on authentic assessment were identified from the fields of child development, early intervention, psychology, special education, physical therapy, pediatrics, and behavioral development. The most studies were conducted on the AEPS, whereas we found the CCP had no studies that met our search criteria. The following information is presented below: total number of studies that met the inclusion criteria, years articles were published, age range included in the studies, and the total number of participants in study samples.

ABAS

- 2 studies
- published between 2006 and 2007
- age range was 33 to 216 months
- 151 children

AEPS

- 8 studies
- published between 1986 and 2005
- age range was birth to 72 months
- 1,516 young children

- *Note-* several studies were found on using the AEPS to develop IEP/IFSPs (Cripe, 1990; Hamilton, 1995; Notari, 1988; Notari & Bricker, 1990; Notari & Drinkwater, 1991; Pretti-Frontczak & Bricker, 2000; Straka, 1994), however we did not include these studies in our synthesis because they did not meet our inclusion criteria which focused on child performance and outcomes.

Carolina

- 1 Italian study
- published in 2006
- mean age for the treatment group was 4.5 months
- 47 children

CCP

- 0 studies

COR

- 3 studies
- published between 1993 and 2005
- age range was 48 to 68 months
- 4,902 children

DOCS

- 5 studies
- published between 1997 and 2005
- age range was 1 to 72 months
- over 2000 children

HELP

- 2 studies
- published between 1995 and 1996
- age range was 22 to 34 months
- 29 children

PEDI

- 2 studies
- published between 1993 and 1998
- age range was 36 to 224 months
- 50 children

TPBA

- 4 studies
- published between 1994 and 2003
- age range was 6 to 46 months
- 74 children

WSS/Ounce

- 1 study
- publication is in review
- age range was 45 to 60 months
- 112 children

Participants

Overall, there were 7,027 young children who participated in the 26 studies. Children's ages ranged from birth to 224 months. Children were identified with various disabilities, and there were several studies that included children without disabilities and children who were at risk for developing a disability. The purpose of Table 3 is to show demographic characteristics of children which include: total sample size, mean age in months, age range in months, and child ability characteristics.

<insert Table 3 here>

Types of Studies

Each study reported in this synthesis examined some aspect of accuracy and/or effectiveness related to one or more of the authentic tests. We found the following types of studies: 13 inter-item/inter-rater reliability, 5 test-retest reliability, 4 sensitivity/specificity, 14 concurrent validity, 3 predictive validity, and 6 construct/criterion validity. Accuracy (reliability) and effectiveness (validity) of the 26 research studies are identified in Table 4.

<insert Table 4 here>

Reported Results

Table 5 incorporates results on the accuracy and effectiveness of authentic assessment measures. A total of 15 studies ($n = 15/26, 57.69\%$) examined the *accuracy* of authentic measures. There were 20 studies ($n = 20/26, 76.92\%$) that examined the *effectiveness* of authentic measures. The number of studies exceeds 26 because some studies examined accuracy and effectiveness.

<insert Table 5 here>

Conclusion

Implications for Practice

Authentic assessments maintain more practice characteristics when compared to conventional tests based on a recent research synthesis (Bagnato et al., submitted). The ABAS, AEPS, Carolina, and HELP had all six practice characteristics recommended for making informed decisions about a child's eligibility for EI/ECSE. The CCP had none of the practice characteristics; while the COR, DOCS, PEDI, TPBA, and WSS/Ounce were missing one or two of the practice characteristics. Areas where some of the authentic measures fell short were: the inclusion of children with disabilities in the validation sample, item density, and graduated scoring. Without these practice characteristics, it is questionable if these authentic measures would be appropriate tools for use in the eligibility determination process as a stand alone tool; however they could be appropriate to corroborate information obtained from a more rigorous measure (Macy & Hoyt-Gonzales, 2007).

State policies can be a barrier to implementing the use of authentic assessments for eligibility determination. Many states have guidelines that may prevent the use of authentic assessments for determining a child eligible for IDEIA services (Layton & Lock, 2007). For example, several states require standard scores to identify EI/ECSE eligibility; therefore conventional tests would probably be favored in those states since most authentic assessments do not provide standard scores. However, as noted in Table 1, some of the authentic assessments do have eligibility features like cut off scores or age intervals that might be useful during the initial eligibility evaluation.

Evaluating a child to determine eligibility, using conventional testing practices, present the following problems not usually found with an authentic assessment approach:

- takes the child away from meaningful learning activities and experiences,
- does not engage the child by building on their interests and motivations,
- and requires the child to attend to tasks that are adult-directed.

Once a child is found eligible using a conventional test, further testing is almost always needed to determine specific supports and services. Time, expenses, and resources can be better utilized by using an authentic eligibility assessment approach due to the benefits of linking the initial eligibility assessment directly to programmatic content (Macy & Hoyt-Gonzales, 2007). When authentic assessments are used, professionals can focus efforts on program (i.e., IEP or IFSP) development and implementation without having to administer additional unnecessary tests.

Children will enter their program with relevant initial eligibility assessment information that will lead to the creation of tailored learning and developmental goals, objectives, and interventions (Macy, 2004; Guralnick, 2005). The eligibility assessment process is a very important time when information is gathered about a child. If the child is determined to be eligible for early intervention services, the assessment information gathered during this phase should be useful for developing goals, intervention content, and evaluation. In order to use resources more effectively, efforts should be directed toward linking eligibility assessment with the programmatic phase of the service delivery system (Bagnato & Neisworth, 1992).

An authentic linked process is needed to identify children in need of services more efficiently (i.e., make use of resources in a different way), as well as increase the effectiveness of intervention with useful information needed to develop goals, interventions, and evaluation. An

authentic assessment process used to determine eligibility has the potential to offer benefits to children and families, professionals, and early intervention systems of service delivery.

