

Scott R. McConnell and Naomi L. Rahn

Assessment

Assessment has long been acknowledged as an essential feature in the provision of early childhood special education (ECSE) services. In this historical position, assessment serves as a major element of eligibility determination (Salvia, Ysseldyke, & Bolt, 2012) and a required element of intervention planning and ongoing evaluation for children and families (McLean, Wolery, & Bailey, 2003). However, in recent years rapid and significant growth has occurred in the sophistication of assessment practices available for young children with disabilities and their families and broader application of these practices in a wide range of settings. The purpose of this chapter is to provide an overview of this historical role and to highlight features of the emerging uses of assessment as a central feature of high-quality services for infants, toddlers, and preschool children with disabilities, their families, and those who serve them.

S.R. McConnell (✉)
Department of Educational Psychology, University of
Minnesota, 56 East River Road, Delivery Code: 4101,
Minneapolis, MN 55455, USA
e-mail: smcconne@umn.edu

N.L. Rahn
College of Education and Human Services, West
Virginia University, Morgantown, WV, USA

Here we make a case for an expanded view of assessment as an essential part of the system that contributes to improved outcomes for children. We argue that assessment practices, carefully aligned and intertwined with formal and informal services and supports, increase the efficacy and efficiency of societal efforts to promote young children's development. Given this approach to assessment, we describe basic features of assessment, including its core purpose and functions and the quality standards for its use, and review four major functions of assessment in early intervention and ECSE: eligibility and identification for specialized intervention, program planning, intervention and fidelity assessment, and progress monitoring.

Assessment's Historical Role and Transition to Contemporary Practice

At the outset of formal services for young children with disabilities, from the 1950s through the 1980s, assessment in early intervention and ECSE focused primarily on individual child description and eligibility evaluation (McLean et al., 2003). Formal and well-evaluated assessment tools are available for this purpose. These tools reflect both comprehensive and broad models of child development and performance such as the *McCarthy Scales of Children's Abilities* (McCarthy, 1970), *Stanford-Binet Intelligence*

Scale (Terman & Merrill, 1972; Thorndike, Hagen, & Sattler, 1986), and *Wechsler Preschool and Primary Intelligence Scale* (Wechsler, 1963, 1967) or more specific criterion-referenced or curriculum-embedded measures of child development such as the *Hawaii Early Learning Profile* (Furuno et al., 1979) or the *Carolina Curriculum* (Johnson-Martin, Jens, Attermeier, & Hacker, 1986).

These assessments, and others like them, became common in part because of federal and state legislation and regulation creating and expanding legal mandates for services to young children with disabilities. Perhaps most noteworthy was early attention to assessment for child find and eligibility determination; from the outset, preschool special education services could only be provided following determination of either an established disability or “developmental delay” (i.e., when a child’s assessed development varies significantly from normative expectations). To meet the needs of teams making this latter determination, researchers and commercial publishers released a new generation of measures to provide norm-referenced assessment across multiple domains (e.g., *Battelle Developmental Inventory*; Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984/1988). While federal and state law and regulations required formal description of intervention targets and assessment of child progress toward them (i.e., as goals and objectives on Individualized Family Service Plans [IFSPs] or Individual Education Plans [IEPs]), this aspect of special education and related services was generally teacher- or therapist- and child-specific, and less dependent on formal practices.

Yet over time, growing evidence of relations between preschool skills and development and later academic, social, and other performance led to continued research and development of refined intervention procedures for infants and preschoolers (Ramey & Ramey, 1998). This led to the development and increased use of assessments that helped specify intervention targets and substantially increased formal attention to assessment and monitoring of the development of individual children receiving specialized supports (e.g., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2002).

Defining Assessment

While the border between *assessment* and *intervention* continues to blur (Bagnato, Neisworth, & Pretti-Frontczak, 2010), assessment’s core features remain consistent. Similar to others (c.f., Anastasi, 1988; McLean et al., 2003; Salvia et al., 2012), we define assessment as *the systematic collection and evaluation of information to determine what, if anything, to do differently*.

Assessment occurs in a wide variety of situations and serves as an essential guide for design, selection, application, and evaluation of intervention practices. “Systematic collection . . . of information” makes clear the intentional, purposeful, and rigorous features of any assessment practice. While the collection can occur in many different ways (e.g., teacher ratings, observational assessments, or evaluations of responses to items on standardized tests), “systematic” means that conditions and procedures for data collection are specified, known, and generally common across time, individuals, and settings.

“Systematic . . . evaluation of information” speaks to specified procedures, rules, and practices for considering collected information and, against some standard or a priori criterion, producing some judgment or appraisal. This step is critical for turning data into actionable information. For many, this element of any assessment is core to its value (Messick, 1990).

“Determining what, if anything, to do differently” follows naturally from this step and brings practical utility to any assessment-based judgment. Specific questions (e.g., does this child’s performance vary from criterion or expectation?) often direct assessments: the answers should lead to actions that bring social utility to the assessment.¹ In particular, determining what to do differently should lead to determining if current services, supports, or conditions are appropriate or if different actions will better serve the individual(s).

¹ This element of our definition, calling for determination of practical actions, may differentiate our perspective from others. We believe this practical determination is a central feature to assessment that contributes to improved outcomes rather than data collection and evaluation that produce nothing more than “assessment reports.”

This last feature, using data to determine specific courses of action, expands existing definitions that focus solely on collecting and analyzing information; here, a central expectation is that assessment contributes to monitoring and improving outcomes. Our system of educational and developmental services, including special education and related services, is predicated on the connected ideas that individual developmental achievement is driven (at least in part) by experience and interaction, and that variations from expected rates of achievement can (and often, must) lead to changes in these experiences and interactions to more successfully support a child's development. By focusing on the status and development of a child *and* the services and supports related to observed development, the central role and contributing features of assessment in a larger system of "special education and related services" become clear.

This notion that assessment practices can contribute to improved outcomes highlights at least three features of a more contemporary set of practices. First, assessment should help identify children who will benefit from additional, or different, services and supports. This identification may be global, as in determining eligibility for special education and related services and providing procedural safeguards; conversely, the identification may be narrow, specific, and dynamic, as in the screening practices of multi-tiered systems of support (Greenwood, Carta, & McConnell, 2011). Second, contemporary practices must provide ongoing dynamic descriptions of children's progress and related practices: we must monitor how children develop and use this information frequently to adjust the intensity, focus, or other features of intervention. Third, contemporary assessment practice must closely align with intervention, which in turn must closely align with intended developmental outcomes. "Intervention" is not a general, nonspecific action but a set of specific practices that produce particular outcomes; thus, assessment practices must guide the selection and use of intervention practices to promote intended outcomes for individuals and groups of children. These three features—identification, monitoring, and alignment with intended

outcomes—maximize the efficacy and efficiency of assessment practices and their contribution to improved outcomes for children.

Basics of Assessment

Purposes of Assessment

Assessment as defined here can occur for various purposes, and characteristics of assessment will vary due to these different purposes. For instance, some assessments in early intervention and ECSE are completed to meet federal, state, or local statutory or regulatory requirements; US regulations require that states develop procedures for determining possible eligibility for early intervention or ECSE, ensuring "a timely, comprehensive, multidisciplinary evaluation of each child, birth through age two, referred for evaluation, including assessment activities related to the child and the child's family" (34 CFR 303.322). In these cases, many assessment features (instruments, characteristics, administrators' qualifications, and interpretation or evaluation standards) may be specified in advance.

