

***Accountability for Early Childhood Intervention:  
Aligning Professional Standards and Policies***

***Advocacy for Sensible and Equitable Outcomes***

Stephen J. Bagnato, Ed.D.  
Kristie Pretti-Frontczak, Ph.D.  
Marisa Macy, Ph.D.  
Jennifer Grisham-Brown, Ed.D.

***PROFESSIONAL POSITION STATEMENT***

**Issues and Conflicts**

The accountability movement associated with No Child Left Behind (NCLB) influences early childhood intervention in the form of a downward extension of a “tests and testing” model employed by school-age programs. Advocates for young children, while proponents of accountability, are concerned that existing models are detrimental not only to children, but to their families, and the programs and personnel who serve them.

Much controversy surrounds the testing of infants, toddlers, and preschoolers, particularly those with disabilities and delays. While the early childhood intervention field supports, generally, the need to monitor the progress of young children in diverse programs, little agreement exists on how desired information should be obtained, who should collect the information, and perhaps most importantly, how the information should be summarized and interpreted. Moreover, there is a dearth of research on accountability assessment practices in early childhood intervention. Much of the current efforts are driven by K-12 models or worse yet, appears to parallel earlier national accountability mandates under the National Reporting System (NRS) initiated by the Head Start Bureau. The now defunct NRS was the first national effort to collect uniform outcomes data in Head Start programs; much justified controversy accompanied the NRS as an “exemplar” of what not to do (i.e., engage in conventional testing procedures, focus on limited areas of development, and measurement with no relationship to curricular goals) when assessing for accountability purposes.

With increasing pressure, government agencies are requiring accountability data from programs serving young children (Harbin, Rous, & McLean, 2005). Much of the efforts in the form of regulations are being proposed and implemented without regard for professional “best practices”, usefulness and benefits to children and families, and the glaring absence of research. In particular, state and federal outcome indicators are emerging to document accountability.

Interdisciplinary professionals in the fields of early childhood intervention (i.e., public and private early care and education, Head Start, and early intervention) have an ethical and moral responsibility to advocate for caution and assurances that sanctioned professional standards will be honored when measurement strategies and standards for accountability and outcomes are designed and mandated by state and federal entities. The over 100,000 membership of our professional organizations (i.e., National Association for the

Education of Young Children; Division for Early Childhood of the Council for Exceptional Children; and the National Head Start Association) must have a strong voice and integral role in this effort.

While accountability methods and standards must meet professional standards, they must also be sensible and equitable. Policies must reflect the uniqueness and diversity of the early childhood intervention field (e.g., settings in which children spend time, education level of teachers), compared to school-age, and the individual needs of its vulnerable young children and families. It is impossible to craft such policies without the integral participation of professionals and researchers who work daily in this field. We advocate for such sensible, equitable, and professionally sanctioned policies and practices.

### **Professional Standards**

Over a 25-year period, the major national professional organizations in early childhood intervention have produced, published, and updated collaborative documents on recommended practice standards for the fields that cover all aspects of teaching and care of all infants, toddlers, and preschool children. These practice standards serve as the foundation for pre-service education of teachers and providers, for daily practice, and for certifying the quality of programs. Professional standards of practice in early childhood intervention distinguish the common and established values of our field and show an emerging applied evidence-base that validates adherence to their principles and practices.

Each of the professional organizations has produced specific cross-referenced practice standards about assessment, curriculum, and program evaluation for all young children. These standards drive our daily work with children and families and must, similarly, drive state and national policies and practices to document the progress of children and the impact of programs. Government policies and practices that ignore or run counter to these standards must be opposed and rejected by the field.

It is beyond scope of this position statement to detail each of these standards. Summarized below, however, are references to these standards and selected examples of their content:

- Bredekamp, S., & Copple, C (2009). *Developmentally appropriate practice in early childhood programs* (3<sup>rd</sup> ed.). Washington, DC: National Association for the Education of Young Children.
- Head Start Bureau. (1992). *Head Start program performance standards* (DHHS Publication No. ACF92-31131). Washing, DC: Department of Health and Human Services.
- NAEYC and NASDE (2003). *Early childhood curriculum, assessment and program evaluation: Building and effective and accountable system in programs for children birth to 8 years of age*. Washington, DC: National Association for the Education of Young Children.
- Sandall, S., Hemmeter, M.L., Smith, B., & McLean, M. (2005). *DEC recommended practices in early intervention/early childhood special education* (2nd ed.). Longmont, CO: Sopris West.