Benefits to children. Children will likely benefit from an authentic eligibility assessment because they will not have to wait an unreasonable amount of time for individualized intervention to begin (Macy & Bricker, 2006), as they often do in situations of typical conventional testing. As a matter of fact, in a study that compared conventional and authentic assessments, researchers found the overall mean number of days to complete the eligibility assessment process per child was $M = 58.3$ ($SD = 42.6$) for the *conventional* assessment, and $M = 36.7$ days ($SD = 20.9$) for the *authentic* assessment (Myers, McBride, & Peterson, 1996). A significant savings of 22 days is an important finding given the legislative mandate to complete eligibility assessment within 45 days. Relevant information focused on goals and intervention would immediately be available. From the completion of eligibility assessment, children will be provided with appropriate services that will address their individual needs (Neisworth & Bagnato, 2004). Earlier and/or sooner is almost always better when addressing developmental needs.

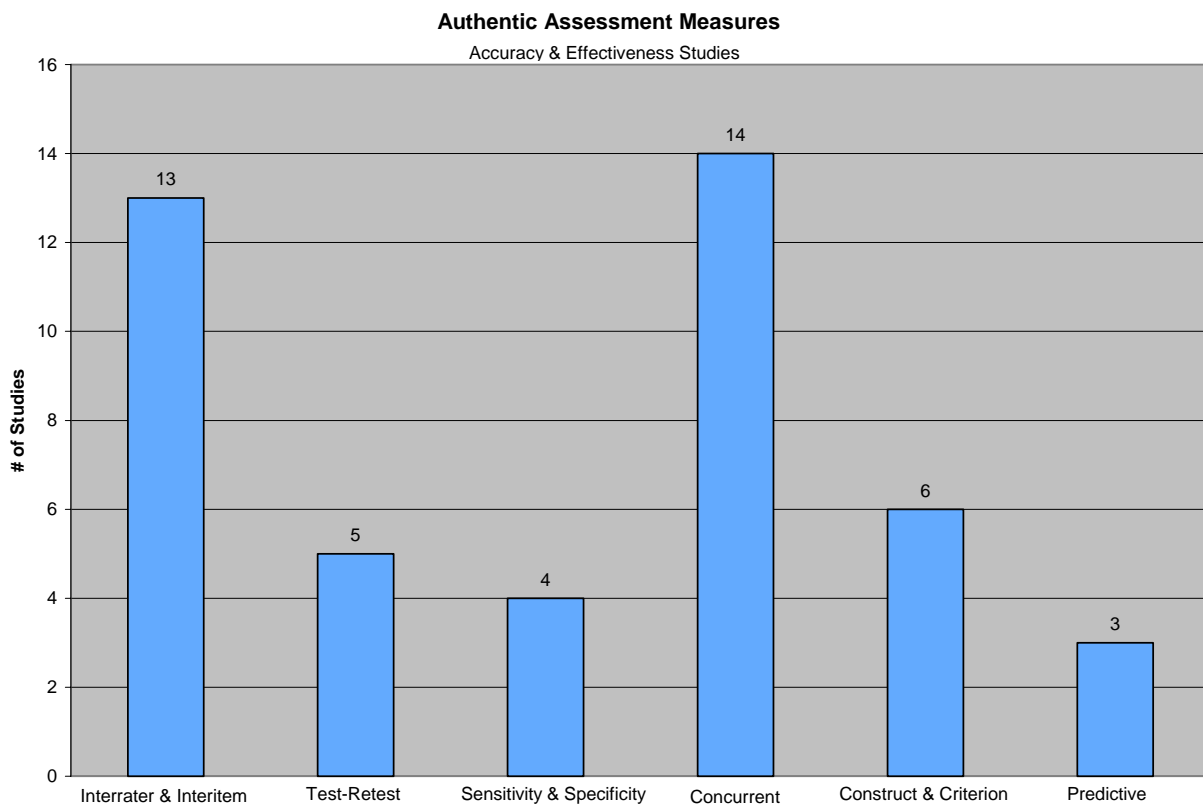
An authentic eligibility assessment approach may also result in more efficient testing of children by building on information gathered from the initial eligibility assessment(s), rather than conducting totally new evaluations once children have begun their EI/ECSE program. In addition, children will benefit from their placement (e.g., classrooms, play groups, and community based programs) because parents (Dunst, 2002; Guralnick, 2006; Kim, Sugawara, & Kim, 2000) and professionals will be better prepared to meet their needs because of the collaborative nature of the authentic assessment process (Bagnato & Neisworth, 1999; 2005).

Benefits to professionals. Professionals (e.g., teachers) may benefit from authentic eligibility assessment for many reasons. First, classroom-based programs often serve large groups of children with varying needs, and professionals often feel they do not have adequate time to spend with children in order to learn their individual strengths and needs. Professionals may benefit from an authentic eligibility measure because from the time a child enters their classroom/program they will have information needed for effective intervention. Professionals can save time because children will enter the program with assessment information that will permit developing goals and objectives without having to administer additional unnecessary assessments and evaluations. Professionals reported high levels of satisfaction with authentic assessment based on several utility studies of the AEPS (Cripe, 1990; Hamilton, 1995; Kim, & Sugawara, 1998; Notari, 1988; Notari & Bricker, 1990; Notari & Drinkwater, 1991; Pretti-Frontczak & Bricker, 2000; Straka, 1994).

Benefits to the early intervention delivery system. Finally, an early intervention system of service delivery would benefit from authentic eligibility assessment because resources will be managed more efficiently. In 2004, there were 293,816 (2.40% of the total population) children under the age of three who received early intervention services across the United States (IDEAdata, 2007). While the population of children continues to grow steadily, so does the cost of services. In 2003, the support appropriated to early intervention systems of service delivery amounted to approximately 434 million dollars (IDEAdata, 2004). By implementing an authentic assessment approach, systems would manage resources effectively by saving time on administering assessments done to determine eligibility and refocus system-wide efforts toward treatment and child/family outcomes.

Implications for Research

The authentic assessment studies reported here in this synthesis are the beginnings of a research base for supporting the use of authentic assessment for eligibility determination. The graph that follows is an illustration of the number of studies on the accuracy and effectiveness of the authentic approach. While this foundation is a good start, more evidence is needed to continue to examine the impact of using authentic measures.