Other times, assessments will take place to inform teachers in the design or evaluation of instructional and other services or to apprise parents and others of children's growth or progress. In these instances, assessment practices are likely to be more specific, more varied across programs and time, and with fewer a priori evaluation or interpretation standards.

Assessments can help describe groups for program planning and evaluation, policy analysis and development, or communication with the public. Kindergarten entry assessments, like those developed as part of federal Race to the Top-Early Learning Challenge Grants to states, will provide "lagging indicator" information about the success of early childhood policy and practice innovations in states and local jurisdictions that will be useful in evaluating and refining early childhood services. Similarly, the *Early Development Instrument* (Janus et al., 2007; Janus & Offord, 2007) was developed as a "leading indicator" to describe the school readiness of

groups of children (in a neighborhood, city, state, or province).

Evaluating, selecting, and implementing any assessment must begin with a clear statement of intent: *who* is the focus—an individual child or family, that child or family in a particular situation or context, or a group of children? What are the *areas of development* that must be assessed? How will the data be *used*? What *degree of confidence and precision* is needed?

Functions of Assessment

These purposes can be sorted and classified as different assessment *functions* (Hawkins, 1979). In most cases, ECSE assessment practices are directed to individual children; at this level, we can identify five primary functions. First, assessments can *identify children who may benefit from additional or compensatory services and determine their eligibility for special education and/or other related services*. This type of assessment (“diagnosis”) is often the first part of a longer sequence for one child and is often a requirement for providing early intervention or ECSE.

Next, children will encounter *assessment for program planning*. Due to the requirements of individualized educational planning in special education and the low fidelity of current diagnostic assessments in education and development, detailed information is needed to determine *what* to teach (and often *how best to teach it*) to children needing preventive or compensatory services. This function of assessment can sometimes benefit from information typically gathered in identification and eligibility determination but generally will also require information that is more detailed, specific, and narrow in scope.

Third, given contemporary problem-solving models (Tilly, 2002) and the challenges in implementing even well-specified interventions (Cook & Odom, 2013), there is increasing attention to *assessment and monitoring of interventions*. This is critical in implementing some interventions and provides vital information during review or revision of intervention services.

Last, best practices require ongoing *assessment of intervention efficacy or monitoring progress toward desired outcomes*. By definition, ECSE is provided in cases where typical experience does not produce desired developmental outcomes: a different type or amount of experience is assumed to be necessary to promote that development. To assure that plans to support development work, ongoing assessment of child growth and development over time (and related evaluation of the rate of change in observed developmental achievement) will determine what (if anything) to change.

Quality Standards for Assessment

To meet statutory and best practice standards, assessment practices must be rigorous and applied with care. However, standards for high-quality assessment vary somewhat by function or purpose. To paraphrase a truism, there are no “good” assessment instruments or practices: rather, we evaluate them for their particular function and intended use.

Reliability. Historically, assessment instruments were evaluated for reliability and validity. *Reliability* is the extent to which any assessment product or score is *trustworthy*, or “the consistency of scores obtained by the same persons when reexamined on different occasions, or with different sets of equivalent items, or under other variable examining conditions” (Anastasi, 1988, p. 109). Measurement, in any case, includes information (the *true score*) and random noise (*error*). If one person takes two measures of the same thing, one after another, some minor difference in results will appear. The size and characteristics of this difference define reliability.

While many forms of reliability exist in the measurement literature, in ECSE reliability is typically evaluated in three ways. First, *internal consistency* describes the empirical cohesiveness of a set of items or scores added into one sum. To the extent that different items correlate with one another, or different assessment instances sample the same broad idea or construct, internal consistency increases.

Second, *interrater* or *interobserver agreement* describes the extent to which two independent assessors score a particular behavior or response similarly. Many early childhood assessments rely on direct observation of discrete child behaviors and classroom events or on slightly more summative ratings completed by teachers or others of child performance, activities, or interactions. When assessments occur this way, it is critical to evaluate the degree to which results reflect characteristics of the assessed content rather than the standards and judgments, or biases, of individuals collecting the data. As people's judgments vary from one another, interrater agreement estimates decline.

Third, ECSE assessment is often evaluated for *temporal consistency* or *test-retest reliability*. Many traditional assessment instruments assume that child characteristics remain relatively stable. While this is not always true in ECSE (Kazdin, 1979; Strain et al., 1992), particularly when monitoring progress or if child behavior is highly variable across situations or conditions, there are many instances where consumers find it important that assessment results appropriately represent a child's achievement or performance at a particular point in time. This is critically important in many normative comparisons, where (for both the child being assessed and scores coming from participants in the norming samples) teachers, parents, and others can assume that observed assessment results minimize any variation due to the day or time of assessment.

Validity. Validity is the meaningfulness of any assessment or information about “*what* the test measures and *how well* it does so” (Anastasi, 1988, p. 139, emphasis in original). Cronbach (1990, p. 145) describes validity as the result of “inquiry into the soundness of the interpretations proposed” from any assessment. Validity studies provide evidence of the extent to which results of assessment can be useful, as generally intended by test developers or for the purposes in a particular case. Unlike most measures of reliability, where explicit standards for evaluating and selecting assessments exist, validity standards are somewhat more general and idiosyncratic to the

purpose of assessment and sometimes the instrument or practice itself (Messick, 1995).

Traditionally, three types of validity are discussed (c.f., Anastasi, 1988). *Content validity* describes the extent to which a particular assessment samples the behaviors, situations, and/or interactions of interest for a particular assessment purpose. Content validity is often determined logically by examining both items and conditions of any assessment and comparing these to generalizations teachers or others might want to make from assessment results. *Criterion validity* evaluates “the effectiveness of a test in predicting an individual's performance in specified activities” (Anastasi, 1980, p. 145), including future tests and life events. Examples might include the extent to which a measure of language and early literacy collected in preschool predicts reading performance in early elementary school or the degree to which a measure of developmental risk status collected at age 3 predicts later performance on a kindergarten entry assessment. Third, *construct validity* is “the extent to which the test may be used to measure a theoretical construct or trait” (Anastasi, 1980, p. 153). While construct validation is often more abstract (c.f., MacCorquodale & Meehl, 1948), the core idea (e.g., “is this assessment measuring what it purports to be measuring?”) pervades all validity discussions, and may be particularly important in diagnostic assessment activities, where statutory or professional standards call for assessment of broadly defined domains like “developmental delay” or “kindergarten readiness.”

Another validity standard is emerging in ECSE. *Treatment validity* (similar to *consequential validity*; Messick, 1988) is the extent to which an assessment practice informs or helps design instructional or other intervention services that produce meaningful changes in an individual's (or group's) performance. Treatment validity can be assessed over a short period of time, as when an assessment practice is used to identify or refine instructional practices or intervention options in a child's current educational programming, or over longer periods of time, as when identifying changes in services or placements to

benefit child development and achievement (Connor et al., 2009).

Modern interpretations. In the latter half of the twentieth century, academic psychometricians developed and deployed new logical and analytic models to the design, evaluation, and use of tests. This “modern test theory” approach, most notably item response theory (IRT: Embretson & Reise, 2000; Wilson, 2005), is now commonplace in education generally and ECSE specifically (Greenwood, Carta et al., 2011; Greenwood & McConnell, 2011; Rodriguez, 2010).