Below are selected examples of the recommended practice standards across purposes for assessment, curriculum, and progress/program evaluation, including accountability:

- Rely upon on developmental observations overtime
- Monitor performance during “authentic, not contrived, activities”
- Integrate and link assessment and curriculum
- Use past performances as the reference, not group norms

- Choose materials that accommodate the child's special functional needs
- Use only measures that have high treatment validity
- Rely on curriculum-based assessments as the foundation or "mutual language" for team assessments
- Defer a diagnosis until evaluation of a child's response to a tailored set of interventions
- Use scales with sufficient item density to detect even small increments of progress
- Use scales which emphasize functional and universal items
- Use scales which promote the inclusion of families in decision- making and infuse cultural aspects in the process
- Use scales which require data collection from multiple sources, settings, people and occasions
- Use scales which have been specifically designed and/or field-validated for specific early childhood intervention purposes (i.e., accountability) and with specific target samples of children

## **Policies**

State and national government policy mandates exist in the form of roughly similar state early learning indicators and, through the US Department of Education, Office of Special Education Programs (OSEP), outcomes for young children with disabilities (i.e., positive social relationships, acquisition of new knowledge and skills, and getting wants and needs met). To date, however, glaring problems exist with: (a) the design of standards or outcome indicators; (b) the fact that individual programs are bound by multiple standards; and (c) the lack of any widely available assessment(s) specifically designed and/or validated to monitor children's progress toward standards, particularly those with disabilities or diverse learning styles. Moreover, evaluation of program impact is not the same as monitoring child outcomes and requires creative designs before diverse early childhood intervention programs can be compared. Before the accountability movement in early childhood intervention becomes universal and codified in questionable policies and practices, early childhood professionals must present a unified "roadmap" to policymakers about the best way to proceed in order to reach a shared destination.

Some federal and state policies on accountability have been crafted without the integral participation of the field. Well-meaning, but misguided summary reports have been published recently which provide the illusion that reliable, valid, and responsive accountability and outcomes evaluation indicators and methods have been developed and should be used (PEW Foundation, 2008; National Academy of Sciences/National Research Council, 2009). Unfortunately, "testing" of our youngest children is the underlying message of these resources.

In the brief discussions below, we present guide-points on policy considerations and caveats for accountability and outcomes evaluation for early childhood intervention. Many of the points below highlight and expand the most important guide-points published previously in the position papers of NAEYC and DEC. The following guide-points operationalize our professional standards and relate them to what we believe should be "best practices" in accountability:

## **Best Practices in Accountability for Early Childhood Intervention**

### ***1. Young children are individuals, so, their programs and performance data must be individualized.***

The distinguishing characteristic of the field of early childhood intervention is that we focus on the strengths and needs of individual children rather than making broad group or age comparisons. At the base, we are a single-subject design field; in this respect, intra-individual progress is the most important criteria for significant change not inter-individual comparisons. Further, the more one aims to compare young children with differing abilities to a normative group, the less valid and trustworthy the conclusions; this fact makes accountability in early childhood, particularly for those children with disabilities, fundamentally different from school-age accountability standards. All young children should be entitled to individualized instruction that meets their unique learning needs. Even children who are at risk, English Language Learners, or with minor articulation concerns, may require individualized programs and their performance over time on family priorities must be the criteria for accountability. Thus, common outcomes should be universally acceptable for the diverse cultural, linguistic, and individual needs for all young children. Further, if individual performance is to be rated and otherwise documented and then aggregated, the sum should be seen as only as valid or trustworthy as its parts.