In order to establish whether or not a child was eligible for early intervention services, conventional assessments were used as a comparison in many of the concurrent and construct validity studies reviewed in this synthesis. McLean, Bailey, and Wolery (2004) suggest that one way to examine construct validity is to establish convergent validity by examining high positive correlations with other tests that measure the same constructs. Future research could be

conducted to examine the construct validity of authentic assessments used for eligibility determination across developmental domains and target behaviors. The field of early intervention has traditionally used a non-linked and conventional approach to determining children eligible for IDEA services, and this is usually done because there are currently no exemplars or measures that can determine eligibility that both meets established policy requirements (e.g., standard scores) and also links to service delivery components in an authentic context. Instead of comparing the ten authentic measures to other good examples of authentic and linked eligibility assessment, many of the studies in our review were forced to make comparisons to conventional, non-linked eligibility measures because there is not a gold standard available. Future research should continue not only to examine the accuracy and effectiveness of authentic measures as well as conventional tests, but also to compare both forms of evaluation to an external standard (e.g., correct identification rates based on expert consensus; the need for services—service-based eligibility; or probability of succeeding/progressing in regular education or typical setting with typical peers without support services). *At this point, in general, neither conventional tests/testing, nor authentic assessment procedures have a sufficient evidence-base to confidently propose their use for early intervention eligibility.*

It would also be helpful to conduct cost/benefit analysis by comparing authentic and conventional testing practices used to determine children eligible for IDEA services. This type of evidence would aid policy makers when re-authorizing policies and updating practices. In addition, it would be important to study the effects of the initial eligibility assessment on child outcomes using a longitudinal design.

Authentic assessment can be used for designing a program for a child, creating interventions, and be used to evaluate the efficacy of a child's individualized program. Not only

does authentic assessment have potential to accurately identify children in need of services, it also has important implications beyond eligibility determination. Using authentic assessment to determine children eligible for IDEA services has the potential to improve services for children (Bagnato, 2007; Grisham-Brown et al., 2006; Gulikers et al., 2004; Layton & Lock, 2007).

References

- * Indicates *studies* used in the synthesis
- ** Indicates *measures* used in the synthesis

- *Anthony, T. L. (2003). *The inclusion of vision-development guidelines in the transdisciplinary play-based assessment: A study of reliability and validity*. Unpublished doctoral dissertation. University of Denver, Denver, CO.
- Bagnato, S. J. (2005). The authentic alternative for assessment in early intervention: An emerging evidence-based practice. *Journal of Early Intervention*, 28(1), 17-22.
- Bagnato, S. J. (2007). *Authentic assessment for early childhood intervention best practices. The Guilford school practitioner series*. New York: Guilford Press.
- Bagnato, S. J., & Neisworth, J. T. (1991). *Assessment for early intervention: Best practices for professionals*. New York: Guilford.
- Bagnato, S. J. & Neisworth, J. T. (1992). The case against intelligence testing in early intervention. *Topics in Early Childhood Special Education*, 12, 1-20.
- Bagnato, S. J., & Neisworth, J. T. (1994). A national study of the social and treatment "invalidity" of intelligence testing for early intervention. *School Psychology Quarterly*, 9(2), 81-102.
- Bagnato, S. J., & Neisworth, J. T. (1999). Collaboration and teamwork in assessment for early intervention. *Comprehensive Psychiatric Assessment of Young Children*, 8(2), 347-363.
- Bagnato, S. J., & Neisworth J.T. (2005). DEC recommended practices: Assessment. In S. Sandall, M.L. Hemmeter, B.J. Smith, & M. McLean (Eds.), *DEC Recommended Practices* (pp.45-50). Longmont, CO: Sopris West.
- Bagnato, S. J., Neisworth, J. T., & Munson, S. (1997). *LINKing assessment and early intervention*. Baltimore: Brookes.
- *Bailey, E., & Bricker, D. (1986). A psychometric study of a criterion-referenced assessment designed for infants and young children. *Journal of the Division of Early Childhood*, 10(2), 124-134.
- *Baird, S.M., Campbell, D., Ingram, R., & Gomez, C. (2001). Young children with Cri-du-Chat: Genetic, developmental, and behavioral profiles. *Infant-Toddler Intervention*, 11(1), 1-14.
- **Bricker, D. (Ed.) et al. (2002). *Assessment, evaluation, and programming system for infants and children* (2nd ed.). Baltimore: Brookes.
- **Bricker, D. (Ed.) et al. (1993). *Assessment, evaluation, and programming system for infants and children* (1st ed.). Baltimore: Brookes.
- *Bricker, D., Bailey, E., & Slentz, K. (1990). Reliability, validity, and utility of the evaluation and programming system: For infants and young children (EPS-I). *Journal of Early Intervention*, 14(2), 147-160.
- *Bricker, D., Yovanoff, P., Capt, B., & Allen, D. (2003). Use of a curriculum-based measure to corroborate eligibility decisions. *Journal of Early Intervention*, 26(1), 20-30.
- *Calhoon, J. M. (1997). Comparison of assessment results between a formal standardized measure and a play-based format. *Infant-Toddler Intervention: The Transdisciplinary Journal*, 7, 201-214.
- *Cody, D. A. (1995). *Play assessment as a measure of cognitive ability in developmentally delayed preschoolers*. Unpublished doctoral dissertation, George Mason University, Fairfax, VA.