While many details of IRT differ markedly from older approaches, core concepts remain the same. In general, assessments should be *stable or consistent* to reflect a child’s performance or development, *meaningful and well specified* as to child or other characteristics, and *useful* in choosing short- or long-term changes in intervention and services.

While technical features of test and measurement construction continue to develop, offering new and perhaps better tools for building and evaluating assessment practices, many of the functional requirements and bases for evaluating them remain the same. End users can, and should, have information to help evaluate the precision, trustworthiness, meaningfulness, and utility of information.

Assessment for Identification and Eligibility

Definition

A common, widely understood, measurement task in ECSE is to identify children who may benefit from additional or compensatory services, including determination of eligibility for special education and/or other related services. In current practice, this stage combines two otherwise distinct functions, *screening* and *diagnosis*. Technically, *screening* refers to a brief, broadscale evaluation used to identify individual cases when more information is needed. In current multi-tiered systems of support, screening often refers to universal and repeated assess-

ment of developmental achievement or growth in a larger group (e.g., classroom, school, program, or community), with the results used to identify those who would benefit from more intensive or supplemental intervention (Christ & Nelson, 2014; Greenwood, Carta et al., 2011). In ECSE, *diagnosis* refers to formal evaluation of individual children for special education eligibility or program enrollment. Diagnostic assessment is typically broad in scope and includes direct or indirect assessment of child status, characteristics, developmental trajectory and, often, environmental supports, comparing this to normative or other a priori standards or criteria and identification of individuals, based on this comparison, who are deemed to meet inclusion criteria for particular program services or procedural supports and safeguards (McLean et al., 2003).

Examples

Screening. McConnell, Wackerle-Hollman, and Bradfield (2014) describe screening practices to identify individual children for intervention in language and early literacy multi-tiered systems of support. These systems are increasingly common in early childhood programs serving children with and without disabilities and can identify individual children who are not acquiring age- or domain-appropriate skills related to the long-term achievement of reading proficiency so that these children can receive further help to improve their achievement (Carta et al., 2016).

The Center for Response to Intervention in Early Childhood (Bradfield, Vue, Rodriguez, & McConnell, 2014; Bradfield, Wackerle-Hollman, Albano, Rodriguez, & McConnell, 2014; Wackerle-Hollman, Schmitt, Bradfield, Rodriguez, & McConnell, 2015) has developed one universal screening approach in this area. In this model, Individual Growth and Development Indicators of four domains of language and early literacy development (i.e., oral language, phonological awareness, alphabet knowledge, and comprehension) have been developed, evaluated, and constructed into three measures, designed to be completed in Fall, Winter, or Spring in the year

before kindergarten. Each consists of 15 items, selected specifically to increase sensitivity to identifying lower-performing students in that season. Adult examiners assess all children in a classroom, comparing individual students' scores to empirically derived "cut scores," or scores for each measure that best identify children who are performing well below expectation and likely to benefit from supplemental instruction.

Evidence from early research and development (Bradfield, Wackerle-Hollman et al., 2014; Wackerle-Hollman et al., 2015) suggests that seasonal screening measures are trustworthy and closely associated with established, standardized measures of language and early literacy. They also identify children in ways that match with teachers' clinical judgments of their needs and intervention candidacy based on longer standardized measures. Classification accuracy can be improved by gathering additional information from teachers (Bradfield, Vue et al., 2014). Research on this model of screening for supplemental intervention proceeds (McConnell, Wackerle-Hollman, Roloff, & Rodriguez, 2014), including expansion to Spanish-speaking preschool children (Wackerle-Hollman et al., 2012).

Screening also occurs in community-wide child-find efforts to identify those who may benefit from early intervention for infants and toddlers with disabilities and ECSE for preschoolers with disabilities and delays. The *Assuring Better Child Health and Development (ABCD) Program*, funded by the Commonwealth Fund and administered by the National Academy for State Health Policy, is an integrated state-level approach to screening for a broad array of health and development concerns (Pelletier & Abrams, 2003). Now implemented in over 25 states, ABCD aligns state policy, funding, and practice development to increase the frequency of screening in primary health care and other settings and to better coordinate response to screening results (National Academy for State Health Policy, 2014). While intentionally varied to reflect unique policy and practice contexts in different states, ABCD focuses on broad, consistent use of a small set of evidence-based screening instruments, along with public/private quality improvement partnerships

to review and improve screening penetration and follow-up. A variety of practices have emerged at state and local levels (Earls & Hay, 2006; Pelletier & Abrams, 2003), but challenges remain in increasing universal screening rates and in using these practices to refer children to more extensive evaluation (Bethell, Reuland, Schor, Abrahms, & Halfon, 2011; King et al., 2010).

Diagnosis. Importantly, ECSE diagnosis is the "front door" to an array of special education safeguards and assurances, coordinated services, and opportunities for effective services and developmental achievement that support future competence. Perhaps due to its central role in providing special education, this aspect of assessment has a long, rich, and well-developed matrix of research, measure development, and practical implementation (Division for Early Childhood, 2014; Shonkoff & Meisels, 1990).

In 2014, the Division for Early Childhood (DEC) of the Council for Exceptional Children published a fully revised compendium of best practices in ECSE, *DEC Recommended Practices in Early Intervention/Early Childhood Special Education*. This document presents "best-in-class" procedures that are supported by research and consistent with the values and expectations of contemporary practice, observable and actionable in practice, applicable to all children with and without disabilities, and appropriate for use in a wide array of settings and situations.

Seven of these best practices describe assessment for diagnosis. In particular, diagnostic assessment should (1) adapt planned assessment practices, to the extent possible, to follow families' preferences for this work; (2) employ a team of professionals and family members to collect and evaluate assessment data; (3) employ assessment practices that are appropriate, given the child's age, individual characteristics, family, and life circumstances; (4) conduct a broadscale assessment that describes "the child's strengths, needs, preferences, and interests; (5) conduct assessment in the child's preferred and most fully developed language or communication system; and (6) report the results in ways that both address a priori evaluation standards for eligibility determination and that are understandable and useful

for families and other professionals. Finally, consistent with other recommendations (American Educational Research Association et al., 2002), when making important decisions that may affect individuals and resources, "...practitioners [must] use a variety of methods, including observations and interviews, to gather information from multiple sources, including the child's family and other significant individuals in the child's life" (DEC, 2014, p. 7).

Key Features of Quality Implementation

Practices in this domain of assessment for identification and eligibility vary widely, as do the questions that assessment and evaluation reports can answer. Across all these practices, three key features of quality implementation can be described.

First, it is essential that all assessment instruments and practices be selected for their "fit" to the evaluation purpose and for the questions at hand. Second, assessment instruments and practices must be easy to understand and must meet appropriate standards of rigor for planned uses. Third, assessment for identification and eligibility should be efficient, both for the immediate task and for subsequent services and supports to children and families. As *DEC Best Practices* suggests, any assessment should be appropriate for the questions to be addressed and the individual case at hand.