For children with disabilities, the goal of a programmatic intervention is not to ensure progress toward normality or a typical level of functioning. Rather, parents and professionals seek to document performance toward individual goals and to alter pre-intervention developmental trajectories. For children with significant disabilities, maintenance of performance or prevention of regression, not progress, is the goal of the intervention. All young children deserve performance criteria and measurement methods that are sensible and equitable:

*"Misrepresenting children by mis-measuring them denies children their rights to beneficial expectations and opportunities" (Neisworth & Bagnato, 2004, p. 198).*

### ***2. Authentic assessment is the most developmentally-appropriate form of measurement for documenting performance, progress, and outcomes in early childhood.***

The early childhood intervention field requires that professionals conduct authentic assessments of children to fulfill major purposes: profiling individual needs; planning individual goals, monitoring individual progress overtime; and documenting the impact of instruction and care (Bagnato, 2007; Meisels, 2007). Authentic assessment focuses upon curricular competencies that form the basis for linking or integrating assessment-instruction-progress evaluation. Authentic observations are recorded by familiar people, in familiar settings and routines, and emphasize the child displaying real-life skills.

No research exists on the specific use of conventional tests and testing for accountability purposes (Macy & Bagnato, 2007) based upon national research conducted through the federally funded TRACE Center for Excellence in Early Childhood Assessment (2002-2007). Some studies have been conducted on the use of authentic and curriculum-based assessment for accountability and other purposes (Macy & Bagnato, 2007; Bricker, et al., 2008). More research is needed in this important area.

**3. Accountability data cannot be interpreted in the absence of additional information about the child.**

A number of variables impact change in children's development and include prior exposure to intervention, regularity of child's participation and engagement in the program, and mediating factors (e.g., serious head injury between entry and exit data collection; uncontrolled seizure activity). As well, cultural expectations will impact behavioral changes in young children. Differing family ideas about when children should learn certain skills will likely impact how quickly children learn them. In addition, the age of the child must be considered. Younger children may show less apparent developmental delays than older children. Given the various ways in which change in child development can be affected, consideration should be given to how to define progress for individual children. For some children with disabilities, developmental changes in some areas are realistic goals. For others, progress may be defined as not acquiring additional disabilities or not regressing in development.

**4. Child progress data for accountability cannot be interpreted in the absence of data on the program, itself.**

Aggregated data on changes in children's acquisition of developmental competencies or changes in trajectory are meaningless unless related to aggregated data about the programs and services in which children participate. There must be a functional interrelationship between each child's patterns of progress and the type, quality, length, and intensity of their programs and the type of teaching and care strategies used. As well, the role of program providers must be considered in the context of analyzing accountability data. The type and amount of educational background of program providers may impact their capacity to deliver high quality interventions with fidelity sufficient to impact child change. Similarly, the consistency with which program providers collect data to measure child change must be considered (i.e., same provider collecting data; assessment fidelity). Larger program variables also have been found to effect change in children's development and should be considered when interpreting accountability data. These include the quality of the environment, the program's leadership, and family involvement in the program.

Other important variables include: (a) prior evidence of response to intervention; (b) the fidelity of intervention; (c) the use of different measures and people conducting the assessment of the child at various testing periods; and (d) the emergence of clear disabilities only after the child matures and fails to acquire critical skills such as pragmatic language. Federal mandates for response-to-intervention (RTI) models (e.g., Recognition and Response, Curriculum Framework, etc.) will transform our approaches in the future for documenting child progress as a result of program engagement. Specifically, RTI models will promote intra-individual change criteria and the functional relationship among child change and program elements rather than normative comparisons.

**5. Developmentally-appropriate accountability data must be used only to improve program quality and practices not to sanction teachers or their programs.**

States are reforming their assessment and evaluation policies to meet the federal mandates for IDEA. Specifically, the child outcomes identified by OSEP include: 1) positive socio-emotional skills, including social relationships; 2) acquisition and use of knowledge and skills; and 3) use of appropriate behaviors to meet needs. For accountability purposes, program personnel are required to assess children's performance

in these three areas near entry and again near exit (Hebbeler, Barton, & Mallik, 2008). Although state agencies have in place procedures for collecting accountability data following federal guidelines, the procedures are highly variable and generally unsubstantiated. For example, the legitimacy of interpreting children's performance near entry and/or exit with regards to the three OSEP child outcomes is open to question for several reasons.