- Cripe, J.W. (1990). *Evaluating the effectiveness of training procedures in a linked system approach to individual family service plan development*. Unpublished doctoral dissertation, University of Oregon, Eugene.
- Danaher, J. (2005). *Eligibility policies and practices for young children under Part B of IDEA* (NECTAC Notes No.15). Chapel Hill: The University of North Carolina, FPG Child Development Institute, National Early Childhood Technical Assistance Center.
- *DelGiudice, E., Brogna, G., Romano, A., Paludetto, R., & Toscano, E. (2006). Early intervention for children with Down syndrome in southern Italy: The role of parent-implemented developmental training. *Infants & Young Children, 19*(1), 50-58.
- *Di Pinto, M. (2006). The ecological validity of the Behavior Rating Inventory of Executive Functioning (BRIEF) in Attention Deficit Hyperactivity Disorder: Predicting academic achievement and social adaptive behavior in the subtypes of ADHD. *Dissertation Abstracts International Section B: The Sciences and Engineering, 67*(3-B), 1696. (UMI No. 3209793).
- Downs, A., & Strand, P. S. (2006). Using assessment to improve the effectiveness of early childhood education. *Journal of Child and Family Studies, 15*(6), 671-680.
- Dunst, C. J. (2002). Family-centered practices: Birth through high school. *Journal of Special Education, 36*(3), 139-147.
- Dunst, C.J., Trivette, C.M., & Cutspe, P.A. (2002). An evidence-based approach to documenting the characteristics and consequences of early intervention practices. *Centerscope, 1*(2), 1-6. Available at <http://www.evidencebasepractices.org/centerscope/centerscopevol1no2.pdf>.
- *Fantuzzo, D., Grim, S., & Montes, G. (2002). Generalization of the Child Observation Record: A validity study for diverse samples of urban, low-income preschool children. *Early Childhood Research Quarterly, 17*, 106-125.
- *Friedli, C. (1994). *Transdisciplinary play-based assessment: A study of reliability and validity*. Unpublished doctoral dissertation. University of Colorado, Denver.
- *Gilbert, S.L. (1997). *Parent and teacher congruency on variations of a screening assessment: An examination* (Report No. H023B50009). AL: Auburn University (ERIC Document Reproduction Service No. ED413726).
- Grisham-Brown, J., Hallam, R., & Brookshire, R. (2006). Using authentic assessment to evidence children's progress toward early learning standards. *Early Childhood Education Journal, 34*(1), 45-51
- Gulikers, T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), 67-86.
- Guralnick, M. J. (2005). *The developmental systems approach to early intervention*. Baltimore: Brookes.
- Guralnick, M. J. (2006). Family influences on early development: Integrating the science of normative development, risk and disability, and intervention. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 44-61). Malden, MA: Blackwell Publishing.
- **Haley, S.M., Coster, W.J., Ludlow, L.H., Haltiwanger, J.T., & Andrellos, P.J. (1992). *Pediatric Evaluation of Disability Inventory (PEDI)*. Boston, MA: New England Medical Center Hospitals, Inc.
- Hamilton, D. A. (1995). *The utility of the assessment evaluation programming system in the development of quality IEP goals and objectives for young children, birth to three, with*

- visual impairments*. Unpublished doctoral dissertation, University of Oregon, Eugene, Oregon.
- **Harrison, P.L., & Oakland, T. (2003). *Adaptive Behavior Assessment System-Second Edition*. San Antonio, TX: The Psychological Corporation.
- **High/Scope Educational Research Foundation. (2003). *Preschool Child Observation Record 2nd Edition*. Ypsilanti, Michigan: High/Scope Press.
- **Hresko, W.P., Miguel, S.A., Sherbenou, R.J., & Burton, S.D. (1994). *Developmental Observation Checklist System*. Austin, TX: PRO-ED.
- *Hsia, T. (1993). *Evaluating the psychometric properties of the Assessment, Evaluation, and Programming System for Three to Six Years: AEPS Test*. Unpublished doctoral dissertation, University of Oregon, Eugene, Oregon.
- **Johnson-Martin, N. M., Attermeier, S. M., & Hacker, B. J. (2004). *The Carolina curriculum for infants and toddlers with special needs (3rd ed.)*. Baltimore: Brookes.
- **Johnson-Martin, N. M., Hacker, B. J., & Attermeier, S. M. (2004). *The Carolina curriculum for preschoolers with special needs (2nd ed.)*. Baltimore: Brookes.
- Kim, Y. M., & Sugawara, A. I. (1998). Perceptions of early childhood special education professionals: Assessment of functional skills among children with special needs. *Early Child Development and Care, 146*, 33-40.
- Kim, Y. M., Sugawara, A. I., & Kim, G. (2000). Parents' perception of and satisfaction with the eligibility assessment of their children with special needs. *Early Child Development and Care, 160*, 133-142.
- *Knox, V., & Usen, Y. (1998). Clinical review of the pediatric evaluation of disability Inventory. *The British Journal of Occupational Therapy, 63*, 29-32.
- Layton, C. A., & Lock, R. H. (2007). Use authentic assessment techniques to fulfill the promise of no child left behind. *Intervention in School & Clinic, 42*(3), 169-173.
- ** Linder, T. W. (1993). *Transdisciplinary play-based assessment: A functional approach to working with young children*. Baltimore: Paul H. Brookes Publishing Co.
- Macy, M. (2004). *Using a curriculum-based assessment to determine eligibility for early intervention services*. Unpublished doctoral dissertation, University of Oregon, Eugene, Oregon.
- Macy, M., & Bricker, D. (2006). Practical applications for using curriculum-based assessment to create embedded learning opportunities for young children. *Young Exceptional Children, 9*(4), 12-21.
- Macy, M., & Hoyt-Gonzales, K. (2007). A Linked System Approach to Early Childhood Special Education Eligibility Assessment. *TEACHING Exceptional Children, 39*(3), 40-44.
- *Macy, M. G., Bricker, D. D., & Squires, J. K. (2005). Validity and reliability of a curriculum-based assessment approach to determine eligibility for part C services. *Journal of Early Intervention, 28*(1), 1-16.
- *McKeating-Esterle, E., Bagnato, S.J., Fevola, A., & Hawthorne, C. (2007). *A comparison of clinical judgment (informed opinion) with performance-based assessment to document early intervention status*. Unpublished masters thesis, University of Pittsburgh, Pittsburgh, PA.
- McLean, M. (2005). Using curriculum-based assessment to determine eligibility: Time for a paradigm shift? *Journal of Early Intervention, 28*(1), 23-27.
- McLean, M., Bailey, D.B., & Wolery, M. (2004). *Assessing infants and preschoolers with special needs (3rd ed.)*. Upper Saddle River, NJ: Pearson.