Assessment for Program Planning

Definition

Assessment for program planning identifies gaps in a child's skills and competencies and areas of need and provides the information needed to identify instructional goals and objectives and to inform teaching methods for immediate intervention use. The assessment compares a child's skills to a set of skills considered important for participating in age-appropriate activities. For infants,

toddlers, and preschoolers, these skills might come from a criterion-referenced assessment such as the *Assessment, Evaluation, and Programming System* (AEPS, Bricker, 2002) or the *Hawaii Early Learning Profile* (HELP; Furuno et al., 1979). Teams might also use state early learning standards or program-specific frameworks (e.g., The Head Start Child Development and Early Learning Framework, U.S. Department of Education, 2011) as a basis for examining children's skills compared to expectations for their age level. For early elementary students, comparison skills typically come from standards such as the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) or state-developed standards (c.f., Texas, Minnesota) that focus on basic academic skills.

Assessment for program planning in ECSE can be characterized by five principles. First, assessment is directly linked to intervention. Information collected and analyzed during this part of the process leads directly to the use of that information for planning individual children's intervention programs (Bagnato et al., 2010; Bricker, 2002; Macy, Bricker, & Squires, 2005). The field of EI/ECSE has long linked assessment and intervention (McConnell, 2000). Assessment must focus on skills important for participating in home, classroom, and community activities that are expected given the child's age and developmental level. Similarly, the assessment outcomes should inform the intervention, with the team teaching skills the child needs.

Second, this assessment examines the child's use of functional or useful skills important for participation in authentic contexts (Neisworth & Bagnato, 2005; Snyder, Wixson, Talapatra, & Roach, 2008). Traditionally, diagnostic assessment occurs in the context of unnatural and even strange situations and often focuses on skills of questionable importance for the child (Bronfenbrenner, 1977). For example, a traditional standardized assessment might require that we assess an infant's grasping and releasing skills. In contrast, an authentic programmatic assessment might instead assess the infant's

ability to stack foam blocks during play with a parent (i.e., an authentic, typical activity for a young child).

Third, assessment for program planning is flexible and adaptable for children with varying needs (Snyder et al., 2008) to ensure equity and “fit” to the individual case (Neisworth & Bagnato, 2005). Typically when administering standardized assessments, all items are administered in the same way and do not allow for modifications. In contrast, when assessing for program planning, the user can modify procedures and materials to accommodate the child’s skills and needs. For example, for a child who is deaf, items might be administered using spoken language and American Sign Language to fully assess functional communication skills.

Fourth, the assessment process is family friendly (Bagnato et al., 2010; DEC, 2014; Grisham-Brown & Pretti-Frontczak, 2011). During early childhood, parents and other family members are integral to a child’s life. Young children typically spend a large part of their day with mom, dad, a grandparent, or other primary caregiver. These relationships are critical for healthy social-emotional development, and caregivers play an important role in helping the child learn new skills. In addition, family members know their child best. Teams gather information from families and listen to their concerns to gain an accurate picture of a child’s strengths and needs.

Fifth, the process is dynamic or repeated, collecting information in an ongoing way to provide information about the child’s skills across time to inform programming decisions (DEC, 2014; Grisham-Brown & Pretti-Frontczak, 2011). Rather than administering a criterion-referenced assessment only once when the child is first eligible for services, teams collect ongoing information by reassessing the child’s skills to inform changes to intervention.

Program planning assessment should focus on skills essential for the child’s success within everyday routines and activities in the home, school, and community. For infants and toddlers, this will often include gross motor, early communication, and self-care in home or care settings. For preschoolers the scope may expand to

include social interaction and pre-academic skills required in inclusive preschool, home, and community settings.

Examples

Given the vast changes in children’s development during early childhood and shifting emphases in programs as they begin school, assessment for planning intervention varies depending on the child’s age. For infants and toddlers with special needs, the focus should be on supporting the family to maximize outcomes (Bruder, 2010; Johnson, Rahn, & Bricker, 2015). Both the Routines-Based Interview (McWilliam, 2003) and the AEPS Family Report (Bricker, 2002) provide a format to gather information about the child’s daily routines and activities (e.g., meals, bathing, and dressing) and family preferences regarding intervention targets. This assists the team in identifying desired outcomes important for the child’s functioning during everyday activities. Intervention strategies can then be embedded within authentic activities and routines using toys and materials available in the home.

In preschool, a child-centered approach is more directly informed by the current and future demands of academic and behavioral school success. The AEPS (Bricker, 2002) is an example of a comprehensive tool used to identify the child’s skills. AEPS items are clustered in developmental areas and arranged hierarchically, from earlier to later developing skills. Each item includes a subset of objectives that develop earlier than the goal and generally precede it. This allows teams to identify where a child lies in the developmental sequence and which skills need focus. Information gathered from both teacher and parent reports are used to identify and set priorities for skills to address in intervention. These skills are taught using evidence-based teaching strategies selected based on the skills being taught and the child’s strengths, needs, and preferences, with instruction and practice embedded across various activities (Wolery, 2005).

In the elementary setting, where demands of the curriculum are more dominant, assessment

for those with academic needs should be individualized based on those needs (Fuchs, Fuchs, & Compton, 2012). Within an RTI framework, assessments include curriculum-based measurement (CBM; Deno, 1985, 2003) or other similar measures (e.g., DIBELS; Good & Kaminski, 2000) that provide information about critical skills. Some children will have needs outside of academics requiring specialized assessment (e.g., a functional behavioral assessment [FBA] for significant behavior issues; c.f. O'Neill, Albin, Storey, Horner, & Sprague, 2014). Based on results, teams then design interventions to address specific skills using evidence-based interventions that could include published curricula and specific evidence-based teaching strategies (see, e.g., Wong et al., 2014).

Key Features of Quality Implementation

Assessment for program planning must reflect the five key principles for this function of assessment. First, the content of assessment must be linked closely to intervention (Neisworth & Bagnato, 2005; Snyder et al., 2008). Information gathered from the tool should allow the child's team to make decisions in two key areas: (1) which skills to target and (2) which teaching methods and strategies to use to meet those targets.

Second, assessment must examine functional skills in authentic contexts. Assessments should focus on gathering information on skills important or essential to children's participation in the home, classroom, and/or community. Assessments are often administered within typical activities, and use toys and materials appropriate for the child's developmental level and available in the child's environment (Snyder et al., 2008; Neisworth & Bagnato, 2005), and identify skills or need that are functional, generalizable, most likely to have the greatest impact on child outcomes, and unlikely to develop without intervention (Johnson et al., 2015; Pretti-Fontczak & Bricker, 2000).

Third, this mode of assessment must be flexible for diverse learners. Assessments should

be flexible enough to allow the team to adapt items for a range of learners. The assessment should be appropriate for and adaptable for use with culturally and linguistically diverse children and families, allowing the child to respond in his or her native language and in ways consistent with family culture.

Fourth, program planning assessment must be family friendly. Families should have meaningful and multiple options for providing information about how the child functions at home and in the community. Also, they should play a key role in making decisions about the focus of intervention efforts.

Fifth, assessment should be dynamic and repeated as needed. Children change rapidly during early childhood and some interventions may not produce desired results. Both factors make repeated administration of an assessment important in determining how to promote skill development over time.