First, each outcome is stated in broad language that makes valid and consistent measurement and comparison over time difficult if not impossible. That is, personnel, and the measures and procedures used may define/conceptualize the three outcomes in very different ways. A cursory comparison of crosswalks that have been created between commonly used assessment instruments and OSEP outcomes indicates startling variability among the specific sets of assessment items that are aligned with each OSEP outcome. Second, measures or procedures for data collection have not been carefully delineated, nor have any measures or procedures been developed for said purpose been adequately tested. Using different measures and collecting information in different ways may lead to child change data that are simply not comparable across children nor for any given child over time. Third, the validity of the two entry categories and the five exit categories is unknown. That is, the reliability of assigning children to these categories has not been investigated—meaning again that comparability across programs and states may be impossible and reported child change suspect.

The lack of empirical verification, in terms of both the validity and reliability, for interpreting and operationalizing the outcomes and the categories is of extreme concern, because critical decisions may rest on accountability findings (i.e., future funding of Part C and 619 programs). Accountability data should represent *developmental performance*, not necessarily developmental progress. Thus, accountability data must not be used inappropriately as an excuse or punishment for professionals, their programs, agencies, or states supplying IDEA services to young children and their families. Safeguard procedures need to be implemented for states and programs that do not meet performance expectations. Individualized professional development and mentoring of teachers must be improved by making accountability data available to teachers and supervisors and by ensuring access to high-quality state technical assistance.

#### ***6. Providers using authentic assessment methods for accountability purposes must be well-trained.***

Obtaining valid, reliable assessment information on young children is considered an important recommended practice (McLean, Wolery, & Bailey, 2004) and is especially important when the data being collected are for high-stakes accountability purposes. In order to ensure that valid, reliable data are collected, it is essential that providers receive adequate training on how to collect assessment information using authentic assessment methods (e.g., observation) (Grisham-Brown, Hallam, & Brookshire, 2006; Neisworth & Bagnato, 2004:). Specifically, providers need training on how to observe children's behaviors and compare those behaviors against some criteria to make judgments about the observed performance (i.e., does the child need further instruction on the skill? or is the child functionally performing the skill?). As well, providers need training on how to implement authentic assessment practices with fidelity (e.g., embedding assessment opportunities into child-initiated activities).

Moreover, training must be sufficient so that providers can collect reliable data using authentic assessment methods. In order to obtain this goal, recommended professional development practices should be employed including training over time, action planning, and classroom follow-up. There is some evidence that this package of practices results in providers ability to collect reliable assessment data with fidelity

(Grisham-Brown, Hallam, & Pretti-Frontczak, 2008). Similarly, well-trained providers are more likely to feel confident in implementing authentic assessment practices (Keilty, LaRocco, & Casell, 2009).

***7. Metrics for profiling child progress and program impact must be sensitive to small increments of individual child performances.***

Standard scores on conventional tests are not sensitive to individual patterns of progress in young children, especially those with disabilities and functional limitations. In contrast, metrics that compare each child's progress to his individual pre-intervention starting point are most sensitive to true progress (i.e., changes in performance over time). Such metrics include expected-actual developmental growth curves, goal-attainment scaling, number of curricular objectives achieved, increases or decreases in the frequency of particular behaviors, and number of skills displayed with and without prompts. All of these metrics are more consistent with a single-subject design approach as recommended previously. Perhaps, most important is the fact that progress metrics must focus upon tangible ultimate criterion standards such as the acquisition of functional competencies that improve independent life functioning, performance, and learning, rather than dubious normative comparisons to non-representative standardization samples.

***8. Measures chosen for accountability purposes must meet developmentally-appropriate professional standards for evidence-based practice.***

Bagnato and colleagues (1997; 2009) have published 8 overarching standards by which consumers can critically examine the developmentally-appropriate quality of an assessment measure considered for accountability and other purposes in early childhood intervention. The assessment standards include: acceptability, authenticity, collaboration, evidence, multi-factors, sensitivity, universality, and utility. Simply, assessment for early childhood intervention must be natural/familiar, involve multiple people working together, truly reflect performance or change, accommodate all children, and be useful for instruction. Additionally, reporting requirements for IDEA Part C (i.e., infants and toddlers) and Part B/Section 619 (preschoolers) must include the collection of information from multiple sources.

Perhaps, however, the most important consideration is that the measures or methods used must have been designed, specifically, for the children being assessed (i.e., age, disability, ethnicity) and validated for the specific purpose of accountability.