- **Meisels, S.J., Dombro, A.L., Marsden, D.B., Weston, D., & Jewkes, A. (2003). *The Ounce Scale: An Observational Assessment for Infants, Toddlers, and Families*. New York: Pearson Early Learning.
- **Meisels, S.J., Jablon, J.R., Marsden, D.B., Dichtelmiller, M.L., & Dorfman, A.B. (2001). *The Work Sampling System* (4th ed.). New York: Pearson Early Learning.
- *Meisels, S.J., Xue, Y., & Shablott, M. (in review). Assessing language, literacy, and mathematics skills with Work Sampling Head Start.
- *Myers, C. L., McBride, S. L., & Peterson, C. A. (1996). Transdisciplinary, play-based assessment in early childhood special education: An examination of social validity. *Topics in Early Childhood Special Education, 16*, 102-126.
- NASP (2005). *NASP position statement on early childhood assessment*. Bethesda, MD: Association of School Psychologists.
- Neisworth, J. T., & Bagnato, S. J. (2004). The mismeasure of young children: The authentic assessment alternative. *Infants & Young Children, 17*(3), 198-212.
- *Noh, J. (2005). *Examining the psychometric properties of the second edition of the Assessment, Evaluation, and Programming System for Three to Six Years: AEPS Test 2nd Edition (3-6)*. Unpublished dissertation. Eugene: University of Oregon.
- Notari, A. R. (1988). *The utility of a criterion-referenced instrument in the development of individualized education plan goals for infants and young children*. Unpublished doctoral dissertation. University of Oregon.
- Notari, A., & Bricker, D. (1990). The utility of a curriculum-based assessment instrument in the development of individualized education plans for infants and young children. *Journal of Early Intervention, 14*(2), 117-132.
- Notari, A., & Drinkwater, S. (1991). Best practice for writing child outcomes: An evaluation of two methods. *Topics in Early Childhood Special Education, 11*, 92-106.
- **Parks, S., Furono, S., O'Reilly, K., Inatsuka, T., Hoska, C. M., & Zeisloft-Falbey, B. (1994). Hawaii early learning profile (HELP). Palo Alto: Vort.
- Pretti-Frontczak, P., & Bricker, D. (2000). Enhancing the quality of Individualized Education Plan (IEP) goals and objectives. *Journal of Early Intervention, 23*(2), 92-105.
- Pretti-Frontczak, K., Kowalski, K., & Brown, R. D. (2002). Preschool teachers' use of assessments and curricula: A statewide examination. *Exceptional Children, 69*(1), 109-123.
- *Sayers, L. K., Cowden, J. E., Newton, M., Warren, B., & Eason, B. (1996). Qualitative analysis of a pediatric strength intervention on the developmental stepping movements of infants with Down syndrome. *Adapted Physical Activity Quarterly, 13*(3), 247-268.
- *Schweinhart, L.J., McNair, S., Barnes, H., & Lerner, M. (1993). Observing young children in action to assess their development: The High/Scope Child Observation Record study. *Educational and Psychological Measurement, 53*, 445-454.
- *Sekina, Y., & Fantuzzo, J. (2005). Validity of the Child Observation Record: An investigation of the relationship between COR dimensions and social-emotional and cognitive outcomes for Head Start children. *Journal of Psychoeducational Assessment, 23*, 242-261.
- Shackelford, J. (2006). *State and jurisdictional eligibility definitions for infants and toddlers with disabilities under IDEA* (NECTAC Notes No. 21). Chapel Hill: The University of North Carolina, FPG Child Development Institute, National Early Childhood Technical Assistance Center.

- *Sher, N. (2000). *Activity-based assessment: Facilitating curriculum linkage between eligibility evaluation and intervention*. Unpublished doctoral dissertation, University of Oregon, Eugene, Oregon.
- *Slentz, K. (1986). *Evaluating the instructional needs of young children with handicaps: Psychometric adequacy of the Evaluation and Programming System-Assessment Level II*. Dissertation Abstracts International, 47(11), 4072A.
- Straka, E. (1994). *Assessment of young children for communication delays*. Unpublished doctoral dissertation, University of Oregon, Eugene, Oregon.
- *Wright, F.V., & Boschen, K. (1993). The Pediatric Evaluation of Disability Inventory (PEDI) validation of a new functional assessment outcome instrument. *Canadian Journal of Rehabilitation*, 7, 41-42.

Table 1

Ten Authentic Measures and their Characteristics

	Age Range	Domains	# of Items	Scoring Features	Family Involvement	Eligibility Features
ABAS	Birth to 89 years 5 rating forms	Communication, Community Use, Functional Pre-Academics, School/Home Living, Health and Safety, Leisure, Self-Care, Self-Direction, Social, Motor (10)	<i>Parent Form</i> (B-5): 241 <i>Teacher Form</i> (2-5): 216	4-point rating scale: 3 = Always when needed 2 = Sometimes when needed 1 = Never when needed 0 = Is not able Includes a box to check if rater guessed. Includes section for rater to make comments regarding a specific item.	Parent/Primary Caregiver form is designed to be completed by parents or other primary caregivers. Two forms are available: Ages birth to 5 and Ages 5-21.	Provides norm-referenced scores based on age.
AEPS	Birth to Six	Adaptive, Cognitive, Fine Motor, Gross Motor, Social-Communication, & Social (6)	B-3: 249 3-6: 217	3- point rating scale: 2 = consistently meets criterion, 1 = in-consistently meets criterion, and 0 = does not meet criterion. Six qualifying notes.	Family Report allows parents to be involved in collecting information and list/ prioritize areas of interest.	Provides cut off scores by age intervals.
Carolina	Birth to Pre-school	Cognition, Cognition/Communication, Personal-Communication, Fine Motor, Gross Motor (6)	B-2: 359 2-5: 400	3-point rating scale: (+) mastery (+/-) inconsistent/emerging skill (-) unable to perform skill Qualifying notes	Families encouraged being involved throughout the assessment and instruction process.	n/a