Assessment for Intervention Monitoring

Definition

Once an intervention has been designed, ongoing assessment of its implementation must follow. Assessment for intervention monitoring is a systematic description and standard-based evaluation of implementation, to ensure that services are meeting a priori or individual standards—or that additional support is provided if the planned intervention is not carried out as intended. While great strides have been made in the field of ECSE since its inception, a large gap exists between best and actual practices (Dunlap, Hemmeter, Kaiser, & Wolery, 2011). Thus, it is important that we monitor implementation of intervention programs to ensure quality services for young children with disabilities.

It is a special educator's legal and ethical responsibility to ensure the intervention plan is carried out as specified. Presumably, children will progress better if the programs we plan are carried out as specified (Wolery, 2004). Lack of

progress may be due to “infrequent, inaccurate, or inconsistent use of interventions” (Wolery, 2004, pp. 573–574). If a child does not progress as expected, the team must ask if services are provided as frequently as the plan specified and with sufficient quality.

In practice, this is likely easier to determine for older children because recommended instructional methods are more direct and easily observable. For example, we can readily observe a teacher implementing a well-specified reading program with a second grader receiving special education. We can observe the time and pacing of instruction and whether the teacher is implementing the program with fidelity.

Observing intervention implementation in early intervention and ECSE may be more challenging when instruction is embedded within the course of daily routines and activities. Learning opportunities are distributed throughout the day, making them more difficult to record. Plus, they are often more difficult for the untrained eye to spot. For example, a teacher may withhold a block from a preschooler and give an expectant look to provide an opportunity for the child to request the block. The observer would need to know the child’s IEP goals and the procedures for using naturalistic communication strategies to notice the teacher facilitate this opportunity. These difficulties compound in home-based programs, which are often implemented by parents or other caregivers. It is challenging to know the frequency and quality of implementation between visits.

It may be challenging to document implementation in ECSE, but it is essential to do so (Fixsen, Blase, Naom, & Wallace, 2009). As noted by Bagnato et al. (2010), even the best intervention is effective only when implemented with fidelity. Assessment for intervention monitoring should follow four general principles. First, the program itself should be clearly articulated before implementation, with special attention to core ingredients. Second, data on implementation should be collected continually to ensure the program meets a priori standards. Third, implementation should closely match the original plan along key dimensions, with formal timelines for review and program revision. More frequent changes may be

based on informal or incomplete information, and thus may increase the odds that intervention is poorly matched to a child’s needs and characteristics. Fourth, the team should examine data on implementation and child progress to determine if the intervention requires changes. Monitoring implementation may help identify key reasons why a child is progressing slower than the team expected, which may then help the team modify services.

Examples

In special education, interventions are tailored specifically to meet the child’s needs and, with infants and toddlers, the family’s needs (Bagnato, McLean, Macy, & Neisworth, 2011). The nature of the program and subsequent monitoring will vary based on the child’s age and needs. In birth-to-three services, an IFSP for a 12-month-old with hearing loss and communication delays might have intervention components that require the family to place hearing aids on the child for particular intervals and that help the family embed communication opportunities into home routines to increase the child’s communication skills. In this example, the team would monitor implementation variables like frequency of home visits, frequency and quality of coaching to teach specific strategies to parents, parent logs of child use of hearing aids, and parent use of communication strategies during and between visits. These data would be reviewed regularly and frequently—perhaps every other week at first, fading to monthly reviews—to identify gaps between planned and implemented intervention; any gaps would be addressed by planned revisions to intervention supports.

For preschool and elementary students, intervention plans are described in the child’s IEP. For a 3-year-old with autism (ASD), this plan might include addressing communication, social-emotional, and adaptive skills with services provided by a co-teaching team in an inclusive preschool classroom. The plan would include specific evidence-based practices (EBPs) for teaching children with ASD (see Wong et al.,

2014) and a checklist for both type and frequency of embedded teaching opportunities to be provided in different contexts. The team would monitor implementation, including checklist recordings of frequency and fidelity of implementing specific components of EBPs. Teachers would review these data weekly and make necessary adjustments when implementation varies from the original plan.

Key Features of Quality Implementation

Quality implementation of assessment for intervention monitoring requires attention to four distinct features. First, special educators must define the intervention plan before implementation, with sufficient detail to communicate clearly its requirements and to set the occasion for thorough review of implementation. Tilly (2008, p. 21) describes a model in which practitioners identify “who will do what, when, and in what manner” as intervention unfolds. Second, special educators must have rigorous and reasonable plans for collecting data regarding the implementation of each core feature (Wolery, 2004). These data should be collected in an ongoing fashion so that any needed changes can be identified quickly, particularly when child progress is “slow, variable, or is not occurring” (Wolery, 2004, p. 579). Third, we must analyze the data collected to ensure the plan and its components are carried out with fidelity or as intended (Greenwood, Carta et al., 2011; Tilly, 2004). Best practices dictate that data describe the extent to which practitioners “implement the frequency, intensity, and duration of instruction needed to address the child’s phase and pace of learning or the level of support needed by the family to achieve the child’s outcomes or goals” (DEC, 2014, p. 10). Finally, if review of implementation data suggests significant variations from original intent, the team has a responsibility to adjust its practices to carry out the plan fully. This may include monitoring whether or not a service is provided with the specified frequency, that modifications to the environment or materials are occurring as

planned, that some core element of an EBP is provided with the specified frequency, or that an EBP is being implemented as intended.

Assessment for Progress Monitoring

Definition

Once individuals have been identified for special or supplemental instruction and intervention, and these specialized services have been planned and implemented, best practices dictate that teachers and other professionals (as well as parents and other interested parties) frequently monitor the degree to which these services or supports promote changes in desired child performance and promote progress toward identified long-term goals. Progress monitoring is frequently repeated and often rather brief. The child’s performance is in turn evaluated against an a priori standard of expected change or rate of growth. This information is used to determine whether to continue the current array and dosage of intervention services or if some change is warranted.

While progress monitoring has been a hallmark of special education for some time (c.f., White, 1986), technical features of this approach to assessment have received substantial attention since the 1990s. In particular, scholarship has identified two paradigmatically different approaches to progress monitoring—*Developmental Skills Mastery Monitoring* and *General Outcome Measurement* (Fuchs & Deno, 1991; McConnell, 2001).

Developmental skills mastery monitoring (or DSMM, termed “mastery monitoring” by Fuchs & Deno, 1991) is common in many ECSE programs. In DSMM, child progress is marked by successive mastery, or skilled performance, on a set of behaviors or items that are ordered developmentally. Mastery criteria for performance are specified and distinct measurement procedures are used for each individual behavior or item. In DSMM, accuracy or appropriateness of the skill or item hierarchy is critically important, measurement is focused and specific, and

(under many conditions) assessment can be very sensitive to short-term effects that are very specific to any given intervention. One specific example of this approach would be to assess, on a weekly or even daily basis, children's performance relative to IFSP or IEP objectives. Assessment of this type can help guide instruction on a particular skill to continue or to advance to more complex skills.

By contrast, general outcome measurement (or GOM) notes changes in performance by repeated assessment of the child in a general, or more global, task. Progress is noted by increases in proficiency (typically, either the extent to which the child completes the full task or some count of successful responses to that task); in other words, assessment describes a child's partial proficiency on a global outcome toward which development and intervention are directed. In GOM, the general or long-term outcome being assessed and the extent to which assessment opportunities sample the child's performance of this outcome are critically important. Measurement is consistent across occasions and provides an integrated view of the child's skill. GOM assessment is broad and outcome referenced, and as a result may be somewhat less sensitive to either intervention specific or small changes in child development but more related to longer-term expectations. GOMs provide metrics of both *status* (i.e., child performance at a given time) and *growth* (i.e., change in that performance across repeated assessments), which can be particular assets to ongoing progress monitoring (Deno, 1997).