## References

- Bagnato, S. J. (2007). *Authentic assessment for early childhood intervention: Best practices*. New York, NY: Guilford Press, Inc.
- Bagnato, S.J., Neisworth J.T., Pretti-Frontczak, K. (2009). *LINKing authentic assessment and early childhood intervention: Best measures for best practices*. (4th Edition). Baltimore, MD, Paul Brookes.
- Bredenkamp, S, & Copple, C (2009). *Developmentally appropriate practice in early childhood programs* (3<sup>rd</sup> ed.). Washington, DC: National Association for the Education of Young Children.
- Bricker, D., Clifford, J., Yovanoff, P., Pretti-Frontczak, K., Waddell, M., Allen, D., & Hoselton, R. (2008). Eligibility determination using a curriculum-based assessment: A further examination. *Journal of Early Intervention, 31*(1), 3-21.
- DEC (2008). *Promoting positive progress outcomes for children with disabilities: Recommendations for curriculum, assessment, and program evaluation*. Missoula, MT: Division for Early Childhood.
- Downs, A., & Strand, P. (2006). Using assessment to improve the effectiveness of early childhood education. *Journal of Child and Family Studies, 15*(6), 671-680.
- Grisham-Brown, J., Hallam, R., & Pretti-Frontczak, K. (2008). Preparing Head Start personnel to use a curriculum-based assessment: An innovative practice in the "age of accountability." *Journal of Early Intervention, 30*(4), 271-281.
- 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.
- Greenwood, C. R., Walker, D., Hornbeck, M., Hebbeler, K., & Spiker, D. (2007). Progress developing the Kansas early childhood special education accountability system: Initial findings using ECO and COSF. *Topics in Early Childhood Special Education, 27*(1), 2-18.
- Harbin, G., Rous, B., & McLean, M. (2005). Issues in designing state accountability systems. *Journal of Early Intervention, 27*(3), 137-164.
- Head Start Bureau. (1992). *Head Start program performance standards* (DHHS Publication No. ACF92-31131). Washing, DC: Department of Health and Human Services.
- Hebbeler, K., Barton, L. R., & Mallik, S. (2008). Assessment and accountability for programs serving young children with disabilities. *Exceptionality, 16*(1), 48-63.
- Keilty, B., LaRocco, D. J., & Casell, F. B. (2009). Early interventionists reports of authentic assessment methods through focus group research. *Topics in Early Childhood Special Education, 28*(4), 244-56.
- Macy, M., & Bagnato, S. J. (in progress). *Authentic assessment for early intervention eligibility determination: Research & practice foundations*.
- McLean, M., Wolery, M., & Bailey, D. B. (Eds.) (2004). *Assessing infants and preschoolers with special needs*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Meisels, S.J. (2007). Accountability in early childhood: No easy answers. In R. C. Pianta, M. J. Cox, and K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 31-47). Baltimore: Brookes.
- NAEYC & NASDE (2003). *Early childhood curriculum, assessment and program evaluation: Building and effective and accountable system in programs for children birth to 8 years of age*. Washington, DC: National Association for the Education of Young Children.
- National Research Council. (2008). *Early childhood assessment: Why, what, and how*. Committee on Developmental Outcomes and Assessments for Young Children, Catherine E. Snow and Susan B. Van Hemel (Eds.). Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

- National Early Childhood Accountability Task Force (2007). Taking stock: Assessing and improving early childhood learning and program quality. *The PEW Charitable Trusts*. Retrieved February 8, 2009, from [http://www.pewtrusts.org/uploadedFiles/wwwpewtrustsorg/Reports/Pre-k\\_education/task\\_force\\_report1.pdf](http://www.pewtrusts.org/uploadedFiles/wwwpewtrustsorg/Reports/Pre-k_education/task_force_report1.pdf).
- Neisworth, J.T., & Bagnato, S.J. (2004). The mismeasure of young children: The authentic assessment alternative. *Infants and Young Children, 17*(3), 198-212.
- Rous, B., McCormick, K., Gooden, C., & Townley, K.F. (2007). Kentucky's early childhood continuous assessment and accountability system: Local decisions and state supports. *Topics in Early Childhood Special Education, 27*(1), 19-33.
- Sandall, S., Hemmeter, M.L., Smith, B.J., & McLean, M.E. (Eds) (2005). DEC-recommended practices: A comprehensive guide for practical application in early intervention/early childhood special education. Longmont: Sopris West.