CCP

	Age Range	Domains	# of Items	Scoring Features	Family Involvement	Eligibility Features
COR	Two and a half to Six	Initiative Social Relations Creative Representation Music and Movement Language and Literacy Mathematics and Science (6)	30 items, each with 5 levels	5-point rating scale: Five descriptive statements that represent a range of functioning from very poor to very superior	Family Report is available to create reports for parents about their child that can be discussed at parent conferences or home visits. Parents are able to record notes about the child based on the parents' observations of the child's behavior at home on the report. Parent Guide available to explain the COR and for parents to record anecdotes based on the COR.	n/a
DOCS	Birth to Six	Developmental Checklist (DC): Language, Social, Motor, Cognition (4); Adjustment Behavior Checklist (ABC); Parental Stress and Support Checklist (PSSC)	DC: 475 ABC: 25 PSSC: 40	<u>DC</u> : Raters check Yes or No (Yes = 1; No =0) <u>ABC</u> : 4-point rating scale: Very Much Like, Somewhat Like, Not Much Like, Not At All Like <u>PSSC</u> : 4-point rating scale: Highly Agree, Somewhat Agree, Sometimes Agree, Do Not Agree	Parents are viewed as primary informant on the DC questionnaire. PSSC assesses parental stress, parental support, child adaptability, parent-child interaction, and environmental impact.	Provides norm-referenced scores based on age.
HELP	Birth to Six	Cognition, Fine Motor, Gross Motor, Language, Social Emotional, Self Help (6)	685	4-point rating scale: (+) skill or behavior is present (-) skill is not present (+/-) skill appears to be emerging (A) skill or behavior is atypical or dysfunctional Qualifying notes	Materials available to increase parent participation in the assessment process; Guidelines for inclusion of parent input are spread throughout.	Provides age ranges for skills.

	Age Range	Domains	# of Items	Scoring Features	Family Involvement	Eligibility Features
PEDI	Six months to seven and a half years	Self-Care Mobility Social Function (3)	Functional Skills: 197 Caregiver Assistance : 20 Modifications: 20	<u>Functional Skills</u> : 0= unable, or limited in capability to perform item in most situations; 1= capable of performing item in most situations <u>Caregiver Assistance</u> : Independent, Supervise/ Prompt/ Monitor, Minimal Assistance, Moderate Assistance, Maximal Assistance, Total Assistance <u>Modifications</u> : No modifications, Child-Oriented, Rehabilitation, Extensive Modifications	Options for administration include parent interview.	Provides norm-referenced scores.
TPBA	Birth to six	Cognitive, Social-Emotional, Communication and Language, Sensorimotor Development	A list of developmental skills observed during play	Scoring system: (+) skill at age level and his/her skills are qualitatively strong (-) skill is below age level and team has qualitative concerns (√) need for further observation and/or testing (NO) no opportunity (NA) not applicable due to age or disability	Includes play with parent in the assessment sequence	n/a

	Age Range	Domains	# of Items	Scoring Features	Family Involvement	Eligibility Features
WSS/ Ounce	WSS: Pre-Grade 5 Ounce: Birth to 3 ½	WSS: Personal and Social Development, Language and Literacy, Mathematical Thinking, Scientific Thinking, Social Studies, The Arts, Physical Development & Health (7) Ounce: Personal Connections, Feelings About Self, Relationships with Other Children, Understanding and Communication, Exploration and Problem Solving, Movement and Coordination (6)	n/a	3 types of ratings: Not Yet In Process Proficient	Ounce: Contains a Family Album element that is used by families to collect observations, photos, and mementos of their child's growth and development.	n/a

Note. Adaptive Behavior Assessment System (ABAS); Assessment Evaluation & Programming System (AEPS); Carolina Curriculum for Infants/Toddlers/Preschoolers with Special Needs (Carolina); Creative Curriculum for Preschool (CCP); Child Observation Record (COR); Developmental Observation Checklist System (DOCS); Hawaii Early Learning Profile (HELP); Pediatric Evaluation of Disability Inventory (PEDI); Transdisciplinary Play-Based Assessment (TPBA); Work Sampling System (WSS)/Ounce.

Table 2

Practice Characteristics of Authentic Measures

Measures (<i>N</i> = 10)	Standardization Includes Children with Disabilities	Procedural Flexibility	Comprehensive Coverage	Functional Content	Item Density	Graduated Scoring
ABAS	Yes	Yes	Yes	Yes	Yes	Yes
AEPS	Yes	Yes	Yes	Yes	Yes	Yes
Carolina	Yes	Yes	Yes	Yes	Yes	Yes
CCP	--	--	--	--	--	--
COR	No	Yes	Yes	Yes	No	Yes
DOCS	Yes	Yes	Yes	Yes	No	No
HELP	Yes	Yes	Yes	Yes	Yes	Yes
PEDI	Yes	Yes	Yes	Yes	Yes	No
TPBA	No	Yes	Yes	Yes	Yes	Yes
WSS/Ounce	No	Yes	Yes	Yes	No	Yes

Note. Adaptive Behavior Assessment System (ABAS); Assessment Evaluation & Programming System (AEPS); Carolina Curriculum for Infants/Toddlers/Preschoolers with Special Needs (Carolina); Creative Curriculum for Preschool (CCP); Child Observation Record (COR); Developmental Observation Checklist System (DOCS); Hawaii Early Learning Profile (HELP); Pediatric Evaluation of Disability Inventory (PEDI); Transdisciplinary Play-Based Assessment (TPBA); Work Sampling System (WSS)/Ounce.