Key Features of Quality Implementation

At least four key features of progress monitoring are required: two relate to application and use and two to instrument or measure selection. First, progress-monitoring measures must be administered frequently. A primary purpose for this assessment is to describe short-term effects on child development and to use this information to adjust intervention as needed. As a

result, progress must be assessed often enough to both describe rates of change in child behavior, and to prompt changes in intervention services, and thus improve long-term outcomes, as often as is possible.

Second, progress-monitoring measures must quickly and easily produce data teachers can analyze to make decisions about ongoing intervention services. This requires both that data provided by measures be accessible and easy to produce and that teachers have and employ rigorous decision-making rules to evaluate these results.

Third, progress-monitoring measures must sample behaviors or competencies that are highly related, conceptually and empirically, to interventions being provided. This is a special case of validity, as noted previously; content and construct validity *for the intervention being implemented* must be very high. Progress-monitoring measures must reveal when intervention is contributing (or not contributing) to changes in child achievement or behavior.

Fourth, progress-monitoring measures must be appropriate for, and sensitive to change in, repeated assessments. This often requires "parallel forms"—empirically equivalent test sets that rely on different items or tasks to evaluate child performance—to reduce contributions from test practice. But it also requires that the measures used be sufficiently robust and reliable (i.e., have small standard errors of measurement) to be able to detect real changes in child performance over the shortest time possible.

Examples

Developmental Skills Mastery Monitoring. DRDP access (McLean, Edelman, & Salcedo, 2011) is a DSMM approach, carefully developed to "observe, assess, and report on the development of [infants, toddlers, and preschoolers] who receive preschool special education services" (McLean et al., 2011, p. 4). *DRDP access* is a universal design adaptation of *Desired Results*, an instrument originally designed for children without disabilities; *DRDP access* provides more

fine-grained assessment of developmental skill, specifies adaptations for children with specific disabling conditions, and has been engineered specifically to report on developmental progress in 6-month intervals as required by US federal special education regulations.

DRDP access is an observational teacher rating scale with content aligned to state early learning standards. Teachers or others who have spent considerable time with an individual child review and rate the child's performance on specific measurable behaviors grouped within slightly broader domains of developmental performance (e.g., "children show growing abilities in communication and language"). *DRDP access* results are presented both descriptively (e.g., skills the child has recently mastered in each developmental domain or "desired result") and summatively, with a scaled score that indexes overall current performance.

General Outcome Measurement. Individual Growth and Development Indicators, or IGDIs, were developed to provide single-point and repeated assessment for both screening/identification and for progress monitoring. Walker, Carta, Greenwood, and Buzhardt (2008) describe the use of the Early Communication Indicator, an infant and toddler measure, for monitoring growth in proto-communication and communication among infants at risk for developmental delays. Using this measure, home visitors repeatedly monitor individual child communication behaviors in a standard play-like activity with a familiar adult and compare observed rates of gestures, vocalizations, and single- and multiple-word utterances to both broad normative standards (based on age) and individual intervention targets (based on parent and professional goal setting). When child progress falls below expected levels of development over time, intervention plans are revised or expanded and monitoring continues. This progress monitoring and intervention adjustment process serves as basis for an online professional support program (Buzhardt et al., 2011) and has been widely used with positive results in statewide home visiting efforts (Greenwood, Buzhardt, Walker, Howard, & Anderson, 2011).

Key Features of Quality Implementation

Three primary factors drive quality implementation of progress-monitoring measures. *Selection of appropriate instruments or measures* is central to positive effects from progress monitoring. The measures or instruments must meet basic standards for psychometric rigor, and they must meet demands unique to this function. In particular, progress-monitoring measures must be (a) related to long-term intervention goals; (b) appropriate for repeated use, with frequency sufficient to provide high-quality information to allow for rapid intervention review and adjustment; (c) sensitive to small changes in child performance over time; and (d) to the extent possible, produce data that teachers and parents can interpret directly and easily. Second, data on child performance must be collected frequently. At its core, progress monitoring provides information to allow rapid and ongoing assessment of intervention efficacy and to support ongoing refinement of intervention services. Finally, high-quality progress monitoring requires *thorough and reliable analysis and use of collected data*. Teachers and parents must reflect on the "moving picture" of intervention effect as this process unfolds and be prepared to revise interventions when warranted. This requires easy-to-use information from progress-monitoring measures and clear, well-implemented rules for data analysis and interpretation.

Future Directions in Assessment Development and Application

As attributed to Niels Bohr, Yogi Berra, and others, we note that it is very difficult to make predictions, particularly about the future. Nonetheless, several modest recommendations for ongoing assessment development and application may be warranted.

First, ECSE practice and policy will benefit from ongoing development and use of a broader, deeper, and more sophisticated array of assessment tools and practices. While assessment resources have certainly expanded as our system

of identification, support, and service to young children and their families has grown, this area is still in its earliest stages of development. More tools and practices, across all functions of assessment, will enable parents and practitioners to better understand the developmental course of young children, to reflect on and plan for interventions to affect that course, and to monitor the effects of these efforts at the individual, group, and population levels over time.

Similarly, we expect to see growing sophistication in the methods and analytic tools used in this research and development work. IRT and other “modern” test development approaches are only recently widely applied in education and psychology (Embretson & Reise, 2000; Wilson, 2005), and these approaches are beginning to be applied to assessing young children (Anthony et al., 2011; Bradfield, Wackerle-Hollman et al., 2014; Phillips, Piasta, Anthony, Lonigan, & Francis, 2012). The growing sophistication in design and evaluation will further support development so that one might expect more and better measures as this work continues.

Finally, technology platforms are apt to become increasingly common; they will bring faster acceleration of sophisticated and new assessment approaches. We are already seeing these effects in language development (Xu, Richards, & Gilkerson, 2014); this century’s explosion of small, affordable, and relatively powerful electronic devices can only expand this area.

Closing

Assessment practices for young children with disabilities are central to identifying children who would benefit from early intervention, the design of services and supports for them and their families and monitoring and evaluating these services over time. A strong reciprocal relation between assessments and intervention effects is clear for individuals and systems. As these practices increase, improve, and expand, this reciprocal effect will likely continue and, in the end, be a vital component for meeting obligations to children, their families, and those who serve them.

Acknowledgments Preparation of this manuscript was supported in part by grant R324C080011, the *Center for Response to Intervention in Early Childhood*, from the Institute of Education Sciences, US Department of Education, to the University of Kansas, Charles Greenwood and Judith Carta, Principal Investigators. However, the opinions and recommendations presented in this paper are those of the authors alone, and this paper does not imply official endorsement from the Institute of Education Sciences.

Dr. McConnell and his colleagues have developed assessment tools and related resources known as *Individual Growth & Development Indicators* and *Get it, Got it, Go!* This intellectual property is subject of technology commercialization and possible licensing agreements through the University of Minnesota. Dr. McConnell may be entitled to royalties for products related to the research described in this paper. The university has reviewed and managed this relationship in accordance with its conflict of interest policies.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2002). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan.
- Anthony, J. L., Williams, J. M., Durán, L. K., Gillam, S. L., Liang, L., Aghara, R., ... Landry, S. H. (2011). Spanish phonological awareness: Dimensionality and sequence of development during the preschool and kindergarten years. *Journal of Educational Psychology, 103*, 857.
- Bagnato, S. J., McLean, M., Macy, M., & Neisworth, J. T. (2011). Identifying instructional targets for early childhood via authentic assessment: Alignment of professional standards and evidence-based practices. *Journal of Early Intervention, 33*, 243–253.
- Bagnato, S. J., Neisworth, J. T., & Pretti-Frontczak, K. (2010). *LINKing authentic assessment and early childhood intervention: Best measures for best practices* (2nd ed.). Baltimore, MD: Brookes.
- Bethell, C., Reuland, C., Schor, E., Abrahms, M., & Halfon, N. (2011). Rates of parent-centered developmental screening: Disparities and links to services access. *Pediatrics, 128*, 146–155.
- Bradfield, T. A., Vue, K., Rodriguez, M., & McConnell, S. R. (2014). *Confident decision-making: Validation of an early childhood response to intervention decision-making framework*. Paper presented at the Conference for Research Innovation in Early Intervention, San Diego, CA.
- Bradfield, T. A., Wackerle-Hollman, A., Albano, A., Rodriguez, M., & McConnell, S. R. (2014). Redefining individual growth and development indicators: Oral language. *Assessment for Effective Intervention, 39*, 233–244.

- Bricker, D. (Series Ed.). (2002). *Assessment, evaluation, and programming system for infants and children* (2nd ed., Vols. 1–4). Baltimore, MD: Brookes.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513–531.
- Bruder, M. B. (2010). Early childhood intervention: A promise to children and families for their future. *Exceptional Children*, 76, 339–355.
- Buzhardt, J., Greenwood, C. R., Walker, D., Anderson, R., Howard, W. J., & Carta, J. J. (2011). Effects of web-based support on early Head Start home visitors' use of evidence-based intervention decision making and growth in children's expressive communication. *NHSA Dialog*, 13, 121–146.
- Carta, J. J., Greenwood, C. R., Goldstein, H., McConnell, S. R., Kaminski, R., Bradfield, T. A., ... Atwater, J. (2016). Advances in multi-tiered systems of support for prekindergarten children: Lessons learned from 5 years of research and development from the Center for Response to Intervention in Early Childhood. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *The handbook of Response to Intervention: The science and practice of multi-tiered systems of support* (2nd ed.). New York, NY: Springer.
- Christ, T. J., & Nelson, P. M. (2014). Developing and evaluating screening systems: Practical and psychometric considerations. In R. J. Kettler, T. A. Glover, C. A. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational settings: Evidence-based decision making for schools. School psychology book series* (pp. 79–110). Washington, DC: American Psychological Association.
- Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., ... Morrison, F. J. (2009). Individualizing student instruction precisely: Effects of child × instruction interactions on first graders' literacy development. *Child Development*, 80(1), 77–100.
- Cook, B. G., & Odom, S. L. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children*, 79, 135–144.
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York, NY: Harper & Row.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219–232.
- Deno, S. L. (1997). Whither thou goest ... perspectives on progress monitoring. In J. W. Lloyd, E. J. Kame'enui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp. 77–99). Mahwah, NJ: Erlbaum.
- Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37, 184–192.
- Division for Early Childhood. (2014). *DEC recommended practices in early intervention/early childhood special education*. Retrieved from <http://www.dec-sp.ed.org>.
- Dunlap, G., Hemmeter, M. L., Kaiser, A. P., & Wolery, M. (2011). Introduction to PL 99-457 anniversary issues. *Journal of Early Intervention*, 33, 239–242.
- Earls, M. F., & Hay, S. S. (2006). Setting the stage for success: Implementation of developmental and behavioral screening and surveillance in primary care practice—the North Carolina Assuring Better Child Health and Development (ABCD) project. *Pediatrics*, 118, e183–e188.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. New York, NY: Psychology Press.
- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, 19, 531–540.
- Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, 57, 488–500.
- Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RTI: A next-generation approach to multilevel prevention. *Exceptional Children*, 78, 263–279.
- Furuno, S., O'Reilly, K. A., Hosaka, C. M., Inatsuka, T. T., Allman, T. L., & Zeisloft, B. (1979). *Hawaii early learning profile*. Palo Alto, CA: VORT Corporation.
- Good, R. H., & Kaminski, R. (2000). *Dynamic indicators of basic early literacy skills* (DIBELS). Retrieved from <http://dibels.uoregon.edu>.
- Greenwood, C. R., Buzhardt, J., Walker, D., Howard, W. J., & Anderson, R. (2011). Program-level influences on the measurement of early communication for infants and toddlers in early Head Start. *Journal of Early Intervention*, 33, 110–134.
- Greenwood, C. R., Carta, J. J., & McConnell, S. R. (2011). Advances in measurement for universal screening and individual progress monitoring of young children. *Journal of Early Intervention*, 33, 254–267.
- Greenwood, C. R., & McConnell, S. R. (2011). JEI guidelines for manuscripts describing the development and testing of an assessment instrument or measure. *Journal of Early Intervention*, 33, 171–185.
- Grisham-Brown, J., & Pretti-Frontczak, K. (2011). *Assessing young children in inclusive settings: The blended practices approach*. Baltimore, MD: Brookes.
- Hawkins, R. P. (1979). The functions of assessment: Implications for selection and development of devices for assessing repertoires in clinical, educational, and other settings. *Journal of Applied Behavior Analysis*, 12, 501–516.
- Janus, M., Brinkman, S., Duku, E., Hertzman, C., Santos, R., Sayers, M., ... Walsh, C. (2007). *The early development instrument: A population-based measure for communities. A handbook on development, properties, and use*. Hamilton, ON: Offord Centre for Child Studies.
- Janus, M., & Offord, D. R. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science*, 39, 1–22.
- Johnson, J. J., Rahn, N. L., & Bricker, D. (2015). *Activity-based intervention: A comprehensive framework for use with young children* (4th ed.). Baltimore, MD: Brookes.