Table 3

Research Studies with Participant Demographic Information

Author(s) and Year (N = 26)	Sample Size	Age Range in Months	Child Characteristics
Anthony (2003)	10	6 to 46	Developmental delay
Bailey & Bricker (1986)	32	n/a	Children with and without developmental delay
Baird, Campbell, Ingram, & Gomez (2001)	13	11 to 47	Cri-du-Chat syndrome
Bricker, Bailey, & Slentz (1990)	335	2 to 72	Children with (mild, moderate, and severe) and without developmental delay, and at risk
Bricker, Yovanoff, Capt, & Allen (2003)	861	1 to 72	Children eligible and not eligible for early intervention/early childhood special education
Calhoon (1997)	4	22 to 35	Language delay
Cody (1995)	25	22 to 34	Previously identified as delayed in the areas of behavior, cognition, and language
DelGiudice, Brogna, Romano, Paludetto, & Toscano (2006)	47	n/a	Down syndrome
Di Pinto (2006)	60	60-216	ADHD (ADHD/PI; ADHD/C)
Fantuzzo, Grim, & Montes (2002)	733 /1,427	n/a	Urban and low income
Friedli (1994)	20	n/a	Children with and without developmental delay
Gilbert (1997)	100	1 to 72	Children with and without developmental delay
Hsia (1993)	82	36 to 72	Children with and without developmental delay
Knox & Usen (1998)	10	45 to 224	Cerebral Palsy
Macy, Bricker, & Squires (2005)	68	6 to 36	Children eligible and not eligible for early intervention/early childhood special education
McKeating-Esterle, Bagnato, Fevola, & Hawthorne (2007)	91	33 to 71	Developmental Delay, Autism, Hearing Impairment/Deafness, Down Syndrome, MR, CP/Muscular Dystrophy, Speech/language impairment, Visual impairment/blindness, Other Health Impairment, Multiple Disabilities
Meisels, Xue, & Shamblott, (in review)	112	45.24 to 59.76	At risk; Children with special needs whose IEPs indicated that they were in the mild to moderate range (speech or physical impairment)
Morgan (2005)	32	4 to 60	Reactive Attachment Disorder

Author(s) and Year (<i>N</i> = 26)	Sample Size	Age Range in Months	Child Characteristics
Myers, McBride, & Peterson (1996)	40	7-36	Developmental delay
Noh (2005)	65	36 to 64	Children eligible and not eligible for early intervention/early childhood special education
Sayers, Cowden, Newton, Warren, & Eason (1996)	4	n/a	Down syndrome
Schweinhart, McNair, Barnes, & Lerner (1993)	2,500	n/a	Low income
Sekina & Fantuzzo (2005)	242	55 to 68	Urban
Sher (2000)	20	36 to 67	Children eligible and not eligible for early intervention/early childhood special education
Slentz (1986)	53	36 to 72	Children with and without developmental delay
Wright & Boschen (1993)	40	36 to 84	Cerebral Palsy

Note. Adaptive Behavior Assessment System (ABAS); Assessment Evaluation & Programming System (AEPS); Carolina Curriculum for Infants/Toddlers/Preschoolers with Special Needs (Carolina); Creative Curriculum for Preschool (CCP); Child Observation Record (COR); Developmental Observation Checklist System (DOCS); Hawaii Early Learning Profile (HELP); Pediatric Evaluation of Disability Inventory (PEDI); Transdisciplinary Play-Based Assessment (TPBA); Work Sampling System (WSS)/Ounce.

Table 4

Research Study Characteristics

Author(s) and Year (N = 26)	Test	Accuracy (Reliability)			Effectiveness (Validity)		
		Inter-item ^a Inter-rater ^b	Test- retest	Sensitivity ^a Specificity ^b	Concurrent	Construct ^a Criterion ^b	Predictive
Anthony (2003)	TPBA		X				
Bailey & Bricker (1986)	AEPS	X ^b	X		X		
Baird, Campbell, Ingram, & Gomez (2001)	DOCS			X ^a			
Bricker, Bailey, & Slentz (1990)	AEPS	X ^b	X		X		
Bricker, Yovanoff, Capt, & Allen (2003)	AEPS			X ^a X ^b			
Calhoon (1997)	TPBA				X		
Cody (1995)	AEPS				X		
DelGiudice, Brogna, Romano, Paludetto, & Toscano (2006)	AEPS				X		
Di Pinto (2006)	ABAS						X
Fantuzzo, Grim, & Montes (2002)	COR					X ^a	
Friedli (1994)	TPBA	X ^b	X		X		
Gilbert (1997)	DOCS	X ^b	X				
Hsia (1993)	AEPS	X ^a X ^b				X ^a	
Knox & Usen (1998)	PEDI			X ^a			
Macy, Bricker, & Squires (2005)	AEPS	X ^b		X ^a X ^b	X		
Meisels, Xue, & Shablott, (in review)	WSS				X		X
McKeating-Esterle, Bagnato, Fevola, & Hawthorne (2007)	ABAS				X		
Morgan (2005)	DOCS						X
Myers, McBride, & Peterson (1996)	TPBA					X ^a	
Noh (2005)	AEPS	X ^a X ^b			X		
Sayers, Cowden, Newton, Warren, & Eason (1996)	AEPS				X		

Studies (<i>N</i> = 26)	Test	Inter-item ^a Inter-rater ^b	Test- retest	Sensitivity ^a Specificity ^b	Concurrent	Construct ^a Criterion ^b	Predictive
Schweinhart, McNair, Barnes, & Lerner (1993)	COR	X ^b			X		
Sekina, & Fantuzzo (2005)	COR					X ^a	
Sher (2000)	AEPS	X ^b			X	X ^a	
Slentz (1986)	AEPS	X ^a X ^b	X		X	X ^a	
Wright & Boschen (1993)	PEDI	X ^a X ^b					

Note. Adaptive Behavior Assessment System (ABAS); Assessment Evaluation & Programming System (AEPS); Carolina Curriculum for Infants/Toddlers/Preschoolers with Special Needs (Carolina); Creative Curriculum for Preschool (CCP); Child Observation Record (COR); Developmental Observation Checklist System (DOCS); Hawaii Early Learning Profile (HELP); Pediatric Evaluation of Disability Inventory (PEDI); Transdisciplinary Play-Based Assessment (TPBA); Work Sampling System (WSS)/Ounce.

Table 5

Reported Research Results

Author(s) and Year (<i>N</i> = 26)	Reported Results
Anthony (2003)	<u>TPBA</u> visual development guidelines were used by raters at Denver's PLAY clinic and had positive inter-rater agreement results.
Bailey & Bricker (1986)	<u>AEPS</u> correlation with the Gesell Developmental Schedule (Knobloch et al., 1980) was strong for the whole test, but not individual areas. AEPS could be successfully administered in a reasonable amount of time.
Baird, Campbell, Ingram, & Gomez (2001)	<u>DOCS</u> may lack sensitivity in detecting variations in development.
Bricker, Bailey, & Slentz (1990)	<u>AEPS</u> correlation across areas was $r = .88$ ($p < .001$).
Bricker, Yovanoff, Capt, & Allen (2003)	<u>AEPS</u> newly established cutoff scores in the 2 nd edition identified eligible children accurately most of the time.
Calhoon (1997)	Children performed better (i.e., higher scores) on the <u>TPBA</u> than the conventional test, and the play-based assessment provided a richer description of children's emerging skills.
Cody (1995)	In the <u>HELP</u> study, the play age obtained from the authentic assessment was highly correlated with the Developmental Age Equivalent of the conventional assessment (i.e., <u>BSID</u>).
DelGiudice, Brogna, Romano, Paludetto, & Toscano (2006)	After one year, children in the <u>Carolina</u> condition made progress and had higher DQ than children in the comparison condition who made slight progress but improvements were not statistically significant.
Di Pinto (2006)	<u>ABAS</u> accurately documents poor social adaptive outcomes for children with ADHD.
Fantuzzo, Grim, & Montes (2002)	The study supports the use of the <u>COR</u> assessment method for low-income urban preschool children, however a three factor model should replace the proposed six-factor model.
Friedli (1994)	<u>TPBA</u> had favorable test retest results, inter-rater agreement, and concurrent validity.
Gilbert (1997)	Significant difference among raters was found: mothers rated the child's skills highest, the fathers next, and the teachers last. Differences among raters on the <u>DOCS</u> may influence eligibility decisions.
Hsia (1993)	The <u>AEPS</u> has strong inter-rater agreement at both domain (ranging from .87 social to .94 adaptive) and total test (.90) levels. Strong relationship between individual domain scores (.64 to .96) and total test (.98) when internal consistency was examined. Findings also showed that the AEPS was sensitive to performance differences of children with delays.
Knox & Usen (1998)	The <u>PEDI</u> is a useful tool for describing the area of functional delay in children with cerebral palsy. It also appears to be sensitive to changes that were observed clinically.
Macy, Bricker, & Squires (2005)	The <u>AEPS</u> accurately classified all eligible children and over 94% ($n = 64/68$) of the non-eligible children. The overall sensitivity was 100%; specificity was 89%. The AEPS used to determine eligibility was positively and significantly correlated with conventional eligibility measures. Finally, the observers who scored the AEPS had strong agreement on observations made about child performance.
Meisels, Xue, & Shablott (in review)	Study found evidence for validity and reliability of <u>WSS</u> , suggesting that WSHS accurately assesses language development, literacy, and mathematics skills in young children.
McKeating-Esterle, Bagnato, Fevola, & Hawthorne (2007)	<u>ABAS-II</u> is correlated with ratings of informed opinion when assessing children for early intervention.
Morgan (2005)	Evidence supports the predictive validity of the <u>DOCS-II</u> in detecting RAD in a randomized sample.
Myers, McBride, & Peterson (1996)	The overall mean number of days to complete the eligibility assessment process took the group using the <u>TPBA</u> 22 days less than it took the group using a conventional test.

Noh (2005)	The <u>AEPS</u> has satisfactory inter-rater reliability agreement in the cognitive and social domains. Strong relationship between individual domain scores and items in the domains.
Sayers, Cowden, Newton, Warren, & Eason (1996)	When children's scores increased on the gross motor domain of the <u>HELP</u> , they did the same for the PSI.
Schweinhart, McNair, Barnes, & Lerner (1993)	The <u>COR</u> was found to be a psychometrically promising tool for the assessment of children's development in developmentally appropriate early childhood programs. Also, the COR helped staff to understand early childhood development and curriculum and to prepare individualized education programs for their children.
Sekina, & Fantuzzo (2005)	Univariate and multivariate results provide support for convergent and divergent validity of the COR dimensions. 15 of the 18 variables differentiated the three <u>COR</u> dimensions, particularly the COR Cognitive and Social Engagement dimensions.
Sher (2000)	Professionals using the <u>AEPS</u> were able to identify eligible children. Moderate interrater reliability for the communication domain and high reliability for other domains.
Slentz (1986)	Positive results of this study support the technical properties of the <u>AEPS</u> . Interrater agreement was very high at .94 for the entire test, and ranged from .84 to .94 for the six domains. Results for two administrations of the AEPS (N=18) revealed strong level of stability across time for the total test (.91), domain scores varied between high (fine motor .86, cognitive .91, social communication .77) to moderate (social .50) to low stability (gross motor .07 and self care .13). Internal consistency was strong for all domains except self care. Concurrent validity was examined by comparing the AEPS to the McCarthy (1972) and the Uniform Performance Assessment System (Haring, White, Edgar, Affleck, & Hayden, 1980) with mixed results ranging from very weak to strong relationships between domains and scales.
Wright & Boschen (1993)	Satisfactory information is provided to confirm the <u>PEDI's</u> usefulness for clinical and research purposes with children with Cerebral Palsy.

Note. Adaptive Behavior Assessment System (ABAS); Assessment Evaluation & Programming System (AEPS); Carolina Curriculum for Infants/Toddlers/Preschoolers with Special Needs (Carolina); Creative Curriculum for Preschool (CCP); Child Observation Record (COR); Developmental Observation Checklist System (DOCS); Hawaii Early Learning Profile (HELP); Pediatric Evaluation of Disability Inventory (PEDI); Transdisciplinary Play-Based Assessment (TPBA); Work Sampling System (WSS)/Ounce.