- Johnson-Martin, N., Jens, K. G., Attermeier, S. M., & Hacker, B. J. (1986). *Carolina curriculum for handicapped infants and infants at risk*. Baltimore, MD: Brookes.
- Kazdin, A. E. (1979). Situational specificity: The double-edged sword of behavioral assessment. *Journal of Behavioral Assessment, 1*, 57–75.
- King, T. M., Tandon, S. D., Macias, M. M., Healy, J. A., Duncan, P. M., Swigonski, N. L., ... Lipkin, P. H. (2010). Implementing developmental screening and referrals: Lessons learned from a national project. *Pediatrics, 125*, 350–360.
- MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55*, 95–107.
- Macy, M. M., 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*, 1–16.
- McCarthy, D. (1970). *Manual for the McCarthy scales of children's abilities*. New York, NY: The Psychological Corporation.
- McConnell, S. R. (2000). Assessment in early intervention and early childhood special education: Building on the past to project into our future. *Topics in Early Childhood Special Education, 20*, 43–48.
- McConnell, S. R. (2001). Distinguishing two approaches to assessment of infants and young toddlers. In S. McConnell (Chair), *Recent advances in developmental assessment of infants and young children: Evidence from two approaches*. Paper presented at the Society for Research in Child Development, Minneapolis, MN.
- McConnell, S. R., Wackerle-Hollman, A., & Bradfield, T. A. (2014). Early childhood literacy screening. In R. Kettler, T. Glover, C. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational settings: Identification, implications, and interpretation* (pp. 141–170). Washington, DC: American Psychological Association.
- McConnell, S. R., Wackerle-Hollman, A. K., Roloff, T. A. B., & Rodriguez, M. (2014). Designing a measurement framework for response to intervention in early childhood programs. *Journal of Early Intervention, 36*(4), 263–280.
- McLean, M., Edelman, L., & Salcedo, P. (2011). *User's guide to the DRDP access for preschool special education* (4th ed.). Sacramento, CA: California Department of Education, Special Education Division.
- McLean, M., Wolery, M., & Bailey, D. B. (2003). *Assessing infants and preschoolers with special needs* (3rd ed.). New York, NY: Prentice Hall.
- McWilliam, R. A. (2003). *RBI report form*. Center for Child Development, Vanderbilt University Medical Center, Nashville, TN. Retrieved from <http://www.siskin.org/www/docs/112.190>.
- Messick, S. (1988). *The once and future issues of validity: Assessing the meaning and consequences of measurement*. New York, NY: Erlbaum.
- Messick, S. (1990). *Validity of test interpretation and use*. Princeton, NJ: Educational Testing Service.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741–749.
- National Academy for State Health Policy. (2014). ABCD resource center. Retrieved October 3, 2014, from <http://www.nashp.org/abcd-welcome>.
- National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common core state standards*. Retrieved from <http://www.corestandards.org/>.
- Neisworth, J. T., & Bagnato, S. J. (2005). DEC recommended practices: Assessment. In S. Sandall, M. L. Hemmeter, B. J. Smith, & M. E. McLean (Eds.), *DEC recommended practices: A comprehensive guide for practice application in early intervention/early childhood special education* (pp. 45–69). Longmont, CO: Sopris West.
- Newborg, J., Stock, J. R., Wnek, L., Guidubaldi, J., & Svinicki, J. (1988). *Battelle developmental inventory*. Allen, TX: DLM (Original work published 1984).
- O'Neill, R. E., Albin, R. W., Storey, K., Horner, R. H., & Sprague, J. R. (2014). *Functional assessment and program development for problem behavior: A practical handbook* (3rd ed.). Stamford, CT: Cengage Learning.
- Pelletier, H., & Abrams, M. (2003). *ABCD: Lessons from a four-state consortium—National Academy for State Health Policy*. Portland, ME: National Academy for State Health Policy.
- Phillips, B. M., Piasta, S. B., Anthony, J. L., Lonigan, C. J., & Francis, D. J. (2012). IRTs of the ABCs: Children's letter name acquisition. *Journal of School Psychology, 50*(4), 461–481.
- Pretti-Frontczak, K., & Bricker, D. (2000). Enhancing the quality of IEP goals and objectives. *Journal of Early Intervention, 23*, 92–105.
- Ramey, C. T., & Ramey, S. L. (1998). Early intervention and early experience. *American Psychologist, 53*, 109–120.
- Rodriguez, M. C. (2010). *Building a validity framework for second-generation IGDIs*. Minneapolis, MN: Technical Reports of the Center for Response to Intervention in Early Childhood. University of Minnesota.
- Salvia, J., Ysseldyke, J. E., & Bolt, S. (2012). *Assessment in special and inclusive education* (12th ed.). Boston, MA: Houghton Mifflin.
- Shonkoff, J. P., & Meisels, S. J. (1990). Early childhood intervention: The evolution of a concept. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (pp. 3–31). Cambridge, United Kingdom: Cambridge University Press.
- Snyder, P. A., Wixson, C. S., Talapatra, D., & Roach, A. T. (2008). Assessment in early childhood: Instruction-focused strategies to support response-to-intervention frameworks. *Assessment for Effective Intervention, 34*, 25–34.
- Strain, P. S., McConnell, S. R., Carta, J. J., Fowler, S. A., Neisworth, J. T., & Wolery, M. (1992). Behaviorism in

- early intervention. *Topics in Early Childhood Special Education*, 12, 121–141.
- Terman, L. W., & Merrill, M. A. (1972). *Stanford-Binet intelligence scale, form L-M* (3rd ed.). Boston, MA: Houghton Mifflin.
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale* (4th ed.). Chicago, IL: Riverside Publishing.
- Tilly, D. W. (2002). Best practices in school psychology as a problem-solving enterprise. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp. 21–36). Bethesda, MD: National Association of School Psychologists.
- Tilly, W. D., III. (2008). The evolution of school psychology to science-based practice: Problem solving and the three-tiered model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 17–36). Bethesda, MD: National Association of School Psychologists.
- U.S. Department of Health and Human Services, Administration for Children and Families, & Office of Head Start. (2011). *The Head Start child development and early learning framework: Promoting positive outcomes in early childhood programs serving children 3–5 years old*. Retrieved from <http://eclkc.ohs.acf.hhs.gov/hslc/hs/sr/approach/cdelf>.
- Wackerle-Hollman, A., Brunner, S., Duran, L., McConnell, S., Palma, J., Kohlmeier, T., ... Rodriguez, M. (2012). *The development of early literacy skills in bilingual and Spanish-speaking preschool-age children: A literature review*. Minneapolis, MN: University of Minnesota.
- Wackerle-Hollman, A., Schmitt, B., Bradfield, T. A., Rodriguez, M., & McConnell, S. R. (2015). Redefining individual growth and development indicators: Phonological awareness. *Journal of Learning Disabilities*, 48(5), 495–510.
- Walker, D., Carta, J. J., Greenwood, C. R., & Buzhardt, J. (2008). The use of individual growth and developmental indicators for progress monitoring and intervention decision-making in early education. *Exceptionality*, 16, 33–47.
- Wechsler, D. (1967). *Manual for the Wechsler preschool and primary scale of intelligence*. New York, NY: The Psychological Corporation (Original work published 1963).
- White, O. R. (1986). Precision teaching-precision learning. *Exceptional Children*, 52(6), 522–534.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Erlbaum.
- Wolery, M. (2004). Monitoring children's progress and intervention implementation. In M. McLean, M. Wolery, & D. B. Bailey Jr. (Eds.), *Assessing infants and preschoolers with special needs* (3rd ed., pp. 545–584). Upper Saddle River, NJ: Pearson.
- Wolery, M. (2005). DEC recommended practices: Child-focused practices. In S. Sandall, M. L. Hemmeter, B. J. Smith, & M. E. McLean (Eds.), *DEC recommended practices: A comprehensive guide for practice application in early intervention/early childhood special education* (pp. 71–106). Longmont, CO: Sopris West.
- Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fetting, A., Kucharczyk, S., & Schultz, T. R. (2014). *Evidence-based practices for children, youth, and young adults with autism spectrum disorder*. Chapel Hill, NC: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.
- Xu, D., Richards, J., & Gilkerson, J. (2014). Automated analysis of child phonetic production using naturalistic recordings. *Journal of Speech, Language, and Hearing Research*, 57, 1638–1650.